Credible Mechanisms[*]

Mohammad Akbarpour[†]   Shengwu Li[‡]

June 21, 2019

Abstract

Consider an extensive-form mechanism, run by an auctioneer who communicates sequentially and privately with agents. Suppose the auctioneer can deviate from the rules provided that no single agent detects the deviation. A mechanism is *credible* if it is incentive-compatible for the auctioneer to follow the rules. We study the optimal auctions in which only winners pay, under symmetric independent private values. The first-price auction is the unique credible static mechanism. The ascending auction is the unique credible strategy-proof mechanism. These results extend naturally when we permit asymmetry and payments from losing bidders.

Keywords: Mechanism design, auction, credible, strategy-proof, sealed-bid.

JEL classification codes: D44, D47, D82

1   Introduction

The standard mechanism design paradigm assumes that the auctioneer has full commitment. She binds herself to follow the rules, and cannot deviate after observing the bids, even when it is profitable *ex post* to renege (McAfee and McMillan, 1987). This contrasts starkly with the way we model participants; incentive compatibility "requires that no one should find it profitable to "cheat," where cheating is defined as behavior that can be made to look "legal" by a misrepresentation of a participant's preferences or endowment"(Hurwicz, 1972).

In this paper, we study incentive compatibility for the auctioneer. We require that the auctioneer, having promised in advance to abide by certain rules, should not find it profitable to "cheat", where cheating is defined as behavior that can be made to look "legal" *to each participant* by misrepresenting the preferences of *the other participants*. For instance, in a second-price auction, the auctioneer can profit by exaggerating the second-highest bid. Thus, as Vickrey (1961) observes, the first-price auction is "automatically self-policing", while the second-price auction requires special arrangements that tie the auctioneer's hands.[1]

To proceed, we must choose a communication structure for the bigger game played by the bidders and the auctioneer. Clearly, if the bidders simultaneously and publicly announce their bids, then the problem is trivial, and reduces to the case of full commitment. However, simultaneous public announcements are uncommon in real-world auctions. Most bidders at high-stakes auction houses do not place bids audibly, and instead use secret signals that other bidders cannot detect. These signals "may be in the form of a wink, a nod, scratching an ear, lifting a pencil, tugging the coat of the auctioneer, or even staring into the auctioneer's eyes – all of them perfectly legal" (Cassady, 1967). Recently, many bidders have ceased to be present in the auction room at all, preferring to bid over the Internet or by telephone.[2] Christie's and Sotheby's are legally permitted to call out fake ('chandelier') bids to give the impression of higher demand; the New York Times reports that, because of this practice, "bidders have no way of knowing which offers are real".[3]

There are several reasons why real-world auctioneers accommodate private communication. First, bidders frequently desire privacy for reasons both intrinsic and strategic. A mobile operator may be unwilling to publicize its value for a band of spectrum, because its rivals will take advantage of this information. In recent spectrum auctions in Ireland, the Netherlands, Austria, and Switzerland, the auctioneer did not disclose the losing bids,

---

[1] Rothkopf et al. (1990) argue that some real-world auctioneers avoid second-price auctions because bidders fear that the auctioneer may cheat.

[2] The Wall Street Journal reports, "Many auction rooms are sparsely attended these days despite widespread interest in the items being sold, with most bids coming in online or, even more commonly, by phone". *Why auction rooms seem empty these days*, The Wall Street Journal, June 15 2014.

[3] *Genteel auction houses turning aggressive*, The New York Times, April 24 2000.

even after the auction (Dworczak, 2017).[4] Second, auctioneers want to prevent collusion. Thus, in many procurement auctions, bidders are forbidden from conferring - they must submit their bids only to the auctioneer. Third, in auctions that take place over the Internet, bidders are anonymous to each other, which prevents them from sharing information. An industry newsletter[5] for online advertising auctions reports:

> In a second-price auction, raising the price floors after the bids come in allows [online auctioneers] to make extra cash off unsuspecting buyers [. . .] This practice persists because neither the publisher nor the ad buyer has complete access to all the data involved in the transaction, so unless they get together and compare their data, publishers and buyers won't know for sure who their vendor is ripping off.

The second-price auction is incentive-compatible for the auctioneer only under strong assumptions about the communication structure, such as simultaneous public communication. In this paper, we instead assume that the auctioneer engages in sequential private communication with the bidders. This enables us to represent auction rules using the tractable and familiar machinery of extensive game forms.

Consider any *protocol*; a pair consisting of an extensive-form mechanism and a strategy profile for the agents. The auctioneer runs the mechanism as follows: Starting from the initial history, she picks up the telephone and conveys a message to the agent who is called to play (an information set), along with a set of acceptable replies (actions). The agent chooses a reply. The auctioneer keeps making telephone calls, sending messages and receiving replies, until she reaches a terminal history, whereupon she chooses the corresponding outcome and the game ends.

Suppose some utility function for the auctioneer. For instance, assume that the auctioneer wants revenue. Suppose that each agent intrinsically observes certain features of the outcome. For instance, each agent observes whether or not he wins the object, and how much he pays, but not how much other agents pay.

By participating in the protocol, each agent observes a sequence of communication between himself and the auctioneer and some features of the outcome. Even if the auctioneer deviates from her assigned strategy, agent i's observation could still have an *innocent explanation*. That is, when the auctioneer plays by the rules, there exist types for the other agents that result in that same observation for i.

Given a protocol, some deviations may be *safe*, in the sense that for every type profile, each agent's observation has an innocent explanation. That is, every observation that an

---

agent might have (under the deviation) is also an observation he might have when the auctioneer is running the mechanism. For instance, when a bidder bids $100 in a second-price auction, receives the object, and is charged $99, that observation has an innocent explanation - it could be that the second-highest value was $99. Thus, in a second-price auction, the auctioneer can safely deviate by exaggerating the second-highest bid.[6]

Instead of just choosing a different outcome, the auctioneer may also alter the way she communicates with agents. For example, consider a protocol in which the auctioneer acts as a middleman between one seller and one buyer. The seller chooses a price for the object, which the auctioneer tells to the buyer. The object is sold to the buyer at that price if and only if the buyer accepts, and the auctioneer takes a 10% commission. The auctioneer has a safe deviation - she can quote a higher price to the buyer, and pocket the difference if the buyer accepts.

A protocol is *credible* if running the mechanism is incentive-compatible for the auctioneer; that is, if the auctioneer prefers playing by the book to any safe deviation. This is a way to think about partial commitment power for any extensive-form mechanism.

The private-communication assumption is crucial for our analysis. The standard mechanism design paradigm assumes that the auctioneer has no room to misrepresent each agent's behavior. We make the opposite assumption, allowing the auctioneer to misrepresent any agent's behavior to any other agent.[7] Many real-world situations are in-between. For instance, in typical auctions for art or wine, the auctioneer reveals the clearing price but hides the identity of the winner (Ashenfelter, 1989).[8] As another example, the US National Resident Matching Program publishes aggregate statistics about the match, but does not publish information that identifies individual doctors or hospitals.[9] In general, which agents share information, and what information they share, depends on context-specific features that are outside our framework. Fully private communication is a tractable benchmark, so it is the focus of the present study.

Having defined the framework, we now consider how credibility interacts with other design features. Most real-world auctions are variations on just a few canonical formats - the first-price auction, the ascending auction, and (more recently) the second-price auction (Cassady, 1967; McAfee and McMillan, 1987; Edelman et al., 2007).[10] The first-

---

[6] An auctioneer running second-price auctions in Connecticut admitted, "After some time in the business, I ran an auction with some high mail bids from an elderly gentleman who'd been a good customer of ours and obviously trusted us. My wife Melissa, who ran the business with me, stormed into my office the day after the sale, upset that I'd used his full bid on every lot, even when it was considerably higher than the second-highest bid." (Lucking-Reiley, 2000)

[7] Section 5.1 formalizes one sense in which private communication is a worst-case scenario for auctioneer incentives.

[8] One justification keeping the winner's identity secret is that it gives bidders incentives to defect from collusive arrangements. However, publishing the clearing price does not rule out cheating, since each losing bidder may believe that someone else placed the second-highest bid.

[9] The match rules also limit the information that participants can share. 2019 Main Residency Match Participation Agreement for Applicants and Programs, Sections 4.4 and 4.6.

[10] The Dutch (descending) auction, in which the price falls until one bidder claims the object, is less

price auction is static ("sealed-bid") – each agent is called to play exactly once, and has no information about the history of play when selecting his action. This yields a substantial advantage: Sealed-bid auctions can be conducted rapidly and asynchronously, thus saving logistical costs.[11] The ascending auction is strategy-proof. Thus, it demands less strategic sophistication from bidders, and does not depend sensitively on bidders' beliefs (Wilson, 1987; Bergemann and Morris, 2005; Chung and Ely, 2007). The second-price auction is static and strategy-proof; it combines the virtues of the first-price auction and the ascending auction (Vickrey, 1961).

We study the implications of credibility in the independent private values (IPV) model (Myerson, 1981). For now, we assume that the value distributions are regular and symmetric, and restrict attention to auctions in which only winning bidders make (or receive) transfers. Under these assumptions, the second-price auction (with reserve) is the unique strategy-proof static optimal auction (Green and Laffont, 1977; Holmström, 1979; Milgrom and Segal, 2002). The second-price auction is not credible, so no optimal auction is strategy-proof, static, and credible. This raises two natural questions: Is any auction static and credible? Is any auction strategy-proof and credible?

Our first result is as follows: The first-price auction (with reserve) is the unique static credible optimal auction. This implies that, in the class of static mechanisms, we must choose between incentive-compatibility for the auctioneer and dominant strategies for the agents.

Static mechanisms include the direct revelation mechanisms, in which each agent simply reports his type. Thus, when designing credible protocols, restricting attention to revelation mechanisms loses generality. The problem is that revelation mechanisms reveal too much. For a bidder to have a dominant strategy, his payment must depend on the other bidders' types. If the auctioneer knows the entire type profile, and the winning bidder's payment depends on the other bidders' types, then the auctioneer can safely deviate to raise revenue. This makes it impossible to run a credible strategy-proof optimal auction. What happens when we look outside the class of revelation mechanisms - when we use the full richness of extensive forms to regulate who knows what, and when?

For the next result, we discretize the type space, so that optimal clock auctions can be represented as extensive form mechanisms.

Our second result is as follows: The ascending auction (with an optimal reserve) is credible. Moreover, it is the *unique* credible strategy-proof optimal auction. No other extensive forms satisfy these criteria.

Notably, this result does not use open outcry bidding to ensure good behavior by the auctioneer. Given an ascending auction with an optimal reserve, the auctioneer prefers to

---

prevalent (Krishna, 2010, p.2).

[11]Using data from U.S. Forest Service timber auctions, Athey et al. (2011) find that "sealed bid auctions attract more small bidders, shift the allocation toward these bidders, and can also generate higher revenue".
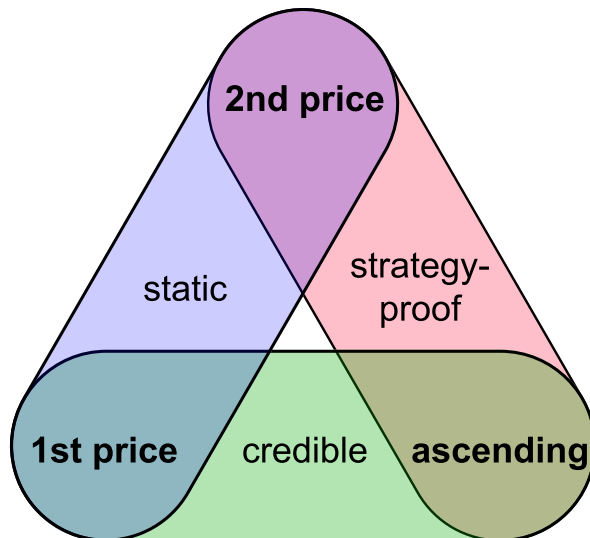
Figure 1: An auction trilemma: In the class of optimal auctions in which only the winner makes transfers, no auction is static, strategy-proof, and credible. Picking two out of three properties uniquely characterizes each standard format.

follow the rules even though she communicates with each bidder individually by telephone. If the auctioneer places chandelier bids, then she runs the risk that bidders will quit. In equilibrium, this deters her from placing chandelier bids at any price above the reserve.

These results imply an auction trilemma. Static, strategy-proof, or credible: An optimal auction can have any two of these properties, but not all three at once. Moreover, picking two out of three characterizes each of the standard auction formats (first-price, second-price, and ascending). Figure 1 illustrates.

Next, we generalize these results by relaxing the assumption that only winners make transfers and that the distributions are symmetric. The credible static auctions are now *twin-bid* auctions. This is a larger class that includes all-pay auctions and first-price auctions with entry fees. In a twin-bid auction, each agent chooses from a set of feasible bids, where a bid is a pair of numbers specifying what he pays if he wins and what he pays if he loses. After taking all bids, the auctioneer chooses a winner that maximizes revenue. Under mild assumptions, twin-bid auctions are not strategy-proof.

Under asymmetry, the static strategy-proof optimal auctions are *virtual* second-price auctions: each bid is scored as its corresponding virtual value, and the winner pays the least bid he could have reported while still having the highest score. Correspondingly, the credible strategy-proof optimal auctions are *virtual* ascending auctions: bids are scored according to their virtual values, so one bidder's price may rise faster than another's. Thus, general extensive forms enable the auctioneer to credibly reject higher bids in favor of lower bids, when it is optimal to do so.

For practical purposes, should an auction be static, strategy-proof, or credible? It depends. Some Internet advertising auctions must be conducted in milliseconds, so latency precludes the use of multi-round protocols. Strategy-proofness matters when bidders

are inexperienced or have opportunities for rent-seeking espionage. Credibility matters especially when bidders are anonymous to each other or require that their bids be kept private. These real-world concerns are outside the model. Our purpose is not to elevate some criterion as essential, but to investigate which combinations are possible.

1.1 Related work

We are far from the first to conceive of games of imperfect information as being conducted by a central mediator under private communication. Von Neumann and Morgenstern exposit such games as being run by "an umpire who supervises the course of play", conveying to each player only such information as is required by the rules (Von Neumann and Morgenstern, 1953, p. 69-84). Similarly, Myerson (1986) considers multi-stage games in which "all players communicate confidentially with the mediator, so that no player directly observes the reports or recommendations of the other players."

The papers closest to ours are Dequiedt and Martimort (2015) and Li (2017). In Dequiedt and Martimort (2015), two agents simultaneously and privately report their types to the principal, who can misrepresent each agent's report to the other agent. If we restrict attention to revelation mechanisms, then our definition of credibility is equivalent to their requirement that the principal report truthfully. However, this restriction loses some generality, so our model instead permits the auctioneer to communicate sequentially with bidders by adopting extensive-form mechanisms. Li (2017) proposes a definition of bilateral commitment power, and also introduces the messaging game that we use here. The definition in Li (2017) is restricted to dominant-strategy mechanisms, whereas credibility allows for Bayes-Nash mechanisms. Also, Li (2017) does not model the incentives faced by the auctioneer, which is the entire subject of the present study.

Our paper is related to the literature on mechanisms with imperfect commitment, in which some parts of the outcome are chosen freely by the designer after observing the agents' reports (Baliga et al., 1997; Bester and Strausz, 2000, 2001). Our paper also relates to the literature that studies multi-period auction design with limited commitment (Milgrom, 1987; McAfee and Vincent, 1997; Skreta, 2006; Liu et al., 2014; Skreta, 2015). In this paradigm, the auctioneer chooses a mechanism in each period, but cannot commit today to the mechanisms that she will choose in future. In particular, if the object remains unsold, then the auctioneer may attempt to sell the object again. Essentially, these papers have a post-auction game, and require that the auctioneer is sequentially rational. Our machinery instead permits the auctioneer to misrepresent bidders' preferences during the auction.

Some papers model auctions as bargaining games in which the auctioneer cannot commit to close a sale (McAdams and Schwarz, 2007a; Vartiainen, 2013). These papers fix a particular stage game, in which players can solicit, make, or accept offers, and study

equilibria of the repeated game. The auctioneer does not promise to obey any rules – she is constrained only by the structure of the repeated game. In our model, the auctioneer instead promises in advance to abide by certain rules, and can only deviate from those rules in ways that have innocent explanations. Thus, if the auctioneer promises to run a first-price auction, then she must conclude the auction after collecting the bids. By contrast, McAdams and Schwarz (2007a) and Vartiainen (2013) permit the auctioneer to restart play in the next period, exploiting the new information that she has learned.

Several papers study auctioneer cheating in specific auction formats, such as shill-bidding in second-price auctions (McAdams and Schwarz, 2007b; Rothkopf and Harstad, 1995; Porter and Shoham, 2005) and in ascending auctions with common values (Chakraborty and Kosmopoulou, 2004; Lamy, 2009). Loertscher and Marx (2017) allow the auctioneer to choose when to stop the clocks in a two-sided clock auction. We contribute to this literature by providing a definition of auctioneer incentive-compatibility that is not tied to a particular format, and can thus be used as a design criterion.

Our paper contributes to the line of research that studies standard auction formats by relaxing various assumptions of the benchmark model (Milgrom and Weber, 1982; Maskin and Riley, 1984; Bulow et al., 1999; Fang and Morris, 2006; Hafalir and Krishna, 2008; Bergemann et al., 2017, 2018). While the usual approach is to compare the standard formats in terms of expected revenue, we instead characterize the standard formats with a few simple desiderata. Of course, the desiderata of Figure 1 do not exhaust the considerations of real-world auctioneers; factors such as interdependent values, risk aversion, and informational robustness importantly affect the choice of format.

2 Model

2.1 Definitions

We now define the model. Proofs omitted from the main text are in Appendix B. The environment consists of:

1. A finite set of agents, N.

2. A set of outcomes, X.

3. A type space, $\Theta_N = \times_{i \in N} \Theta_i$, endowed with σ-algebra $\mathcal{F}$.

4. A probability measure $D : \mathcal{F} \to [0, 1]$.

5. Agent utilities $u_i : X \times \Theta_N \to \mathbb{R}$

6. A partition Ω_i of X for each $i \in N$. (ω_i denotes a cell of Ω_i.)

The partition Ω_i represents what agent i *directly observes* about the outcome. Conceptually, these partitions represent physical facts about the world, which are not objects of design. They capture the bare minimum that each agent observes about the outcome, regardless of the choice of mechanism.[12]

We represent the rules of the mechanism as an extensive game form with imperfect information. This specifies the information that will be provided to each agent, the choices each agent will make, and the outcomes that will result, *assuming that the auctioneer follows the rules.* Crucially, we are not yet modeling the ways that the auctioneer can deviate.

Formally, a mechanism is an extensive game form with consequences in X. This is an extensive game form for which each terminal history is associated with some outcome. Formally, a mechanism G is a tuple $(H, \prec, P, A, \mathcal{A}, (\mathcal{I}_i)_{i \in N}, g)$, where each part of the tuple is as specified in Table 1. The full definition of extensive forms is familiar to most readers, so we relegate further detail to Appendix A. We restrict attention to mechanisms with perfect recall and finite depth (that is, there exists some $K \in \mathbb{Z}$ such that no history has more than K predecessors).

Table 1: Notation for Extensive Game Forms

Name	Notation	Representative element
histories	H	h
precedence relation over histories	$\prec$	
reflexive precedence relation	$\preceq$	
initial history	$h_\emptyset$	
terminal histories	Z	z
player called to play at h	$P(h)$	
actions	A	a
most recent action at h	$\mathcal{A}(h)$	
information sets for agent i	$\mathcal{I}_i$	I_i
outcome resulting from z	$g(z)$	
immediate successors of h	$\text{succ}(h)$	
actions available at I_i	$A(I_i)$	

An **interim strategy** is a function from information sets to available actions, $\sigma_i : \mathcal{I}_i \rightarrow A$, satisfying $\sigma_i(I_i) \in A(I_i)$. Let Σ_i denote the set of i's interim strategies, and denote an interim strategy profile by $\sigma_N = (\sigma_i)_{i \in N}$. An ***ex ante* strategy** is a function from types to interim strategies, $S_i : \Theta_i \rightarrow \Sigma_i$. An *ex ante* strategy profile is $S_N = (S_i)_{i \in N}$, which implies an interim strategy profile for each type profile, $S_N(\theta_N) = (S_i(\theta_i))_{i \in N}$. We use $S_i(I_i, \theta_i)$ to denote the action played under *ex ante* strategy S_i at information set I_i by type θ_i.

---

[12]In the application that follows, we will assume that each bidder in an auction knows how much he paid and whether he receives the object. In effect, this rules out the possibility that the auctioneer could hire pickpockets to raise revenue, or sell the object to multiple bidders by producing counterfeit copies.

By convention, many papers make statements about mechanisms that implicitly refer to a particular equilibrium of the mechanism, such as the claim "second-price auctions are efficient". To reduce ambiguity, we will state our results explicitly for pairs (G, S_N) consisting of a mechanism and a strategy profile, which we refer to as a *protocol*.

Let $x^G(\sigma_N)$ denote the outcome in G, when agents play according to σ_N. Let $u_i^G(\sigma_i, \sigma_{-i}, \theta_N) \equiv u_i(x^G(\sigma_i, \sigma_{-i}), \theta_N)$.

Definition 2.1. (G, S_N) *is **Bayes incentive-compatible (BIC)** if, for all $i \in N$, for all $\theta_i \in \Theta_i$:*

$$S_i(\theta_i) \in \underset{\sigma_i}{\text{argmax}}\, \mathbb{E}_{\theta_{-i}}[u_i^G(\sigma_i, S_{-i}(\theta_{-i}), (\theta_i, \theta_{-i}))] \tag{1}$$

2.2 Pruning

At first glance, when constructing extensive-form mechanisms, it may seem important to keep track of off-path beliefs. However, if certain histories are off-path at every type profile, then we can delete those histories without altering the mechanism's incentive properties. Similarly, if an agent is called to play, but reveals no outcome-relevant information about his type, we can skip that step without undermining incentives. Thus, we restrict attention to the class of *pruned* protocols.[13] This technique allows us to remove redundant parts of the game tree, and implies cleaner definitions for the theorems that follow. In words, a pruned protocol has three properties.

1. For every history h, there exists some type profile such that h is on the path of play.

2. At every information set, there are at least two actions available (equivalently, every non-terminal history has at least two immediate successors).

3. If agent i is called to play at history h, then there are two types of i compatible with his actions so far, that could lead to different eventual outcomes.

Let $z(\sigma_N)$ denote the terminal history that results from interim strategy profile σ_N. Formally:

Definition 2.2. (G, S_N) *is **pruned** if, for any history h:*

1. *There exists θ_N such that $h \preceq z(S_N(\theta_N))$*

2. *If $h \notin Z$, then $|succ(h)| \geq 2$.*

3. *If $h \notin Z$, then for $i = P(h)$, there exist $\theta_i, \theta_i', \theta_{-i}$ such that*

   (a) *$h \prec z(S_N(\theta_i, \theta_{-i}))$*

   (b) *$h \prec z(S_N(\theta_i', \theta_{-i}))$*

---

[13] This is stronger than the definition of pruning in Li (2017), which includes only the first requirement.

(c) $x^G(S_N(\theta_i, \theta_{-i})) \neq x^G(S_N(\theta_i', \theta_{-i}))$

By the next proposition, when our concern is to construct a BIC protocol, it is without loss of generality to consider only pruned protocols.

Proposition 2.3. *If (G, S_N) is BIC, then there exists (G', S_N') such that (G', S_N') is pruned and BIC and for all θ_N, $x^{G'}(S_N'(\theta_N)) = x^G(S_N(\theta_N))$.*

Hence, from this point onwards we restrict attention to pruned (G, S_N). If the type space Θ_N is finite and the probability measure D has full support, then every information set in a pruned protocol is reached with positive probability, which implies that any Bayes-Nash equilibrium survives equilibrium refinements that restrict off-path beliefs.

2.3 A messaging game

We now explicitly model the auctioneer[14] as a player (denoted 0) . The auctioneer has utility $u_0 : X \times \Theta_N \to \mathbb{R}$.

The auctioneer promises in advance to run some protocol (G, S_N). We now describe a messaging game G^* that includes the auctioneer as a player. In G^*, the auctioneer contacts players privately and sequentially. At each step, she contacts some agent i, sending a message that corresponds to one of i's information sets in the mechanism G. Agent i replies with one of the actions available at that information set. At any step, the auctioneer can choose an outcome x and end the game. Thus, the auctioneer can deviate from G by altering the sequence of players or information sets, or by choosing different outcomes.

Formally, the messaging game generated by protocol (G, S_N) is defined as follows: Let the auctioneer's message space be $M = \cup_i \mathcal{I}_i$.

1. The auctioneer chooses to:

   (a) Either: Select outcome $x \in X$ and end the game.

   (b) Or: Go to step 2.

2. The auctioneer chooses some agent $i \in N$ and sends a message $m = I_i \in \mathcal{I}_i$.

3. Agent i privately observes message I_i and chooses reply $r \in A(I_i)$.

4. The auctioneer privately observes r.

5. Go to step 1.

---

[14]We use the term 'auctioneer' to refer to the mediator, but this could be any mediator who runs a mechanism, such as a school choice authority or the National Resident Matching Program.

There exists an auctioneer strategy in the messaging game that 'follows the rules' of the mechanism G. These rules prescribe which agents to contact, in what order, what messages to send, when to end the game, and what outcome to choose.

We use S_0^G to denote the rule-following auctioneer strategy. Formally, S_0^G is defined by the following algorithm: Initialize $\hat{h} := h_\emptyset$. At each step, if $\hat{h}$ is a terminal history in G, end the game and choose outcome $g(\hat{h})$. Else, contact agent $P(\hat{h})$ and send message $m = I_{P(\hat{h})}$ such that $\hat{h} \in I_{P(\hat{h})}$. Upon receiving reply r, update $\hat{h} := h' \in \text{succ}(\hat{h}) \mid \mathcal{A}(h') = r$, and iterate.[15]

We make a crucial restriction: The auctioneer can only deviate in ways that no agent can detect. Formally, in the messaging game, agent i observes the sequence of communication between himself and the auctioneer $(m_i^k, r_i^k)_{k=1}^{T_i}$, and directly observes some details of the outcome, as specified by the partition Ω_i. An observation for i is a tuple $((m_i^k, r_i^k)_{k=1}^{T_i}, \omega_i)$, where ω_i is the cell of Ω_i that contains the outcome.[16] Let $o_i(S_0, S_N, \theta_N)$ be i's observation when the auctioneer plays S_0, the agents play S_N, and the type profile is θ_N.

Definition 2.4. *Given some promised strategy profile (S_0, S_N), auctioneer strategy $\hat{S}_0$ is **safe** if for all agents $i \in N$ and all type profiles θ_N, there exists $\hat{\theta}_{-i}$ such that $o_i(\hat{S}_0, S_N, \theta_N) = o_i(S_0, S_N, (\theta_i, \hat{\theta}_{-i}))$. $\mathcal{S}_0^*(S_0, S_N)$ denotes the set of safe strategies.*

G^* is the messaging game restricted to $\mathcal{S}_0^*(S_0^G, S_N)$; this constrains the auctioneer to only play safe deviations from the rule-following strategy.

Definition 2.4 permits the auctioneer to deviate only if every agent's observation has an **innocent explanation**; there must exist $\hat{\theta}_{-i}$ such that i's observation is consistent the auctioneer playing S_0^G, the agents playing S_N, and the other agents' types being $\hat{\theta}_{-i}$.

Definition 2.5. *(G, S_N) is **credible** if*

$$S_0^G \in \underset{S_0 \in \mathcal{S}_0^*(S_0^G, S_N)}{\text{argmax}} \mathbb{E}_{\theta_N}[u_0(S_0, S_N, \theta_N)] \tag{2}$$

where $u_0(S_0, S_N, \theta_N)$ is the utility to the auctioneer from the outcome that results from (S_0, S_N) when the type profile is θ_N.

This parallels the definition of agent incentive compatibility in Hurwicz (1972):

> *In effect, our concept of incentive compatibility merely requires that no one should find it profitable to "cheat," where cheating is defined as behavior that*

---

[15] We have not defined S_0^G at information sets in the messaging game that are ruled out by S_0^G. Since we are not considering trembles by the auctioneer, all such strategies are outcome-equivalent, and this omission is harmless.

[16] Note the lack of calendar time: The agent observes the sequence of past communications between himself and the auctioneer, *not* a sequence of periods in which he either sees some communication or none.
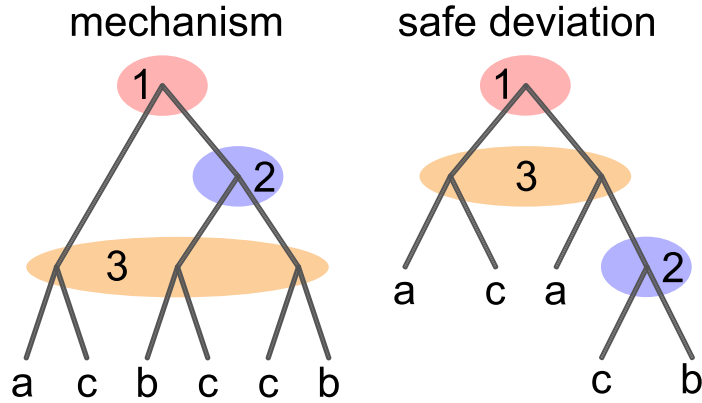
Figure 2: A mechanism and a deviation. If agent 1 cannot distinguish outcomes a and b, then the deviation is safe.

> can be made to look "legal" by a misrepresentation of a participant's preferences or endowment, with the proviso that the fictitious preferences should be within certain "plausible" limits.

In our definition, the auctioneer is allowed to behave in ways that can be made to look "legal" by misrepresenting the preferences of the other agents, with the proviso that the fictitious preferences should be within certain "plausible" limits". These limits are defined by the type space.

Instead of just choosing different outcomes, Definition 2.5 permits the auctioneer to modify G by altering the sequence of information sets. This may materially expand the auctioneer's strategic opportunities, as the following example illustrates.

Example 2.6. *Consider the mechanism on the left side of Figure 2. Each agent has one information set, two moves (left and right), and two types (l_i and r_i) that play the corresponding moves. Agent 1 is assumed to observe whether the outcome is in the set $\{a, b\}$ or in $\{c\}$. Agents 2 and 3 perfectly observe the outcome.*

The right side of Figure 2 illustrates a safe deviation: If agent 1 plays left, then the auctioneer follows the rules. If agent 1 plays right, then instead of querying agent 2, the auctioneer queries agent 3. If agent 3 then plays left, the auctioneer chooses outcome a. If agent 3 plays right, only then does the auctioneer query agent 2, choosing c if 2 plays left and b if 2 plays right.

For every type profile, each agent's observation has an innocent explanation. The most interesting case is when the type profile is (r_1, l_2, l_3). In this case, following the rules results in outcome b, but the deviation results in outcome a. Agent 1 cannot distinguish between a and b, so (l_2, l_3) is an innocent explanation for 1. (l_1, l_3) is an innocent explanation for 2, and (l_1, l_2) is an innocent explanation for 3. Thus, if the auctioneer prefers outcome a to any other outcome, then the mechanism is not credible.

Notably, this deviation involves not just choosing different outcomes, but communicating differently even before a terminal history is reached. Indeed, when the type profile is (r_1, l_2, l_3), the auctioneer can only get outcome a by deviating midway. If she waited until the end and then deviated to choose a, then agent 2's observation would not have an innocent explanation. Once agent 2 is called to play, he knows that outcome a should not occur.

Definition 2.5 takes the expectation of θ_N with respect to the *ex ante* distribution D. However, when Θ_N is finite and D has full support, Definition 2.5 implicitly requires the auctioneer to best-respond to her *updated* beliefs in the course of running G. Recall that a strategy for the auctioneer is a complete contingent plan. Suppose that in the course of running G, the auctioneer discovers new information about agents' types, such that she can profitably change her continuation strategy. There exists a deviating strategy that adopts this new course of action contingent on the auctioneer discovering this information, and plays by the rules otherwise. Thus, if S_0 is an *ex ante* best response, then its corresponding continuation strategies are also best responses along the equilibrium path-of-play.

When our concern is to construct a credible protocol, it is also without loss of generality to consider only pruned protocols.

Proposition 2.7. *If (G, S_N) is credible and BIC, then there exists (G', S_N') such that (G', S_N') is pruned, credible, and BIC, and for all for all θ_N, $x^{G'}(S_N'(\theta_N)) = x^G(S_N(\theta_N))$.*

Observation 2.8. *(G, S_N) is credible and BIC if and only if (S_0^G, S_N) is a Bayes-Nash equilibrium of G^*.*

Credibility restricts attention to 'promise-keeping' equilibria of the messaging game. However, any equilibrium can be turned into a promise-keeping equilibrium by altering the promise.

Observation 2.9. *If $S_0' \in \mathcal{S}_0^*(S_0, S_N)$, then $\mathcal{S}_0^*(S_0', S_N) \subseteq \mathcal{S}_0^*(S_0, S_N)$. Thus, if (S_0', S_N) is a Bayes-Nash equilibrium of the messaging game restricted to $\mathcal{S}_0^*(S_0, S_N)$, then it is also a Bayes-Nash equilibrium of the messaging game restricted to $\mathcal{S}_0^*(S_0', S_N)$.*

Definition 2.5 is stated for pure strategies, but can be generalized to allow the auctioneer to mix. To do so, we simply extend the definition of extensive game forms so that G includes chance moves. We then specify that $\hat{S}_0$ is **safe** if for all agents $i \in N$ and all type profiles θ_N, for any observation of agent i that occurs for some realization of the auctioneer's randomization under $(\hat{S}_0, S_N, \theta_N)$, there exists $\hat{\theta}_{-i}$ so that the same observation occurs for some realization of the auctioneer's randomization under $(S_0^G, S_N, \theta_i, \hat{\theta}_{-i})$.

In some settings, auctioneer randomization is needed to deliver the right incentives for the agents. However, randomization does not improve auctioneer incentives: We cannot

construct a credible protocol (G, S_N) by randomizing over deterministic non-credible protocols. Given randomized (G, S_N), let (G', S_N) be a deterministic protocol in which we fix a particular realization of the auctioneer's randomization. Suppose (G, S_N) is credible, so the auctioneer is indifferent between S_0^G and $S_0^{G'}$. Switching from G to G' shrinks the set of innocent explanations, and therefore the set of safe deviations. The auctioneer preferred S_0^G to any safe deviation in the larger set, and therefore prefers $S_0^{G'}$ to any safe deviation in the smaller set, so (G', S_N) is credible.

In the settings we are about to consider, randomization is not helpful for agent incentives.[17] Thus, we will restrict attention to deterministic protocols.

3 Credible Optimal Auctions

We now study credible auctions in the independent private values (IPV) model (Myerson, 1981). We make this choice for two reasons: Firstly, this is a benchmark model in auction theory, so using it shows that the results are driven by credibility, and not by some hidden feature of an unusual model.[18] Secondly, in the symmetric IPV model, revenue equivalence implies that the standard auctions start on an equal footing – the value distribution does not tip the scales in favor of a particular format, unlike the model with affiliated signals (Milgrom and Weber, 1982) or the model with risk aversion (Maskin and Riley, 1984).

Assume there are at least two bidders. An outcome $x = (y, t_N)$ consists of a winner $y \in N \cup \{0\}$ and a profile of payments (one for each bidder) $t_N \in \mathbb{R}^{|N|}$, so $X = (N \cup \{0\}) \times \mathbb{R}^{|N|}$.

Agents have private values, that is:

$$u_i((y, t_N), \theta_N) = 1_{i=y}(\theta_i) - t_i \tag{3}$$

Ω_i is as follows: Each bidder observes whether he wins the object and observes his own payment. That is, $(y, t_N), (y', t'_N) \in \omega_i$ if and only if:

1. Either: $y \neq i$, $y' \neq i$, and $t_i = t'_i$

2. Or: $y = y' = i$ and $t_i = t'_i$.

The auctioneer desires revenue, and her value for the object is normalized to zero[19]:

$$u_0((y, t_N), \theta_N) = \sum_{i \in N} t_i \tag{4}$$

---

[17]In auctions with independent private values, there always exists a deterministic mechanism that maximizes expected revenue. For instance, we can run a second-price auction that scores bids according to their ironed virtual value, breaking ties deterministically (Myerson, 1981).

[18]As Brooks and Du (2018) observe, "The IPV model has been broadly accepted as a useful benchmark when values are private, but there is no comparably canonical model when values are common."

[19]The results that follow would require only small modifications if the auctioneer's payoff was a weighted average of revenue and social welfare.

An allocation rule is a function $\tilde{y} : \Theta_N \to N \cup \{0\}$, and a transfer rule is a function $\tilde{t}_N : \Theta_N \to \mathbb{R}^{|N|}$. For any protocol (G, S_N), we can consider its induced allocation rule and transfer rule $(\tilde{y}^{G,S_N}(\cdot), \tilde{t}_N^{G,S_N}(\cdot))$. Where it is clear, we suppress the dependence on (G, S_N) to ease notation.

Let $\pi(G, S_N) = \mathbf{E}_{\theta_N}[\sum_{i \in N} \tilde{t}_i^{G,S_N}(\theta_N)]$ denote the expected revenue of (G, S_N). We will specify the relevant distribution shortly.

Definition 3.1. (G, S_N) *is* ***optimal*** *if it maximizes* $\pi(G, S_N)$ *subject to the constraints:*

1. *Incentive compatibility:* (G, S_N) *is BIC.*

2. *Voluntary participation: For all* i, *there exists* σ_i' *that ensures that* i *does not win and has a zero net transfer, regardless of* σ_{-i}'.[20]

3.1 Credible static optimal auctions

We now characterize credible static optimal auctions. Assume that $\Theta_i = [0, 1]$ and that θ_i is independently distributed according to continuous full-support density $f_i : [0, 1] \to \mathbb{R}$.

We restrict attention to protocols such that:

1. For all θ_N, $\tilde{y}(\cdot)$ and $\tilde{t}_i(\cdot)$ are measurable functions (with respect to the Borel σ-algebra on Θ_N).[21]

2. For all θ_i, $\tilde{y}(\theta_i, \cdot)$ and $\tilde{t}_i(\theta_i, \cdot)$ are measurable functions (with respect to the Borel σ-algebra on Θ_{-i}).

3. For all θ_{-i}, $\tilde{y}(\cdot, \theta_{-i})$ and $\tilde{t}_i(\cdot, \theta_{-i})$ are measurable functions (with respect to the Borel σ-algebra on Θ_i).

These conditions ensure that expected transfers and allocations are well-defined, both *ex ante* and *interim*. These are implicit in almost all papers with continuum type spaces and transferable utility. We make these restrictions explicit because the proof of Theorem 3.3 runs into some measure-theoretic subtleties.[22]

Definition 3.2. (G, S_N) *is* ***static*** *if for each agent* i, i *has exactly one information set, and for every terminal history* z, *there exists* $h \prec z$ *such that* $P(h) = i$.

---

[20]There are several standard ways of defining participation constraints, not entirely equivalent for our purposes. This definition appears in Maskin and Riley (1984). The existence of this non-participating strategy is used in the proof of Proposition B.1.

[21]That is, for any $J \subseteq N \cup \{0\}$, its preimage $\{\theta_N \mid \tilde{y}(\theta_N) \in J\}$ is a Borel set, and for any Borel set $J \subseteq \mathbb{R}$, its preimage $\{\theta_N \mid \tilde{t}_i(\theta_N) \in J\}$ is a Borel set.

[22]The restrictions rule out, for instance, that we can fix a Vitali set V, and specify that agent 1 has transfer 1 if $\theta_1 \in V$ and $\theta_2 = .5$, and transfer 0 otherwise, in which case 1's expected transfer conditional on $\theta_2 = .5$ is not defined since V is not measurable.

Next, we prove that, in a credible static auction, the winner makes a payment that essentially depends only on his own type. Thus, we can regard each agent as placing bids, with the assurance that if he wins the object, he pays exactly his bid.

Theorem 3.3 (pay-as-bid)**.** *If (G, S_N) is credible and static, then for each agent i, there exists a function $\tilde{b}_i : \Theta_i \to \mathbb{R}$ such that almost everywhere in Θ_N, if $\tilde{y}(\theta_i, \theta_{-i}) = i$ then $\tilde{t}_i(\theta_i, \theta_{-i}) = \tilde{b}_i(\theta_i)$.*

Proof. If the pay-as-bid property does not hold, then we can construct a safe deviation that raises payments on a positive-measure set. However, we cannot simply charge the 'highest safe payment' point-by-point, because there may be uncountably many opponent type profiles consistent with i winning the object, and the pointwise supremum of an uncountable family of measurable functions may not be measurable.

Lemma 3.4 (Hajłasz and Malý (2002)[23])**.** *Let Φ be a family of measurable functions defined on a set $E \subseteq \mathbb{R}^n$. There exists a countable subfamily $\hat{\Phi} \subseteq \Phi$ such that for all $\phi \in \Phi$, $\sup \hat{\Phi} \geq \phi$ almost everywhere.*

Consequently, let $(\theta_{-i}^k)_{k=1}^\infty$ be a countable subset of opponent type profiles, such that for all θ_{-i}, $\sup_k \tilde{t}_i(\cdot, \theta_{-i}^k) \geq \tilde{t}_i(\cdot, \theta_{-i})$ almost everywhere in Θ_i. $\sup_k \tilde{t}_i(\cdot, \theta_{-i}^k)$ is measurable.

We assert that $\tilde{b}_i(\cdot) = \sup_k \tilde{t}_i(\cdot, \theta_{-i}^k)$. Suppose the set

$$\{\theta_N \mid \tilde{y}(\theta_N) = i \text{ and } \tilde{t}_i(\theta_i, \theta_{-i}) \neq \sup_k \tilde{t}_i(\theta_i, \theta_{-i}^k)\} \tag{5}$$

has positive measure. Then the set

$$Q = \{\theta_N \mid \tilde{y}(\theta_N) = i \text{ and } \tilde{t}_i(\theta_i, \theta_{-i}) < \sup_k \tilde{t}_i(\theta_i, \theta_{-i}^k)\} \tag{6}$$

has positive measure. Since transfers and allocations can only change when the action profile changes, Q is measurable with respect to the equilibrium action profiles.

We now construct a safe deviation: Fix some finite K. If the agents' chosen actions are consistent with any type profile $(\theta_i, \theta_{-i}) \in Q$, then instead charge agent i $\max_{k \leq K} \tilde{t}_i(\theta_i, \theta_{-i}^k)$. Let k^* denote the arg max. If $\tilde{y}(\theta_i, \theta_{-i}^{k^*}) = i$, then allocate the object to i; else keep the object. Otherwise, play according to S_0^G. This deviation takes the maximum of finitely many measurable functions, so the resulting transfer $\tilde{t}_i^K : \Theta_N \to \mathbb{R}$ is measurable.

For K large enough, this deviation is profitable. In particular, for any $(\theta_i, \theta_{-i}) \in Q$, $\tilde{t}_i^K(\theta_i, \theta_{-i})$ is non-decreasing in K and converges as $K \to \infty$ to $\sup_k \tilde{t}_i(\theta_i, \theta_{-i}^k)$. Thus, by

---

[23]Lemma 2.6 in Hajłasz and Malý (2002), which is a special case of Lemma 2.6.1 in Meyer-Nieberg (1991).

the monotone convergence theorem,

$$\lim_{K \to \infty} \mathbb{E}_{\theta_N} \left[\tilde{t}_i^K(\theta_N) \mid \theta_N \in Q \right] = \mathbb{E}_{\theta_N} \left[\sup_k \tilde{t}_i(\theta_i, \theta_{-i}^k) \mid \theta_N \in Q \right] \tag{7}$$
$$> \mathbb{E}_{\theta_N} \left[\tilde{t}_i(\theta_N) \mid \theta_N \in Q \right]$$

which completes the proof. $\qquad\square$

Definition 3.5. (G, S_N) *is a **first-price auction** if (G, S_N) is static, and each agent i either chooses a bid b_i from a set $B_i \subset \mathbb{R}_0^+$ or declines to bid, such that:*

1. *Each bidder i pays b_i if he wins and 0 if he loses.*

2. *If any bidder places a bid, then some maximal bidder wins the object. Otherwise, no bidder wins.*

*If clauses 1 and 2 hold almost everywhere in Θ_N, then (G, S_N) is **a first-price auction almost everywhere**.*

We represent a reserve price by restricting the set B_i.

For the next theorem, we assume that the distributions are symmetric, *i.e.* $f_i(\cdot) = f_j(\cdot)$ for all i, j, and regular, *i.e.* $\theta_i - \frac{1 - F_i(\theta_i)}{f_i(\theta_i)}$ is strictly increasing. We also restrict attention to winner-paying protocols.

Definition 3.6. (G, S_N) *is **winner-paying** if, for all θ_N, if $\tilde{t}_i(\theta_N) \neq 0$ then $\tilde{y}(\theta_N) = i$.*

Theorem 3.7. *Assume the distributions are symmetric and regular. Assume (G, S_N) is winner-paying and optimal. If (G, S_N) is a first-price auction, then (G, S_N) is credible and static. If (G, S_N) is credible and static, then (G, S_N) is a first-price auction almost everywhere.*

Proof. Suppose (G, S_N) is a first-price auction. (G, S_N) is static by definition. Every safe deviation that sells the object involves charging some bidder his bid, so no safe deviation yields more revenue than following the rules. Thus, (G, S_N) is credible.

Suppose (G, S_N) is credible and static. By Theorem 3.3, there exists a function $\tilde{b}_i : \Theta_i \to \mathbb{R}$ such that, almost everywhere in Θ_N, if type θ_i wins, then i pays $\tilde{b}_i(\theta_i)$. (G, S_N) is optimal, so the participation constraint of the lowest type binds, and we can pick a non-negative function $\tilde{b}_i : \Theta_i \to \mathbb{R}_0^+$. We now partition i's actions into bidding actions $B_i = \{\tilde{b}_i(\theta_i) \mid \theta_i \in \Theta_i \text{ and } \exists \theta_{-i} : \tilde{y}(\theta_i, \theta_{-i}) = i\}$, and actions that decline. (G, S_N) is winner-paying, so Clause 1 of Definition 3.5 holds almost everywhere.

(G, S_N) is optimal, which determines the allocation rule and i's interim expected transfer almost everywhere (Myerson, 1981). By BIC, i's interim expected transfer is increasing in i's type, and the distributions are symmetric and regular, so almost everywhere the winner has a maximal type, and thus a maximal bid. Thus, Clause 2 of Definition 3.5 holds almost everywhere. Thus (G, S_N) is a first-price auction almost everywhere. $\qquad\square$

We now relax the assumption that the distributions are symmetric and regular, and that the protocol is winner-paying and optimal. In particular, rather than requiring that the protocol be optimal, we will require that, with probability 1, no bidder knows at the interim stage that he will win for sure.

Definition 3.8. (G, S_N) *is* ***contestable*** *if, almost everywhere in Θ_N, if $\tilde{y}(\theta_i, \theta_{-i}) = i$, then there exists θ'_{-i} such that $\tilde{y}(\theta_i, \theta'_{-i}) \neq i$.*

Since $\Theta_i = \Theta_j = [0, 1]$, optimal auctions are contestable.

The first-price auctions of Theorem 3.7 generalize to a larger class that permits the auctioneer to extract transfers from losing bidders, though each losing bidder's transfer must depend only on his own bid.

Whether this class is of more than technical interest will vary from case to case. Most economically important auctions, such as those for art, for mineral rights, for spectrum, or for online advertising, do not extract payments from losing bidders. Some real-world auctions may need to respect *ex post* individual rationality, since otherwise one party will try to annul the contract afterwards. The resulting transaction costs may constrain the auctioneer to use winner-paying protocols.

We now state the definition that generalizes first-price auctions.

Definition 3.9. (G, S_N) *is a* ***twin-bid auction*** *if (G, S_N) is static, and each agent chooses a two-dimensional bid (b_i^W, b_i^L) from a set $B_i \subset \mathbb{R}^2$ such that:*

1. *Each bidder i pays b_i^W if he wins and b_i^L if he loses.*

2. *If any agent places a bid such that $b_i^W - b_i^L > 0$, then some agent wins the object.*

3. *If i wins the object, then $b_i^W - b_i^L \geq \max\{0, \max_{j \neq i} b_j^W - b_j^L\}$.*

If clauses 1, 2, and 3 hold almost everywhere in Θ_N, then (G, S_N) is ***a twin-bid auction almost everywhere***.

Twin-bid auctions include first-price auctions and all-pay auctions, though the credibility of all-pay auctions is sensitive to the assumption that the object is costless to provide. (More generally, $b_i^W - b_i^L$ must be no less than the auctioneer's cost of provision, which rules out standard all-pay auctions.[24]) Twin-bid auctions also encompass first-price auctions with entry fees (b_i^L is the entry free), and first-price auctions in which losing bidders are paid fixed compensation ($b_i^L < 0$). Bidders who place higher bids may also receive more compensation if they lose; under the assumptions of Maskin and Riley (1984), this is the form of the optimal auction for symmetric bidders with constant absolute risk aversion.[25]

---

[24]The case when $b_i^W - b_i^L$ is exactly equal to the cost of provision is studied in Dequiedt and Martimort (2006), an early draft of Dequiedt and Martimort (2015).

[25]Theorem 14 (Maskin and Riley, 1984, p. 1506-1507). This claim follows from their Equations 75 and 77, since μ is non-decreasing.

Theorem 3.10. *Assume (G, S_N) is contestable. If (G, S_N) is a twin-bid auction, then (G, S_N) is credible and static. If (G, S_N) is credible and static, then (G, S_N) is a twin-bid auction almost everywhere.*

The proof of Theorem 3.10 does not rely on independence, so the characterization holds even with correlated types. Twin-bid auctions are not strategy-proof, except in degenerate cases.

Definition 3.11. *(G, S_N) is **strategy-proof** if, for all $i \in N$, for all S'_{-i}, for all $\theta_i \in \Theta_i$:*

$$S_i(\theta_i) \in \underset{\sigma_i}{\operatorname{argmax}} \, \mathbb{E}_{\theta_{-i}}[u_i^G(\sigma_i, S'_{-i}(\theta_{-i}), (\theta_i, \theta_{-i}))] \tag{8}$$

The definition above requires that S_i is a best response to all S'_{-i}, taking the expectation with respect to θ_{-i}. It is natural to consider a stronger definition that requires S_i to be a best response to all S'_{-i} and all θ'_{-i}. Under private values these definitions are equivalent.

Proposition 3.12. *Let (G, S_N) be such that there exist $\theta_i < \theta'_i < \theta''_i < \theta'''_i$, θ_{-i}, and θ'_{-i} such that $\tilde{y}(\theta_i, \theta_{-i}) \neq i = \tilde{y}(\theta'_i, \theta_{-i})$ and $\tilde{y}(\theta''_i, \theta'_{-i}) \neq i = \tilde{y}(\theta'''_i, \theta'_{-i})$. If (G, S_N) is a twin-bid auction, then (G, S_N) is not strategy-proof.*

What happens to Theorem 3.10 if we remove the assumption that the protocol is contestable? In that case, then some bidder i could have actions that win the object for sure, even when the difference $b_i^W - b_i^L$ is not high enough to satisfy Clause 3 of definition 3.9. Since there is only one object for sale, at most one bidder can have incontestable actions. The characterization of credible static mechanisms is otherwise unchanged. We omit the proof, since it is an easy modification of the proof of Theorem 3.10.

3.2 Credible and strategy-proof optimal auctions

We now characterize credible strategy-proof optimal auctions. In particular, we will show that certain ascending auctions are credible and strategy-proof.

We must make a modeling choice, because ascending auctions with discrete steps are not optimal for continuum type spaces. We could proceed by introducing a model for continuous-time auctions, as in Milgrom and Weber (1982). However, we wish to argue that credibility and strategy-proofness select ascending auctions out of a general class, and there is not yet any theory of continuous-time games that rivals the generality of extensive-form games.[26]

Consequently, our approach is to discretize the type space, so that clock auctions (and many other dynamic protocols) can be optimal. Let $\Theta_i = \{\theta_i^1, \ldots, \theta_i^{K_i}\}$. Each

---

[26]For an explanation of some difficulties involved in continuous-time game theory, see Simon and Stinchcombe (1989).

type is associated with a real number $v(\theta_i^k)$. Assume $v(\theta_i^1) = 0$, $v(\theta_i^{K_i}) = 1$ and that $v(\theta_i^{k+1}) - v(\theta_i^k) > 0$. We will abuse notation slightly, and use θ_i^k to refer both to i's kth type, and to the real number associated with that type.

Types are independently distributed, with probability mass function $f_i : \Theta_i \to (0, 1]$ and corresponding $F_i(\theta_i^k) = \sum_{l=1}^k f_i(\theta_i^l)$.

The virtual values machinery in Myerson (1981) applies *mutatis mutandis* to the discrete setting.

Definition 3.13. *For each k, we define the **virtual value** of θ_i^k to be:*[27]

$$\eta_i(\theta_i^k) \equiv \theta_i^k - \frac{1 - F_i(\theta_i^k)}{f_i(\theta_i^k)}(\theta_i^{k+1} - \theta_i^k) \tag{9}$$

$F_N = (F_i)_{i \in N}$ *is **regular** if, for all i, $\eta_i(\cdot)$ is strictly increasing.*

Optimal auctions have a characterization in terms of virtual values when certain constraints bind. $\check{u}_i^{G,S_N}(k, k')$ denotes the expected utility of agent i when his type is θ_i^k and he plays as though his type is $\theta_i^{k'}$. $\tilde{y}(\theta_N)$ denotes the allocation at type profile θ_N.

Proposition 3.14. *(Elkind, 2007) Assume F_N is regular and (G, S_N) satisfies the constraints in Definition 3.1. (G, S_N) is optimal if and only if:*

1. *Participation constraints bind for the lowest types. $\forall i : \check{u}_i^{G,S_N}(1, 1) = 0$*

2. *Incentive constraints bind locally downward. $\forall i : \forall k \geq 2 : \check{u}_i^{G,S_N}(k, k) = \check{u}_i^{G,S_N}(k, k-1)$*

3. *The allocation maximizes virtual value. $\forall \theta_N$:*

   (a) *If $\max_i \eta_i(\theta_i) > 0$, then $\tilde{y}(\theta_N) \in \operatorname{argmax}_i \eta_i(\theta_i)$.*

   (b) *If $\eta_i(\theta_i) < 0$, then $i \neq \tilde{y}(\theta_N)$.*

Ties occur with positive probability under discrete type spaces, although the probability goes to 0 as we make the discretization finer. For convenience, we will assume that the protocol breaks ties deterministically according to a fixed priority order.

Definition 3.15. *Consider a strict total order $\rhd$ on N. This generates a strict total order on all agent types, as follows: $\theta_i \rhd \theta_j$ if and only if $\theta_i \geq \theta_j$ and either $\theta_i > \theta_j$ or $i \rhd j$. We also include a reserve ρ in this total order: $\theta_i \rhd \rho$ if and only if $\theta_i \geq \rho$. We use $\overset{\rhd}{\min}$ to denote the minimum of a set with respect to this $\rhd$, and $\overset{\rhd}{\max}$ similarly.*

*(G, S_N) is **orderly** if, for some strict total order $\rhd$ on N and some reserve price ρ, i wins the object if and only if $\theta_i \rhd \overset{\rhd}{\underset{j \neq i}{\max}} \theta_j$ and $\theta_i \rhd \rho$.*

---

[27]Since $1 - F_i(\theta_i)$ is equal to 0 at the upper bound, we can define $\theta_i^{K_i+1}$ arbitrarily for the purposes of Equation 9.

We now characterize credible and strategy-proof optimal auctions.

Definition 3.16. (G, S_N) is an **ascending auction** *(with reserve price ρ) if:*

1. *All bidders start as **active**, with initial bids $(b_i)_{i \in N} := (\theta_i^1)_{i \in N}$.*

2. *The **high bidder** is the active bidder with the highest bid that is weakly above ρ (breaking ties according to $\rhd$).*

3. *At each non-terminal history, some active bidder i (other than the high bidder) is called to play, and he chooses between actions that place a **bid** in Θ_i and actions that **quit**. Bidder i knows whether an action quits, and knows the bid associated with each non-quitting action.*

   (a) *Each bid is no less than the last bid that i placed.*

   (b) *Each bid is no more than is necessary for i to become the high bidder.*

   (c) *If i quits, then he is no longer active.*

   (d) *At each information set, there is a unique action that places a bid, with one exception: If the reserve has not yet been met, and there is exactly one active bidder left, there may be multiple actions that place bids.*[28]

4. *i's strategy specifies:*

   (a) *If i's type is strictly below a bid, he does not place that bid.*

   (b) *If i's type is weakly above ρ and there is no high bidder, he places a bid.*

   (c) *If i's type is above the current high bid (breaking ties with $\rhd$), he places a bid.*[29]

5. *The auction ends if one of three conditions obtains:*

   (a) *If there are no active bidders. In that case, the object is not sold.*

   (b) *If only the high bidder is active. In that case, the object is sold to the high bidder at his last bid.*

   (c) *If the high bidder has bid θ_i^K, and no active bidder has higher tie-breaking priority. In that case, the object is sold to the high bidder at his last bid.*

In stating Definition 3.16, we have deliberately omitted what each bidder is told about the other bidders. The protocol could require that each bidder is informed about the

---

[28]This exception is here because we will shortly state a characterization theorem. If there is exactly one bidder left and the reserve has not been met, then it is as though that bidder simply faces a posted price equal to the reserve. Provided that bidder knows that he wins for sure if he bids the reserve, distinct types above the reserve can take distinct actions without allowing the auctioneer to profitably deviate.

[29]Notice that, since i's strategy must be measurable with respect to i's information sets, this implies that if i's type is above the *least possible high bid associated with that information set*, he places a bid.

number of active bidders or the identities of the active bidders. The protocol could specify that each bidder places an increasing sequence of bids, receiving no other information until he quits or is the last bidder left. These all count as ascending auctions for the purposes of the definition. We just require that each bidder's strategy satisfies Clause 4 of Definition 3.16.

Observation 3.17. *If the type distributions are regular and symmetric, then there exists an optimal ascending auction. In any ascending auction, participation constraints bind for the lowest types and incentive constraints bind locally downward. Given an optimal reserve $\rho^* = \min_k \theta_i^k \mid \eta_i(\theta_i^k) > 0$, the ascending auction maximizes the virtual value of the winning bidder. By Proposition 3.14, such an auction is optimal.*

Notably, optimality and strategy-proofness together imply that the protocol is winner-paying. Thus, we do not need to make that assumption separately in the results that follow.

The definition of extensive-form mechanisms permits the auctioneer to communicate with agents in any order, to convey information to the agent called to play, and to ask that agent to report any partition of his type space. Thus, there are many optimal auctions. However, the optimal auctions that are credible and strategy-proof are exactly the ascending auctions. To be precise:

Theorem 3.18. *Assume the type distributions are regular and symmetric and (G, S_N) is orderly and optimal. (G, S_N) is credible and strategy-proof if and only if (G, S_N) is an ascending auction.*

Proof overview. Suppose (G, S_N) is credible and strategy-proof. To prove that (G, S_N) is an ascending auction, we must show that for *any* extensive form that is not an ascending auction, there exists a profitable safe deviation for the auctioneer. Fix a protocol and a history h where agent i is called to play. Consider the types θ_i consistent with h, such that there exists θ_{-i} consistent with h, such that i wins at (θ_i, θ_{-i}). A key feature of ascending auctions is that, at each history, these types pool on the same action, unless every other agent has quit. This is stated precisely in Proposition B.8, and is closely related to unconditional winner privacy as defined by Milgrom and Segal (2017). If at some history these types do not pool, then the auctioneer can exploit one type by deviating to charge him a higher price. In the case of a second-price auction, the auctioneer simply exaggerates the value of the second-highest bid. In general, however, the deviation must be more subtle in order to be safe - instead of just choosing a different outcome, the auctioneer may systematically misrepresent agents' actions midway through the extensive form. We construct an algorithm that produces a profitable safe deviation for any such extensive form.

Suppose (G, S_N) is an ascending auction. By inspection, it is strategy-proof. What remains is to show that it is credible. Suppose that the auctioneer has a profitable safe

deviation. For every agent i, S_i remains a best response to any safe deviation by the auctioneer. Thus, since the auctioneer has a profitable safe deviation, she can openly commit to that deviation without altering the agents' incentives - we can define a new protocol (G', S'_N) that is BIC and yields strictly more expected revenue than (G, S_N). But (G, S_N) is optimal, a contradiction. (The full proof is in the Appendix.) $\qquad \square$

By Theorem 3.7, restricting attention to revelation mechanisms forces a sharp choice between incentives for the auctioneer and strategy-proofness for the agents. Theorem 3.18 shows that allowing other extensive forms relaxes this trade-off.

The characterization in Theorem 3.18 assumes optimality. This is not just a feature of our proof technique: the ascending auction is credible because it is optimal. If the reserve price is below-optimal, then the auctioneer could profitably deviate by chandelier bidding up to the optimal reserve. If the type distributions are asymmetric, then the auctioneer may profitably deviate by enforcing bidder-specific reserve prices.[30] We characterize the asymmetric case in Theorem 3.22.

While first-price auctions and ascending auctions seem to be disparate formats, they share a common feature. In both formats, if an agent might win the auction without being called to play again, then that agent knows exactly how much he will pay for the object. Thus, we can regard each agent as placing bids in the course of the auction, with the assurance that if he wins without further intervention, he will pay his bid. This 'pay-as-bid' feature is shared by all credible auctions:

Theorem 3.19 (extensive pay-as-bid). *Assume (G, S_N) is credible. Suppose that, with positive probability, i is called to play at information set I_i, takes some action a, and wins without being called to play again. Conditional on that event, there is a price $t_i(I_i, a)$ that i will pay with probability 1.*

Proof. Suppose that the event obtains, and there are two distinct prices $t_i < t'_i$, such that i pays each with positive conditional probability. The auctioneer has a profitable safe deviation: when i is meant to pay t_i, she can deviate to charge t'_i, so (G, S_N) is not credible. $\qquad \square$

Theorem 3.19 provides a consideration in favor of multi-stage auctions. Suppose we wish to have bidder i's payment depend on bidder j's private information. In order for the auction to be credible, bidder i must place a bid that incorporates that information, which requires i to learn that information during the auction. The converse of Theorem 3.19 is not true. For a counterexample, consider a 'pay-as-bid' static auction that allocates the object to the bidder who placed the second-highest bid.

---

[30]Symmetric beliefs may seem like a knife-edge case. However, in some real-world auctions, strong bidders can mask their identities and bid through proxies so as to avoid discriminatory pricing. When faced with anonymous bidders, it is quite reasonable for auctioneers to hold symmetric beliefs.

Theorem 3.18 assumed that the distribution was symmetric; we now state a version that allows asymmetry. To proceed, we define a technical condition on the distribution. Clause 1 and 2 of the following definition require that the distribution is generic, which removes distractions from tie-breaking. Clause 3 states that for any $\eta_i(\theta_i')$ in the interior of the convex hull of $\eta_j(\Theta_j)$, we can find θ_j with virtual value 'just below' $\eta_i(\theta_i')$. This is implied by continuum type spaces and continuous densities, but must be assumed separately for finite type spaces.

Definition 3.20. F_N is **interleaved** if, $\forall i \neq j :$

1. $\forall \theta_i, \theta_j : \eta_i(\theta_i) \neq \eta_j(\theta_j)$

2. $\forall \theta_i : \eta_i(\theta_i) \neq 0$

3. $\forall \theta_i, \theta_i' :$ if $\eta_i(\theta_i) < \eta_i(\theta_i')$ and $\eta_j(\theta_j^1) < \eta_i(\theta_i') < \eta_j(\theta_j^{K_j})$, then $\exists \theta_j : \eta_i(\theta_i) < \eta_j(\theta_j) < \eta_i(\theta_i')$.

Under asymmetry, we can construct an optimal auction by modifying the ascending auction to score bids according to their corresponding virtual values, and to sell only when the high bidder's virtual value is positive.

Definition 3.21. (G, S_N) is a **virtual ascending auction** if:

1. All bidders start as active, with initial bids $(b_i)_{i \in N} := (\theta_i^1)_{i \in N}$.

2. If $\max_i \eta_i(b_i) > 0$, the high bidder is $\operatorname{argmax}_i \eta_i(b_i)$. Otherwise there is no high bidder.

3. At each non-terminal history, some active bidder i (other than the high bidder) is called to play, and he chooses between actions that place a bid $b_i \in \Theta_i$ and actions that quit. Bidder i knows whether an action quits, and knows the bid associated with each non-quitting action.

    (a) Each bid is no less than the last bid that i placed.

    (b) Each bid is no more than is necessary for i to become the high bidder.

    (c) If i quits, then he is no longer active.

    (d) At each information set, there is a unique action that places a bid, with one exception: If $\max_i \eta_i(b_i) < 0$ and there is exactly one active bidder left, there may be multiple actions that place bids.

4. i's strategy specifies:

    (a) If i's type is strictly below a bid, he does not place that bid.

    (b) If $\eta_i(\theta_i) > \max\{0, \max_{j \neq i} \eta_j(b_j)\}$, he places a bid.

5. *The auction ends if one of three conditions obtains:*

   (a) *If there are no active bidders. In that case, the object is not sold.*

   (b) *If only the high bidder is active. In that case, the object is sold to the high bidder at his last bid.*

   (c) *If no active bidder can beat the current high bid b_i, that is, for every active bidder $j \neq i$, $\eta_j(\theta_j^{K_j}) < \eta_i(b_i)$. In that case, the object is sold to the high bidder at his last bid.*

Theorem 3.22. *Assume F_N is regular and interleaved and (G, S_N) is optimal. (G, S_N) is credible and strategy-proof if and only if (G, S_N) is a virtual ascending auction.*

Virtual ascending auctions score bids asymmetrically: Bidder i may be asked to bid $100 in order to beat j's bid of $50, and then to bid $101 to beat j's bid of $51. Since the auctioneer is communicating privately, she could safely deviate to equalize the prices that bidders face (provided Θ_i and Θ_j overlap enough). Nonetheless, it is incentive-compatible for the auctioneer to follow the rules. For each bidder, truthful bidding is a best-response to any safe deviation. Thus, if the auctioneer has a profitable safe deviation, then she could openly promise to deviate without undermining bidders' incentives. In that case, the original protocol was not optimal, a contradiction. It may seem intuitive that the auctioneer cannot credibly reject higher bids in favor of lower bids, but multi-round communication permits her to do so.

The virtual ascending auction can be modified to deal with irregular distributions: we simply alter Definition 3.21 to use ironed virtual values instead of virtual values, following the construction in Elkind (2007). In effect, if we iron virtual values in the interval θ_i^k to $\theta_i^{k'}$, the auctioneer promises ahead of time to jump i's price directly from θ_i^k to $\theta_i^{k'+1}$. The proof that this is credible is the same as in the regular case.

Finally, the virtual ascending auction can be used to construct a static credible optimal auction. Consider a modified all-pay auction; each type θ_i makes a bid equal to the expected payment of θ_i in the virtual ascending auction, to be paid regardless of whether he wins. The winner is the bidder with the highest virtual value. This twin-bid auction is BIC and optimal, but neither strategy-proof nor *ex post* individually rational.[31]

3.3 A note on the Dutch auction

The Dutch (descending) auction is neither strategy-proof nor static, but it is credible. In a Dutch auction, the price falls until one bidder claims the object. Thus, each bidder in sees a sequence of descending prices $(p_i^1, p_i^2, p_i^3, \ldots)$; once he claims the object, he wins

---

[31]This format is closely related to the 'all-pay' procurement auctions studied in Dequiedt and Martimort (2015).

at that price. Consequently, once one bidder makes a claim, it is not safe to deviate - the auctioneer must sell to that bidder at his current price. Fixing S_N, each bidder has a claim-price $p_i(\theta_i)$ at which he will agree. For a given θ_N, the rule-following auctioneer strategy yields revenue $\max_{i \in N} p_i(\theta_i)$. No safe deviation results in bidder i paying more than $p_i(\theta_i)$, so the revenue from following the rules first-order stochastically dominates the revenue from any safe deviation.

4 Extensions

Appendix C studies a number of extensions to the benchmark model.

Appendix C.1 relaxes the assumption that bidders' types are independent, so that the optimal auction extracts full surplus (Cremer and McLean, 1988). Static Cremer-Maclean mechanisms are not credible, since two type profiles with the same winning bidder may have different profiles of transfers. Even using extensive form mechanisms does not in general allow credible full-surplus extraction.

Next, we assume symmetric and affiliated type spaces, and constrain the auctioneer to use *ex post* incentive compatible and *ex post* individually rational mechanisms. In this setting, a modified ascending auction is optimal (Roughgarden and Talgam-Cohen, 2013), and is also credible.

Appendix C.2 assumes independent private values, and relaxes the assumption that there is a single object for sale. Instead, the feasible sets of winning bidders are a matroid. We prove that there exists a credible strategy-proof optimal auction.

5 Alternative definitions

5.1 Group-credible mechanisms

Our main purpose in this paper is to study auctioneer incentives under private communication. Nonetheless, it is natural to consider what happens under other communication structures. Here we develop an extension that permits agents to share information in groups, and show that increasing information-sharing makes it harder for the auctioneer to deviate.

Essentially, we partition agents into groups in advance, and permit each group of agents to share information after the auction, so that the auctioneer can only hide deviations by misrepresenting the behavior of other groups. Let Λ be a partition on N, and let λ denote a cell of Λ.

Definition 5.1. *Given some promised strategy profile (S_0, S_N), auctioneer strategy $\hat{S}_0$ is Λ-safe if for all groups $\lambda \in \Lambda$ and all type profiles θ_N, there exists $\hat{\theta}_{-\lambda}$ such that for*

all $i \in \lambda$, $o_i(\hat{S}_0, S_N, \theta_N) = o_i(S_0, S_N, (\theta_\lambda, \hat{\theta}_{-\lambda}))$. $\mathcal{S}_0^\Lambda(S_0, S_N)$ *denotes the set of* Λ-*safe strategies.*

Definition 5.1 permits the auctioneer to deviate only if every group's observations have an **innocent explanation**; there must exist $\hat{\theta}_{-\lambda}$ such that all observations by agents in λ are consistent the auctioneer playing S_0^G, the agents playing S_N, and the other groups' types being $\hat{\theta}_{-\lambda}$. Notably, the order of quantifiers in Definition 5.1 requires a single explanation to be offered to the entire group, which is more demanding than if we permit each observation in the group to have a different explanation.

Coarser partitions imply more information sharing between agents.

Definition 5.2. (G, S_N) *is* Λ-***credible*** *if*

$$S_0^G \in \operatorname*{argmax}_{S_0 \in \mathcal{S}_0^\Lambda(S_0^G, S_N)} \mathbb{E}_{\theta_N}[u_0(S_0, S_N, \theta_N)] \tag{10}$$

Proposition 5.3. *If* Λ *is coarser than* Λ' *and* (G, S_N) *is* Λ'-*credible, then* (G, S_N) *is* Λ-*credible.*

Proof. We will prove that, if Λ is coarser than Λ', then $\mathcal{S}_0^\Lambda(S_0^G, S_N) \subseteq \mathcal{S}_0^{\Lambda'}(S_0^G, S_N)$. From that, Proposition 5.3 follows immediately.

Take any $\hat{S}_0 \in \mathcal{S}_0^\Lambda(S_0^G, S_N)$, any group $\lambda' \in \Lambda'$ and any θ_N. Since Λ is coarser than Λ', we can find a group $\lambda \in \Lambda$ such that $\lambda \supseteq \lambda'$. Let $\hat{\theta}_{-\lambda}$ be such that for all $i \in \lambda$, $o_i(\hat{S}_0, S_N, \theta_N) = o_i(S_0^G, S_N, (\theta_\lambda, \hat{\theta}_{-\lambda}))$. Observe that $(\hat{\theta}_{-\lambda}, \theta_{\lambda \setminus \lambda'})$ is an innocent explanation for group λ' at type profile θ_N. Thus, $\hat{S}_0 \in \mathcal{S}_0^{\Lambda'}(S_0^G, S_N)$. $\qquad\square$

One interpretation of Proposition 5.3 is that starting with a Λ-credible mechanism and increasing information-sharing does not undermine auctioneer incentives. Equivalently, starting with a mechanism that is not Λ-credible and reducing information-sharing does not restore auctioneer incentives. When Λ is the finest partition, then Definition 5.2 is equivalent to Definition 2.5.

The second-price auction is not Λ-credible, unless Λ is the coarsest partition. If even a single bidder is unwilling to share information about his bids, then the auctioneer can profitably deviate by misrepresenting that bidder's behavior.

5.2 A 'Prior-free' Definition

The definition of credibility depends on the joint distribution of agent types (Definition 2.5). It may be useful to have a definition that is 'prior-free', for settings such as matching or maxmin mechanism design.

Definition 5.4. *Given* (G, S_N), $S_0 \in \mathcal{S}_0^*(S_0^G, S_N)$ *is* ***always-profitable*** *if, for all* θ_N:

$$u_0(S_0, S_N, \theta_N) \geq u_0(S_0^G, S_N, \theta_N) \tag{11}$$

with strict inequality for some θ_N.

(G, S_N) *is **prior-free credible** if no safe deviation is always-profitable.*

For comparison, (G, S_N) is credible if no safe deviation is profitable in expectation. Prior-free credibility allows one to dispense with strong assumptions about the auctioneer's beliefs.

With continuum type-spaces, credibility neither implies nor is implied by prior-free credibility. This is because some always-profitable deviations are strictly profitable only on a zero-measure set.

Replacing credibility with prior-free credibility does not essentially change any of our characterizations. Indeed, for the continuum type-spaces, requiring prior-free credibility sharpens the results, since it pins down the payment rule even on measure-zero sets:

Proposition 5.5. *Suppose the continuum type-space model of Section 3.1. Assume the distributions are symmetric and regular. Assume (G, S_N) is winner-paying and optimal. (G, S_N) is a first-price auction if and only if (G, S_N) is credible and static.*

With finite type-spaces, every credible protocol is prior-free credible. Nonetheless, prior-free credibility is enough to pin down the extensive form of the ascending auction:

Proposition 5.6. *Suppose the finite type-space model of Section 3.2. Assume the type distributions are regular and symmetric and (G, S_N) is orderly and optimal. (G, S_N) is prior-free credible and strategy-proof if and only if (G, S_N) is an ascending auction.*

6 Discussion

It is worth considering why real-world auctioneers might lack full commitment power. Vickrey (1961) suggests that the seller could delegate the task of running the auction to a third-party who has no stake in the outcome. However, auction houses such as Sotheby's, Christie's, and eBay charge commissions that are piecewise-linear functions of the sale price.[32] Running an auction takes effort, and many dimensions of effort are not contractible. Robust contracts reward the auctioneer linearly with revenue (Carroll, 2015), so it is difficult to employ a third-party who is both neutral and well-motivated.[33]

When an auctioneer makes repeated sales, reputation could help enforce the full-commitment outcome. However, the force of reputation depends on the discount rate

---

[32]http://www.sothebys.com/en/news-video/blogs/all-blogs/sotheby-s-at-large/2016/10/important-update-regarding-sothebys-buyers-premium.html, http://www.christies.com/buying-services/buying-guide/financial-information/, http://pages.ebay.com/help/sell/fees.html, accessed 11/5/2017.

[33]As Myerson (2009) observes, "The problems of motivating hidden actions can explain why efficient institutions give individuals property rights, as owners of property are better motivated to maintain it. But property rights give people different vested interests, which can make it more difficult to motivate them to share their private information with each other."

and the detection rate of deviations. Safe deviations are precisely those that a bidder could not detect immediately. Online advertising auctions are repeated frequently, so it is plausible that bidders could examine the statistics to detect foul play.[34] However, some economically important auctions are infrequent or not repeated at all - for instance, auctions for wireless spectrum or for the privatization of state-owned industries. Even established auction houses such as Christie's and Sotheby's have faced regulatory scrutiny, based in part on concerns that certain deviations are difficult for individual bidders to detect.

Modern auctioneers could use cryptography to prove that the rules of the auction have been followed, without disclosing additional information to bidders. Cryptographic verification relies on digital infrastructure: Participants typically need access to a public bulletin board, a sound method of creating and sharing public keys, and a time-lapse encryption service that provides public keys and commits to release the corresponding decryption keys only at pre-defined times (Parkes et al., 2015).[35] It can be costly to construct this infrastructure, and to persuade bidders that it works as the auctioneer claims. By using credible mechanisms, auctioneers may increase the resources and attention available for substantive purposes.

Not all auctioneers have full commitment power, just as not all firms are Stackelberg leaders. When the auctioneer lacks full commitment, it can be hazardous for bidders to reveal all their information at once. In a first-price auction, a bidder 'reveals' his value in return for a guarantee that his report completely determines the price he might pay.[36] In an ascending auction, a bidder reports whether his value is above b only when the auctioneer (correctly) asserts that bids below b are not enough to win. Credibility is a shared foundation for these seemingly disparate designs. How this principle extends to other environments is an open question.

References

Ashenfelter, O. (1989). How auctions work for wine and art. *Journal of Economic Perspectives*, 3(3):23–36.

---

[34]However, bidders in online advertising auctions have expressed concerns that supply-side platforms (SSPs) are deviating from the rules of the second-price auction. The industry news website Digiday alleged, "Rather than setting price floors as a flat fee upfront, some SSPs are setting high price floors after their bids come in as a way to squeeze out more money from ad buyers who believe they are bidding into a second-price auction". https://digiday.com/marketing/ssps-use-deceptive-price-floors-squeeze-ad-buyers/, accessed 11/30/2017.

[35]Bidders may even need special training or software assistance to play their part in a cryptographic protocol.

[36]This property is generalized in a natural way by the 'first-price' menu auction (Bernheim and Whinston, 1986).

Athey, S., Levin, J., and Seira, E. (2011). Comparing open and sealed bid auctions: Evidence from timber auctions. *The Quarterly Journal of Economics*, 126(1):207–257.

Baliga, S., Corchon, L. C., and Sjöström, T. (1997). The theory of implementation when the planner is a player. *Journal of Economic Theory*, 77(1):15–33.

Bergemann, D., Brooks, B., and Morris, S. (2017). First-price auctions with general information structures: Implications for bidding and revenue. *Econometrica*, 85(1):107–143.

Bergemann, D., Brooks, B., and Morris, S. (2018). Revenue guarantee equivalence. *working paper*.

Bergemann, D. and Morris, S. (2005). Robust mechanism design. *Econometrica*, 73(6):1771–1813.

Bernheim, B. D. and Whinston, M. D. (1986). Menu auctions, resource allocation, and economic influence. *The Quarterly Journal of Economics*, 101(1):1–31.

Bester, H. and Strausz, R. (2000). Imperfect commitment and the revelation principle: the multi-agent case. *Economics Letters*, 69(2):165–171.

Bester, H. and Strausz, R. (2001). Contracting with imperfect commitment and the revelation principle: the single agent case. *Econometrica*, 69(4):1077–1098.

Bikhchandani, S., De Vries, S., Schummer, J., and Vohra, R. V. (2011). An ascending Vickrey auction for selling bases of a matroid. *Operations research*, 59(2):400–413.

Brooks, B. and Du, S. (2018). Optimal auction design with common values: An informationally-robust approach. *working paper*.

Bulow, J., Huang, M., and Klemperer, P. (1999). Toeholds and takeovers. *Journal of Political Economy*, 107(3):427–454.

Carroll, G. (2015). Robustness and linear contracts. *American Economic Review*, 105(2):536–563.

Cassady, R. (1967). *Auctions and auctioneering*. Univ of California Press.

Chakraborty, I. and Kosmopoulou, G. (2004). Auctions with shill bidding. *Economic Theory*, 24(2):271–287.

Chung, K.-S. and Ely, J. C. (2007). Foundations of dominant-strategy mechanisms. *The Review of Economic Studies*, 74(2):447–476.

Cremer, J. and McLean, R. P. (1988). Full extraction of the surplus in bayesian and dominant strategy auctions. *Econometrica: Journal of the Econometric Society*, pages 1247–1257.

Dequiedt, V. and Martimort, D. (2006). Mechanism design with private communication. *working paper*.

Dequiedt, V. and Martimort, D. (2015). Vertical contracting with informational opportunism. *American Economic Review*, 105(7):2141–2182.

Dworczak, P. (2017). Mechanism design with aftermarkets: Cutoff mechanisms. *working paper*.

Edelman, B., Ostrovsky, M., and Schwarz, M. (2007). Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American Economic Review*, pages 242–259.

Elkind, E. (2007). Designing and learning optimal finite support auctions. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 736–745. Society for Industrial and Applied Mathematics.

Fang, H. and Morris, S. (2006). Multidimensional private value auctions. *Journal of Economic Theory*, 126(1):1–30.

Green, J. and Laffont, J.-J. (1977). Characterization of satisfactory mechanisms for the revelation of preferences for public goods. *Econometrica*, pages 427–438.

Hafalir, I. and Krishna, V. (2008). Asymmetric auctions with resale. *American Economic Review*, 98(1):87–112.

Hajłasz, P. and Malý, J. (2002). Approximation in sobolev spaces of nonlinear expressions involving the gradient. *Arkiv för Matematik*, 40(2):245–274.

Hartline, J. D. (2016). Mechanism design and approximation. *Book draft*, 122.

Holmström, B. (1979). Groves' scheme on restricted domains. *Econometrica*, pages 1137–1144.

Hurwicz, L. (1972). On informationally decentralized systems. In McGuire, C. B. and Radner, R., editors, *Decision and Organization*, chapter 14, pages 297–336. North-Holland, Amsterdam.

Krishna, V. (2010). *Auction Theory*. Academic Press, San Diego, USA.

Lamy, L. (2009). The shill bidding effect versus the linkage principle. *Journal of Economic Theory*, 144(1):390–413.

Li, S. (2017). Obviously strategy-proof mechanisms. *American Economic Review*, 107(11):3257–87.

Liu, Q., Mierendorff, K., Shi, X., et al. (2014). Auctions with limited commitment. Technical report, Working Paper.

Loertscher, S. and Marx, L. M. (2017). Optimal clock auctions. *working paper*.

Lucking-Reiley, D. (2000). Vickrey auctions in practice: From nineteenth-century philately to twenty-first-century e-commerce. *The Journal of Economic Perspectives*, 14(3):183–192.

Maskin, E. and Riley, J. (1984). Optimal auctions with risk averse buyers. *Econometrica: Journal of the Econometric Society*, pages 1473–1518.

McAdams, D. and Schwarz, M. (2007a). Credible sales mechanisms and intermediaries. *American Economic Review*, 97(1):260–276.

McAdams, D. and Schwarz, M. (2007b). Who pays when auction rules are bent? *International Journal of Industrial Organization*, 25(5):1144–1157.

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

McAfee, R. P. and Vincent, D. (1997). Sequentially optimal auctions. *Games and Economic Behavior*, 18(2):246–276.

Meyer-Nieberg, P. (1991). *Banach lattices.* Springer-Verlag.

Milgrom, P. (1987). Auction theory. In Bewley, T., editor, *Advances in Economic Theory: Fifth World Congress*, pages 1–32. Cambridge University Press, Cambridge.

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

Milgrom, P. and Segal, I. (2017). Deferred-acceptance auctions and radio spectrum reallocation. Technical report, working paper.

Milgrom, P. R. and Weber, R. J. (1982). A theory of auctions and competitive bidding. *Econometrica: Journal of the Econometric Society*, pages 1089–1122.

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

Myerson, R. B. (1986). Multistage games with communication. *Econometrica: Journal of the Econometric Society*, pages 323–358.

Myerson, R. B. (2009). Fundamental theory of institutions: a lecture in honor of leo hurwicz. *Review of Economic Design*, 13(1-2):59.

Parkes, D. C., Thorpe, C., and Li, W. (2015). Achieving trust without disclosure: Dark pools and a role for secrecy-preserving verification. In *Third Conference on Auctions, Market Mechanisms and Their Applications*.

Porter, R. and Shoham, Y. (2005). On cheating in sealed-bid auctions. *Decision Support Systems*, 39(1):41–54.

Rothkopf, M. H. and Harstad, R. M. (1995). Two models of bid-taker cheating in vickrey auctions. *Journal of Business*, pages 257–267.

Rothkopf, M. H., Teisberg, T. J., and Kahn, E. P. (1990). Why are Vickrey auctions rare? *Journal of Political Economy*, pages 94–109.

Roughgarden, T. and Talgam-Cohen, I. (2013). Optimal and near-optimal mechanism design with interdependent values. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 767–784. ACM.

Simon, L. K. and Stinchcombe, M. B. (1989). Extensive form games in continuous time: Pure strategies. *Econometrica*, pages 1171–1214.

Skreta, V. (2006). Sequentially optimal mechanisms. *The Review of Economic Studies*, 73(4):1085–1111.

Skreta, V. (2015). Optimal auction design under non-commitment. *Journal of Economic Theory*, 159:854–890.

Vartiainen, H. (2013). Auction design without commitment. *Journal of the European Economic Association*, 11(2):316–342.

Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *The Journal of Finance*, 16(1):8–37.

Von Neumann, J. and Morgenstern, O. (1953). *Theory of Games and Economic Behavior: 3d Ed.* Princeton University Press.

Wilson, R. (1987). Game theoretic analysis of trading processes. In Bewley, T., editor, *Adv. Econ. Theory*. Cambridge University Press.

A   Definition of Extensive Game Forms with Consequences in X

An *extensive game form with consequences in X* is a tuple $(H, \prec, P, A, \mathcal{A}, (\mathcal{I}_i)_{i \in N}, g)$, where:

1. H is a set of histories, along with a binary relation $\prec$ on H that represents precedence.

   (a) $\prec$ is a partial order, and $(H, \prec)$ form an arborescence[37].

   (b) We use $h \preceq h'$ if $h = h'$ or $h \prec h'$.

   (c) $h_\emptyset$ denotes $h \in H : \neg \exists h' : h' \prec h$.

   (d) $Z \equiv \{h \in H : \neg \exists h' : h \prec h'\}$

   (e) $\text{succ}(h)$ denotes the set of immediate successors of h.

2. P is a player function. $P : H \setminus Z \to N$.

3. A is a set of actions.

4. $\mathcal{A} : H \setminus h_\emptyset \to A$ labels each non-initial history with the last action taken to reach it.

   (a) For all h, $\mathcal{A}$ is one-to-one on $\text{succ}(h)$.

   (b) $A(h)$ denotes the actions available at h.

$$A(h) \equiv \bigcup_{h' \in \text{succ}(h)} \mathcal{A}(h') \tag{12}$$

5. $\mathcal{I}_i$ is a partition of $\{h : P(h) = i\}$ such that:

   (a) $A(h) = A(h')$ whenever h and h' are in the same cell of the partition.

   (b) For any $I_i \in \mathcal{I}_i$, we denote: $P(I_i) \equiv P(h)$ for any $h \in I_i$. $A(I_i) \equiv A(h)$ for any $h \in I_i$.

   (c) Each action is available at only one information set: If $a \in A(I_i)$, $a' \in A(I'_j)$, $I_i \neq I'_j$ then $a \neq a'$.

6. g is an outcome function. It associates each terminal history with an outcome. $g : Z \to X$

---

[37]That is, a directed rooted tree such that every edge points away from the root.

B Proofs omitted from the main text

B.1 Proposition 2.3

Suppose that (G, S_N) does not satisfy Clause 1 of Definition 2.2. We can modify (G, S_N) so that it satisfies Clause 1, remains BIC, and results in the same outcomes for each type profile.

In particular, suppose there exists h such that there is no θ_N such that $h \preceq z(S_N(\theta_N))$. Since the game tree has finite depth, we can locate an earliest possible h; that is, an h such that no predecessor satisfies this property. Consider h' that immediately precedes h, and the information set I_i' such that $h \in I_i'$. There is some action a' at I_i' that is not played by any type of i that reaches I_i'. We can delete all histories that follow i playing a' at I_i' (and define $(\prec', \mathcal{A}', P', (\mathcal{I}_i')_{i \in N}, g')$ and S_N' so that they are as in G, but restricted to the new smaller set of histories H'). Since these histories were off the path of play, their deletion does not affect the incentives of agents in $N \setminus i$. Since each type θ_i preferred $S_i(\theta_i)$ to any interim strategy that played a' at I_i', his new interim strategy $S_i'(\theta_i)$ remains incentive-compatible. Thus, the transformed (G', S_N') is BIC. We do this for all such histories simultaneously, to produce a protocol that satisfies Clause 1.

Suppose that (G, S_N) satisfies Clause 1 but not Clause 2. We now modify (G, S_N) so that it satisfies Clause 1 and Clause 2, remains BIC, and results in the same outcomes for each type profile.

Suppose there exists $h \notin Z$ such that $|\text{succ}(h)| = 1$. We simply rewrite the transformed game (G', S_N') that deletes h (and all the other histories in that same information set) and 'automates' i's singleton action at h. That is, for all $h' \in I_i$ for I_i such that $h \in I_i$, we remove h' from the set of histories, and define $(\prec', \mathcal{A}', P', (\mathcal{I}_i')_{i \in N}, g')$ and S_N' so that they are as in G, but restricted to $H \setminus I_i$. We do this for all singleton-action histories simultaneously, to produce a protocol that satisfies Clause 1 and Clause 2.

We now take (G, S_N) that satisfies Clauses 1 and 2, and transform it to satisfy Clause 3. Informally, our argument proceeds as follows: Suppose there is some h at which Clause 3 is not satisfied, where we denote $i = P(h)$. Upon reaching h, i's continuation strategy no longer affects the outcome. Consider a modified protocol (G', S_N'): Play proceeds exactly as in (G, S_N), except after history h is reached. Whenever, under (G, S_N), i would be called to play at h' where $h \preceq h'$, we instead skip i's turn and continue play as though i chose the action that would be selected by some type θ_i.

Formally, suppose Clause 1 and 2 hold for (G, S_N), but there exists $h \notin Z$, such that for $i = P(h)$, there does not exist $\theta_i, \theta_i', \theta_{-i}$ such that

1. $h \prec z(S_N(\theta_i, \theta_{-i}))$

2. $h \prec z(S_N(\theta_i', \theta_{-i}))$

3. $x^G(S_N(\theta_i, \theta_{-i})) \neq x^G(S_N(\theta_i', \theta_{-i}))$

Since Clause 1 holds, there exists (θ_i, θ_{-i}) such that $h \prec z(S_N(\theta_i, \theta_{-i}))$. Upon reaching h, we can henceforth 'automate' play as though i had type θ_i. First, we delete any history h' such that $h \preceq h'$ and $P(h') = i$; this ensures that i is no longer called to play after h. Next, we delete any history h' such that $h \preceq h'$ and there does not exist θ_{-i}'' such that $h' \preceq z(S_N(\theta_i, \theta_{-i}''))$; this has the effect of 'automating' play as though i has type θ_i. Given the new smaller set of histories H', we again define $(\prec', \mathcal{A}', P', (\mathcal{I}_i')_{i \in N}, g')$ and S_N' so that they are as in G, but restricted to H'. We perform this deletion simultaneously for all histories that violate Clause 3.

By construction, for all θ_i', if i is playing as though his type is θ_i' and we would have reached some deleted history h under (G, S_N), then the outcome is the same under (G', S_N') as when i is playing as though his type is θ_i under (G, S_N) (which by hypothesis is the same as when i is playing as though his type is θ_i' under (G, S_N)). Plainly, if we would not have reached history a deleted history under (G, S_N), then the outcomes under (G, S_N) and (G', S_N') are identical. Thus, (G', S_N') is BIC, satisfies Clauses 1, 2, and 3, and results in the same outcomes for each type profile.

This completes the proof of Proposition 2.3.

B.2 Proposition 2.7

To prove Proposition 2.7, we show that each of the three transformations we used in the proof of Proposition 2.3 also preserve credibility. That is, for each (G', S_N') that is produced from (G, S_N) by one of the three transformations, if the auctioneer has a profitable safe deviation from $S_0^{G'}$, then she also has a profitable safe deviation from S_0^G.

Consider the first transformation (deleting all histories that are not reached at any type profile). Suppose the auctioneer had a profitable safe deviation S_0' from $S_0^{G'}$. The auctioneer could make that same deviation in the messaging game generated by (G, S_N) (with her play specified arbitrarily after actions that correspond to deleted histories). At every type profile, the agents never reply with actions corresponding to the deleted histories, so the auctioneer's deviation is in $S_0^*(S_0^G, S_N)$.

Consider the second transformation (deleting all histories with singleton action sets). Suppose the auctioneer had a profitable safe deviation S_0' from $S_0^{G'}$. The auctioneer could make that same deviation from S_0^G, except that for any deleted information set, the auctioneer delays sending the corresponding message until the last possible moment. That is consider S_0 that is the same as S_0', except that:

1. If agent i last received message I_i, and S_0' specifies that the auctioneer sends I_i' to i, let $(I_i^1, I_i^2, \ldots, I_i^K)$ denote the sequence of deleted information sets that i would have encountered between I_i and I_i' under G (this sequence is possibly empty, and is

37

unique by perfect recall). S_0 specifies that the auctioneer first sends $(I_i^1, I_i^2, \ldots, I_i^K)$ and then (immediately thereafter) sends I_i'.

2. If agent i last received message I_i, and S_0' specifies that the auctioneer chooses outcome x, let $(I_i^1, I_i^2, \ldots, I_i^K)$ denote the (possibly empty) sequence of deleted information sets that i would have encountered (under G) between I_i and some terminal history $z \succ I_i$ such that $\exists \omega_i \in \Omega_i : \{g(z)\} \cup \{x\} \in \omega_i$. At least one such history exists because S_0' is a safe deviation. S_0 specifies that the auctioneer sends $(I_i^1, I_i^2, \ldots, I_i^K)$ before choosing x.

S_0 is a profitable safe deviation from S_0^G.

Consider the third transformation (deleting histories where i is called to play, following any history h such that, for any two types of i that reach h, both types of i result in the same outcome). Suppose S_0' was a profitable safe deviation from $S_0^{G'}$. The auctioneer can make that same deviation from S_0^G, except that she delays any 'outcome-irrelevant' queries to i until just before she selects the outcome.

Formally, take any θ_N, i, and $\hat{\theta}_{-i}$ such that $o_i(S_0', S_N, \theta_N) = o_i(S_0^{G'}, S_N, (\theta_i, \hat{\theta}_{-i}))$. If $o_i(S_0^{G'}, S_N, (\theta_i, \hat{\theta}_{-i})) \neq o_i(S_0^G, S_N, (\theta_i, \hat{\theta}_{-i}))$, then this can only be because $o_i(S_0^G, S_N, (\theta_i, \hat{\theta}_{-i}))$ contains additional communication at the end of the sequence that corresponds to deleted histories at which i is called to play. Let h be the earliest such deleted history that would be encountered under (G, S_N) at type profile $(\theta_i, \hat{\theta}_{-i})$. We can 'fill in' the missing communication for agent i, as follows. Initialize $\hat{h} := h$.

1. If $\hat{h} \in Z$, then terminate.

2. Else if $P(\hat{h}) \neq i$, then for $I_{P(\hat{h})}$ such that $\hat{h} \in I_{P(\hat{h})}$:

   (a) $\hat{h} := h' \mid h' \in \text{succ}(\hat{h})$ and $S_{P(\hat{h})}(I_{P(\hat{h})}, \hat{\theta}_{P(\hat{h})}) = \mathcal{A}(h')$.

   (b) Go to step 1.

3. Else:

   (a) Send (to agent i) message I_i such that $\hat{h} \in I_i$.

   (b) Upon receiving reply a, choose $\hat{h} := h' \mid \mathcal{A}(h') = a$ and $h' \in \text{succ}(\hat{h})$.

   (c) Go to step 1.

Since (under S_N), i's play in the deleted histories makes no difference to the outcome, delaying communication with i until the outcome is about to be selected results in a safe deviation. Thus, whenever S_0' would select an outcome, we can run the above algorithm for every agent whose resulting observation would not have an innocent explanation, and then select the same outcome, thus producing a profitable safe deviation from S_0^G. This completes the proof of Proposition 2.7.

B.3   Theorem 3.10

Suppose (G, S_N) is a twin-bid auction. (G, S_N) is static by definition. Given any profile of bids $(b_i^W, b_i^L)_{i \in N}$, every safe deviation charges b_i^W if agent i wins and b_i^L if he loses, so the auctioneer prefers S_0^G to any safe deviation. Thus, (G, S_N) is credible.

Suppose (G, S_N) is credible and static. By Theorem 3.3, there exists a function $\tilde{b}_i^W :$ $\Theta_i \to \mathbb{R}$ such that, almost everywhere in Θ_N, if type θ_i wins, then i pays $\tilde{b}_i^W(\theta_i)$.

By Lemma 3.4, let $(\underline{\theta}_{-i}^k)_{k=1}^\infty$ be a countable subset such that for all θ_{-i}, $\inf_k \tilde{t}_i(\cdot, \underline{\theta}_{-i}^k) \leq$ $\tilde{t}_i(\cdot, \theta_{-i})$ almost everywhere in Θ_i.

Define:

$$\tilde{t}_i^L(\theta_i, \theta_{-i}) = \begin{cases} \tilde{t}_i(\theta_i, \theta_{-i}) & \text{if } \tilde{y}(\theta_i, \theta_{-i}) \neq i \\ \inf_k \tilde{t}_i(\theta_i, \underline{\theta}_{-i}^k) - 1 & \text{otherwise} \end{cases} \tag{13}$$

Intuitively, the function constructed above 'penalizes' the auctioneer's revenue from i unless the type profile is consistent with i losing.

Since $\tilde{y}(\cdot, \theta_{-i})$, $\tilde{t}_i(\cdot, \theta_{-i})$, and $\inf_k \tilde{t}_i(\cdot, \underline{\theta}_{-i}^k)$ are measurable, it follows that $\tilde{t}_i^L(\cdot, \theta_{-i})$ is measurable. Again applying Lemma 3.4, let $(\overline{\theta}_{-i}^k)_{k=1}^\infty$ be a countable subset of opponent type profiles, such that for all θ_{-i}, $\sup_k \tilde{t}_i^L(\cdot, \overline{\theta}_{-i}^k) \geq \tilde{t}_i^L(\cdot, \theta_{-i})$ almost everywhere in Θ_i.

We now assert that, almost everywhere in Θ_N, if type θ_i does not win, then that type is charged $\tilde{b}_i^L(\theta_i) = \sup_k \tilde{t}_i^L(\cdot, \overline{\theta}_{-i}^k)$. Suppose the set

$$\{\theta_N \mid \tilde{y}(\theta_N) \neq i \text{ and } \tilde{t}_i(\theta_i, \theta_{-i}) \neq \sup_k \tilde{t}_i^L(\theta_i, \overline{\theta}_{-i}^k)\} \tag{14}$$

has positive measure. Observe that for (θ_i, θ_{-i}) in the above set, $\tilde{t}_i(\theta_i, \theta_{-i}) = \tilde{t}_i^L(\theta_i, \theta_{-i})$. Consequently, the set

$$Q = \{\theta_N \mid \tilde{y}(\theta_N) \neq i \text{ and } \inf_k \tilde{t}_i(\theta_i, \underline{\theta}_{-i}^k) \leq \tilde{t}_i(\theta_i, \theta_{-i}) < \sup_k \tilde{t}_i^L(\theta_i, \overline{\theta}_{-i}^k)\} \tag{15}$$

has positive measure. Q is measurable with respect to the equilibrium action profiles.

We now construct a profitable safe deviation. Fix some finite K. If the agents' chosen actions are consistent with any type profile $(\theta_i, \theta_{-i}) \in Q$, charge $\max\{\tilde{t}_i(\theta_i, \theta_{-i}), \max_{k \leq K} \tilde{t}_i^L(\theta_i, \overline{\theta}_{-i}^k)\}$, without changing the allocation or the other agents' transfers. Otherwise, play according to S_0^G. The resulting transfer $\tilde{t}_i^K : \Theta_N \to \mathbb{R}$ is measurable. Notice that our construction of Q and $\tilde{t}_i^L(\cdot)$ means that we charge more than $\tilde{t}_i(\theta_i, \theta_{-i})$ only if $\max_{k \leq K} \tilde{t}_i^L(\theta_i, \overline{\theta}_{-i}^k)$ is consistent with i losing.

For K large enough, this deviation is profitable. In particular, for all $(\theta_i, \theta_{-i}) \in Q$, $\tilde{t}_i^K(\theta_i, \theta_{-i})$ is non-decreasing in K and converges as $K \to \infty$ to $\sup_k \tilde{t}_i^L(\theta_i, \overline{\theta}_{-i}^k)$. Thus, by

the monotone convergence theorem,

$$\lim_{K \to \infty} \mathbb{E}_{\theta_N} \left[\tilde{t}_i^K(\theta_N) \mid \theta_N \in Q \right] = \mathbb{E}_{\theta_N} \left[\sup_k \tilde{t}_i^L(\theta_i, \theta_{-i}^k) \mid \theta_N \in Q \right]$$
$$> \mathbb{E}_{\theta_N} \left[\tilde{t}_i(\theta_N) \mid \theta_N \in Q \right] \tag{16}$$

which establishes that the deviation is profitable.

We have shown that there exist $\tilde{b}_i^W : \Theta_i \to \mathbb{R}$ and $\tilde{b}_i^L : \Theta_i \to \mathbb{R}$ such that, almost everywhere in θ_N, i pays $\tilde{b}_i^W(\theta_i)$ if $\tilde{y}(\theta_i, \theta_{-i}) = i$ and $\tilde{b}_i^L(\theta_i)$ if $\tilde{y}(\theta_i, \theta_{-i}) \neq i$. If for all θ_{-i}, $\tilde{y}(\theta_i, \theta_{-i}) \neq i$, then we set $\tilde{b}_i^W(\theta_i)$ to be equal to $\tilde{b}_i^L(\theta_i) - 1$. We then define $B_i = \{ (\tilde{b}_i^W(\theta_i), \tilde{b}_i^L(\theta_i)) \mid \theta_i \in \Theta_i \}$, which implies that Clause 1 of Definition 3.9 holds almost everywhere. Let Υ denote the subset of Θ_N on which Clause 1 holds.

Suppose then that Clause 2 does not hold on a positive measure set. Then, for some agent i, the set

$$\{ \theta_N \mid \tilde{y}(\theta_N) = 0 \text{ and } \tilde{b}_i^W(\theta_i) - \tilde{b}_i^L(\theta_i) > 0 \} \cap \Upsilon \tag{17}$$

has positive measure. The auctioneer can raise expected revenue by deviating at all type profiles in this set, allocating the object to i and charging $\tilde{b}_i^W(\theta_i)$. Thus Clause 2 holds almost everywhere.

Suppose then that Clause 3 does not hold on a positive measure set. (G, S_N) is contestable, so for some agent i, the set

$$Q' = \{ \theta_N \mid \tilde{y}(\theta_N) = i \text{ and } \tilde{b}_i^W(\theta_i) - \tilde{b}_i^L(\theta_i) < \max\{0, \max_{j \neq i} \tilde{b}_j^W(\theta_j) - \tilde{b}_j^L(\theta_j)\}$$
$$\text{and } \exists \theta_{-i}' : \tilde{y}(\theta_i, \theta_{-i}') \neq i \} \cap \Upsilon \tag{18}$$

has positive measure. The auctioneer can raise expected revenue by deviating at all type profiles in this set. Take any type profile in $\theta_N \in Q'$.

1. If $\tilde{b}_i^W(\theta_i) - \tilde{b}_i^L(\theta_i) < 0$, then keep the object and changes i's payment to $b_i^L(\theta_i)$.

2. Else, if $\tilde{b}_i^W(\theta_i) - \tilde{b}_i^L(\theta_i) < \max_{j \neq i} \tilde{b}_j^W(\theta_j) - \tilde{b}_j^L(\theta_j)$, then award the object to the agent who maximizes the right-hand side, changes i's payment to $\tilde{b}_i^L(\theta_i)$ and the other agent's payment to $\tilde{b}_j^W(\theta_j)$.

Hence, Clause 3 holds almost everywhere, which completes the proof.

B.4 Proposition 3.12

Suppose (G, S_N) is a twin-bid auction and strategy-proof. Strategy-proofness requires:

$$\tilde{t}_i(\theta_i', \theta_{-i}) - \tilde{t}_i(\theta_i, \theta_{-i}) \leq \theta_i' \tag{19}$$

$$\theta_i'' \leq \tilde{t}_i(\theta_i''', \theta_{-i}') - \tilde{t}_i(\theta_i'', \theta_{-i}') \tag{20}$$

$\tilde{y}(\cdot)$ is non-decreasing in θ_i, so $\tilde{y}(\theta_i, \theta'_{-i}) \neq i$ and $\tilde{y}(\theta'''_i, \theta_{-i}) = i$. It follows that:

$$\tilde{t}_i(\theta_i, \theta_{-i}) = \tilde{t}_i(\theta_i, \theta'_{-i}) = \tilde{t}_i(\theta''_i, \theta'_{-i}) \tag{21}$$

$$\tilde{t}_i(\theta'''_i, \theta'_{-i}) = \tilde{t}_i(\theta'''_i, \theta_{-i}) = \tilde{t}_i(\theta'_i, \theta_{-i}) \tag{22}$$

where the first equality in each line follows from the definition of a twin-bid auction and the second equality follows from strategy-proofness. Substituting into Equation 20 yields

$$\theta''_i \leq \tilde{t}_i(\theta'_i, \theta_{-i}) - \tilde{t}_i(\theta_i, \theta_{-i}) \tag{23}$$

which contradicts Equation 19.

B.5   Theorem 3.18

B.5.1   credible, strategy-proof → ascending

We start by deriving several properties of credible strategy-proof optimal (G, S_N), without assuming that F_N is regular or symmetric. Since we are mostly holding fixed (G, S_N), we will drop the superscripts on $\tilde{y}^{G,S_N}$ and $\tilde{t}_i^{G,S_N}$ to reduce clutter.

Proposition B.1. *If (G, S_N) is optimal and strategy-proof, then (G, S_N) is winner-paying.*

Proof. For all (θ_i, θ_{-i}), if $\tilde{y}(\theta_i, \theta_{-i}) \neq i$ then $\tilde{t}_i(\theta_i, \theta_{-i}) \leq 0$. Suppose not. (G, S_N) satisfies voluntary participation. When i's opponent's imitate θ_{-i},[38] type θ_i can profitably deviate to non-participation if $\tilde{t}_i(\theta_i, \theta_{-i}) > 0$, contradicting strategy-proofness.

$\theta^1_i \leq 0$, so $\eta_i(\theta^1_i) < 0$. (G, S_N) is optimal, so θ^1_i never wins (by Proposition 3.14). θ^1_i's participation constraint binds, so for all θ_{-i}, $\tilde{t}_i(\theta^1_i, \theta_{-i}) = 0$.

Take any (θ_i, θ_{-i}). If $\tilde{y}(\theta_i, \theta_{-i}) \neq i$ and $\tilde{t}_i(\theta_i, \theta_{-i}) > 0$, then when i's opponents imitate θ_{-i}, θ^1_i can profitably imitate θ_i, contradicting strategy-proofness. Thus, (G, S_N) is winner-paying. $\square$

Proposition B.2. *If (G, S_N) is strategy-proof, then the allocation rule is monotone. That is, if $\theta_i < \theta'_i$ and $\tilde{y}(\theta_i, \theta_{-i}) = i$, then $\tilde{y}(\theta'_i, \theta_{-i}) = i$.*

Proof. Suppose not, so $\tilde{y}(\theta'_i, \theta_{-i}) \neq i$. By strategy-proofness, $-\tilde{t}_i(\theta'_i, \theta_{-i}) \geq \theta'_i - \tilde{t}_i(\theta_i, \theta_{-i})$, which implies $-\tilde{t}_i(\theta'_i, \theta_{-i}) > \theta_i - \tilde{t}_i(\theta_i, \theta_{-i})$, so θ_i can profitably imitate θ'_i, a contradiction. $\square$

Definition B.3. *(G, S_N) has **threshold pricing** if:*

$$\tilde{t}_i(\theta_N) = \begin{cases} \min_{\theta'_i \in \Theta_i} \theta'_i \mid \tilde{y}(\theta'_i, \theta_{-i}) = i & \text{if } \tilde{y}(\theta_N) = i \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

---

[38] Formally, define S'_{-i} such that for all $j \neq i$, I_j, and θ'_j, $S'_j(I_j, \theta'_j) = S_j(I_j, \theta_j)$

Proposition B.4. *If (G, S_N) is optimal and strategy-proof, then (G, S_N) has threshold pricing.*

Proof. Proposition B.1 pins down the payments whenever $\tilde{y}(\theta_N) \neq i$.

We prove the rest by induction. (G, S_N) is optimal, so θ_i^1's participation constraint binds. Thus, Equation 24 holds when for θ_i^1. Suppose that Equation 24 holds for all $\theta_i^{k'}$ such that $k' \leq k$. We prove it holds for θ_i^{k+1}.

Take any θ_{-i}. There are three cases to consider.

If $\tilde{y}(\theta_i^k, \theta_{-i}) = i$, then strategy-proofness implies that $\tilde{y}(\theta_i^{k+1}, \theta_{-i}) = i$ and $\tilde{t}_i(\theta_i^{k+1}, \theta_{-i}) = \tilde{t}_i(\theta_i^k, \theta_{-i}) = \min_{\theta_i' \in \Theta_i} \theta_i' \mid \tilde{y}(\theta_i', \theta_{-i}) = i$.

If $\tilde{y}(\theta_i^{k+1}, \theta_{-i}) \neq i$, then $\tilde{t}_i(\theta_i, \theta_{-i}) = 0$.

Notice that, in the previous two cases, θ_i^{k+1} is exactly indifferent between S_i and deviating to imitate type θ_i^k. Finally, suppose $\tilde{y}(\theta_i^k, \theta_{-i}) \neq i$ and $\tilde{y}(\theta_i^{k+1}, \theta_{-i}) = i$. $\tilde{t}_i(\theta_i^{k+1}, \theta_{-i}) \leq \theta_i^{k+1}$, since (G, S_N) is strategy-proof. If $\tilde{t}_i(\theta_i^{k+1}, \theta_{-i}) < \theta_i^{k+1}$, then (G, S_N) is not optimal, since the incentive constraints do not bind locally downward (Proposition 3.14). Thus, $\tilde{t}_i(\theta_i^{k+1}, \theta_{-i}) = \theta_i^{k+1}$, and the inductive step is proved. $\square$

Given (G, S_N), let Θ_i^h denote the types of i that are consistent with i's actions up to history h, that is:

$$\Theta_i^h = \{\theta_i \mid \forall h', h'' \preceq h : [h' \in I_i, h'' \in \text{succ}(h')] \rightarrow [S_i(I_i, \theta_i) = \mathcal{A}(h'')]\} \tag{25}$$

For $\hat{N} \subseteq N$, let $\Theta_{\hat{N}}^h = \times_{i \in \hat{N}} \Theta_i^h$.

Proposition B.5. *If $h \prec h'$ then $\Theta_i^h \supseteq \Theta_i^{h'}$. If $h \in I_i$ and $h' \in I_i$, then $\Theta_i^h = \Theta_i^{h'}$.*

The first is clear by inspection. The second follows because the definition of Θ_i^h invokes only i's past information sets and actions, and G has perfect recall. Thus, we define $\Theta_i^{I_i} = \Theta_i^h \mid h \in I_i$. Define also:

$$\underline{\theta}_i^h = \min\{\theta_i \in \Theta_i^h\} \tag{26}$$

$$\overline{\theta}_i^h = \max\{\theta_i \in \Theta_i^h\} \tag{27}$$

The next proposition states that strategy-proofness constrains what agents can learn about each others' play midway through the protocol. In essence, it says that if, at some history h where i is called to play, i can affect whether or not θ_j wins, then i cannot (at this information set) rule out the possibility that j's type is instead some $\theta_j' > \theta_j$.

Proposition B.6. *Assume (G, S_N) is optimal and strategy-proof. Take any information set I_i and history $h \in I_i$. Take any $\theta_i, \theta_i' \in \Theta_i^h$, $\theta_j \in \Theta_j^h$, and $\theta_{N \setminus \{i,j\}} \in \Theta_{N \setminus \{i,j\}}^h$.*

If $\tilde{y}(\theta_i, \theta_j, \theta_{N \setminus \{i,j\}}) = j$ and $\tilde{y}(\theta_i', \theta_j, \theta_{N \setminus \{i,j\}}) \neq j$, then $\forall \theta_j' > \theta_j : \exists h' \in I_i : \theta_j' \in \Theta_j^{h'}$ and $\theta_{N \setminus \{i,j\}} \in \Theta_{N \setminus \{i,j\}}^{h'}$.

Proof. Suppose not. We construct a strategy profile S'_{-j} such that θ'_j has a profitable deviation. For $l \in N \setminus \{i, j\}$, let l imitate θ_l; that is $\forall I_l : \forall \theta'_l : S'_l(I_l, \theta'_l) = S_l(I_l, \theta_l)$. Let i imitate θ'_i unless he encounters I_i, and let him imitate type θ_i if he has encountered I_i. Formally:

$$\forall I'_i : \forall \theta''_i : S'_i(I'_i, \theta''_i) = \begin{cases} S_i(I'_i, \theta_i) & \text{if } \exists h'' \in I'_i : \exists h''' \in I_i : h''' \preceq h'' \\ S_i(I'_i, \theta'_i) & \text{otherwise} \end{cases} \tag{28}$$

By Proposition B.4, (G, S_N) has threshold pricing. If type θ'_j deviates to imitate θ_j, then (when facing S'_j), the path of play passes through I_i, so j wins at price $\min_{\theta''_j \in \Theta_j} \theta''_j \mid \tilde{y}(\theta_i, \theta''_j, \theta_{N\setminus\{i,j\}}) = j$, for a positive surplus since $\theta'_j > \theta_j$. On the other hand, if type θ'_j plays according to S_j, then the path of play does not pass through I_i, so j either wins at a strictly higher price $\min_{\theta''_j \in \Theta_j} \theta''_j \mid \tilde{y}(\theta'_i, \theta''_j, \theta_{N\setminus\{i,j\}}) = j$, or does not win and has zero surplus. Thus, j has a profitable deviation, and (G, S_N) is not strategy-proof, a contradiction. $\qquad\square$

Let W^h_i denote the subset of i's types that might reach h and then win. Similarly, let L^h_i denote the subset of i's types that might reach h and then lose.

$$W^h_i = \{\theta_i \in \Theta^h_i \mid \exists \theta_{-i} \in \Theta^h_{-i} : \tilde{y}(\theta_i, \theta_{-i}) = i\} \tag{29}$$

$$L^h_i = \{\theta_i \in \Theta^h_i \mid \exists \theta_{-i} \in \Theta^h_{-i} : \tilde{y}(\theta_i, \theta_{-i}) \neq i\} \tag{30}$$

Definition B.7. (G, S_N) *is **winner-pooling** if for all I_i, $h \in I_i$:*

1. *Either: $\forall \theta_i, \theta'_i \in W^h_i : S_i(I_i, \theta_i) = S_i(I_i, \theta'_i)$*

2. *Or: $W^h_i \cap L^h_i = \emptyset$*

Proposition B.8. *Assume F_N is symmetric and regular, and (G, S_N) is optimal, orderly, and strategy-proof. If (G, S_N) is credible, then (G, S_N) is winner-pooling.*

Before starting the proof of Proposition B.8, we highlight that this is the reason that we have assumed regularity and orderliness in the statement of Theorem 3.18. Together, regularity and orderliness imply that, if there are two distinct types $\theta_i < \theta'_i$ in W^h_i that do not pool on the same action, then there exists θ_{-i} such that θ_i loses when facing θ_{-i}, but θ'_i wins. This enables us to construct profitable safe deviations for the auctioneer.[39]

Proof. Under the assumptions of Proposition B.8, we will show that if (G, S_N) is not winner-pooling, then the auctioneer has a profitable safe deviation, so (G, S_N) is not credible.

---

[39]If type spaces were continuous, regularity would by itself imply the desired property for *every* optimal allocation rule. However, for discrete types, we need to pick a particular allocation rule - and the orderly one will do.

Let h^* be some history at which the winner-pooling property does not hold; we pick h^* such that, for all $h \prec h^*$, h is not a counterexample to winner-pooling. Since (G, S_N) is orderly and the winner-pooling property held at all predecessors to h^*, it follows that for all i, either $W_i^{h^*} = \emptyset$ or $W_i^{h^*} = \{\theta_i \mid \theta_i \vartriangleright \max_{j \neq i}^{\vartriangleright} \underline{\theta}_j^{h^*} \text{ and } \theta_i \vartriangleright \rho\}$.

Let i^* denote $P(h^*)$, and I_{i*}^* the corresponding information set. Since the winner-pooling property doesn't hold at h^*, $W_{i*}^{h^*} \cap L_{i*}^{h^*} \neq \emptyset$ and there exist two distinct actions taken by types in $W_{i*}^{h^*}$ at I_{i*}^*.

Since (G, S_N) is orderly, $\min^{\vartriangleright} W_{i*}^{h^*} \in W_{i*}^{h^*} \cap L_{i*}^{h^*}$. Define

$$\theta_{i*}^* = \min^{\vartriangleright} \theta_{i*} \in W_{i*}^{h^*} \mid S_{i*}(I_{i*}^*, \theta_{i*}) \neq S_{i*}(I_{i*}^*, \min^{\vartriangleright} W_{i*}^{h^*}) \tag{31}$$

We are going to squeeze extra revenue out of agent i^* when his type is θ_{i*}^*: by his actions at h^*, he hints that his type is *more than high enough to win*. Let h^{**} be the immediate successor of h^* that would be reached by θ_{i*}^*, that is

$$h^{**} = h \mid h \in \text{succ}(h^*) \text{ and } \theta_{i*}^* \in \Theta_{i*}^h \tag{32}$$

Since $W_{i*}^{h^*} \cap L_{i*}^{h^*} \neq \emptyset$ and (G, S_N) is orderly, $\{j \in N \mid W_j^{h^*} \neq \emptyset\}$ includes i^* and at least one other agent. For each $i \in N$, we assign a nemesis:

$$\psi(i) = \max^{\vartriangleright}\{j \in N \setminus \{i\} \mid W_j^{h^*} \neq \emptyset\} \tag{33}$$

By choosing i's nemesis in this way, we ensure a useful property; given any θ_j, we can find $\theta_{\psi(i)}$ such that i has the same allocation and transfer when the highest opponent type is θ_j and when it is $\theta_{\psi(i)}$. Similarly, given any θ_i, we can find $\theta_{\psi(i)}$ that forces i to pay exactly θ_i if he wins (by threshold pricing). Formally, we say $\theta_{\psi(i)}$ is *i*-**equivalent to** θ_j if

$$\{\theta_i \mid \theta_i \trianglerighteq \theta_j\} = \{\theta_i \mid \theta_i \trianglerighteq \theta_{\psi(i)}\} \tag{34}$$

where $\trianglerighteq$ is the reflexive order implied by the strict order $\vartriangleright$.

Given S_0^G, we now exhibit a (partial) behavioral strategy that deviates from S_0^G upon encountering h^{**} and is strictly profitable. We describe this algorithmically. The description is lengthy, because it must produce a safe deviation for any extensive game form in a large class. We start by defining several subroutines for the algorithm.

The algorithm calls the following subroutine: Given some variable $\hat{h}$ that takes values in the set of histories, we can start at the initial value of $\hat{h}$ and communicate with i as though the opponent types were θ_{-i}, updating $\hat{h}$ as we go along. When we do this, we say that we **simulate** θ_{-i} against i starting from $\hat{h}$, until certain specified conditions are met. Formally,

1. If [**conditions**], STOP.

2. Else if $P(\hat{h}) \neq i$, set $\hat{h} := h \in \text{succ}(\hat{h}) \mid \theta_{-i} \in \Theta_{-i}^h$

3. Else if $P(\hat{h}) = i$:

    (a) Send message $I_i \mid \hat{h} \in I_i$ to i.

    (b) Upon receiving $r \in A(I_i)$, set $\hat{h} := h \mid (h \in \text{succ}(\hat{h}) \text{ and } \mathcal{A}(h) = r)$.

    (c) Go to step 1.

The algorithm also calls the following subroutine: Given some history h and some θ_{-i}, where i was called to play at h's immediate predecessor, we may **find the cousin** of h consistent with θ_{-i}. This is the history that immediately follows from the same information set, is consistent with the action i just took, but is also consistent with the opponent types being θ_{-i}. Formally, let $\text{cousin}(h, \theta_{-i})$ be equal to h' such that $\exists I_i :$ $\exists h'', h''' :$

1. $h'', h''' \in I_i$

2. $h \in \text{succ}(h'')$

3. $h' \in \text{succ}(h''')$

4. $\mathcal{A}(h) = \mathcal{A}(h')$

5. $\theta_{-i} \in \Theta_{-i}^{h'}$

Clearly, it is not always possible to find such a history. But we will be careful to prove that $\text{cousin}(h, \theta_{-i})$ is well-defined when we call it.

Our algorithm keeps track of several variables:

1. A best offer, initialized $\beta := \theta_i^*$.

2. A set of 'active' agents, initialized $\hat{N} := N$.

3. The agent we are currently communicating with, $\hat{i} := i^*$.

4. A simulated history, for each agent: $\hat{h}_{i^*} := h^{**}$ and for $i \in N \setminus \{i^*\}$, $\hat{h}_i := h^*$.

The algorithm proceeds in three stages. At h^{**}, i^*'s type could be at least $\theta_{i^*}^*$, but it could also be too low to exploit (if some types not in $W_{i^*}^{h^*}$ took the same action as $\theta_{i^*}^*$ at h^*). In Stage 1, we check whether i^*'s type is at least $\theta_{i^*}^*$. If it is, we set β to be the least type consistent with i^*'s responses, and go to Stage 2. Otherwise, we lower β appropriately, and proceed to Stage 2. In Stage 2, we cycle through the bidders, updating β to be equal to the highest type we've confirmed so far, until we have found the bidder with the highest type (breaking ties with $\rhd$). Finally, in Stage 3, we sell to the bidder with the highest type (if it's above the reserve), at a price greater than or equal to the

price in the original protocol. We use := for the assignment operator, and :$\in$ to assign an arbitrary element in the set on the right-hand side.

Stage 1

1. Pick $\theta_{\psi(i^*)}$ that is i^*-equivalent to β.

2. Simulate $(\theta_{\psi(i^*)}, \underline{\theta}^{h^*}_{N\setminus\{i^*,\psi(i^*)\}})$ against i^* starting from $\hat{h}_{i^*}$, until either $\underline{\theta}^{\hat{h}_{i^*}}_{i^*} \trianglerighteq \beta$ or $\hat{h}_{i^*} \in Z$.

3. If $\underline{\theta}^{\hat{h}_{i^*}}_{i^*} \trianglerighteq \beta$, then set $\beta := \underline{\theta}^{\hat{h}_{i^*}}_{i^*}$ and go to Stage 2.

4. Else, set $\hat{N} := \hat{N} \setminus \{i^*\}$, $\beta := \min_{i \neq i^*, \theta_i}^{\triangleright} \theta_i \mid \theta_i \in W^{h^*}_i$ and go to Stage 2.

Stage 2

1. If $\hat{N} = 1$, go to Stage 3.

2. Set $\hat{i} :\in \{i \in \hat{N} \mid \underline{\theta}^{\hat{h}_i}_i \triangleleft \beta\}$.

3. Pick $\theta_{\psi(\hat{i})}$ that is $\hat{i}$-equivalent to β.

4. If $(\theta_{\psi(\hat{i})}, \underline{\theta}^{h^*}_{N\setminus\{\hat{i},\psi(\hat{i})\}}) \notin \Theta^{\hat{h}_{\hat{i}}}_{-\hat{i}}$, set $\hat{h}_{\hat{i}} := \text{cousin}(\hat{h}_{\hat{i}}, (\theta_{\psi(\hat{i})}, \underline{\theta}^{h^*}_{N\setminus\{\hat{i},\psi(\hat{i})\}}))$.

5. Simulate $(\theta_{\psi(\hat{i})}, \underline{\theta}^{h^*}_{N\setminus\{\hat{i},\psi(\hat{i})\}})$ against $\hat{i}$ starting from $\hat{h}_{\hat{i}}$, until either $\underline{\theta}^{\hat{h}_{\hat{i}}}_{\hat{i}} \trianglerighteq \beta$ or $\hat{h}_{\hat{i}} \in Z$.

6. If $\underline{\theta}^{\hat{h}_{\hat{i}}}_{\hat{i}} \trianglerighteq \beta$, set $\beta := \underline{\theta}^{\hat{h}_{\hat{i}}}_{\hat{i}}$ and go to Step 1 of Stage 2.

7. Else, set $\hat{N} := \hat{N} \setminus \{\hat{i}\}$ and go to Step 1 of Stage 2.

Stage 3

1. Set $\hat{i} := i \mid i \in \hat{N}$.

2. Pick $\theta_{\psi(\hat{i})}$ that is $\hat{i}$-equivalent to β.

3. If $(\theta_{\psi(\hat{i})}, \underline{\theta}^{h^*}_{N\setminus\{\hat{i},\psi(\hat{i})\}}) \notin \Theta^{\hat{h}_{\hat{i}}}_{-\hat{i}}$, set $\hat{h}_{\hat{i}} := \text{cousin}(\hat{h}_{\hat{i}}, (\theta_{\psi(\hat{i})}, \underline{\theta}^{h^*}_{N\setminus\{\hat{i},\psi(\hat{i})\}}))$.

4. Simulate $(\theta_{\psi(\hat{i})}, \underline{\theta}^{h^*}_{N\setminus\{\hat{i},\psi(\hat{i})\}})$ against $\hat{i}$ starting from $\hat{h}_{\hat{i}}$, until $\hat{h}_{\hat{i}} \in Z$.

5. Choose the outcome that corresponds to that terminal history, $x = g(\hat{h}_{\hat{i}})$, and terminate.

Since (G, S_N) is orderly, the deviation does not change the allocation. In particular, some agent $\hat{i}$ is removed from $\hat{N}$ only when we know that $\theta_{\psi(\hat{i})} \triangleright \theta_{\hat{i}}$, since $\theta_{\psi(\hat{i})}$ is $\hat{i}$-equivalent to β, the latter implies that $\beta \triangleright \theta_{\hat{i}}$.[40] Moreover, since (G, S_N) is orderly and

---

[40]Since (G, S_N) is orderly, we must eventually learn either that $\theta_{\hat{i}} \triangleright \theta_{\psi(\hat{i})}$ or *vice versa*, since this information is necessary to determine whether $\hat{i}$ or $\psi(\hat{i})$ should win when the other agents' types are $\underline{\theta}^{h^*}_{N\setminus\{\hat{i},\psi(\hat{i})\}}$. Thus, reaching Step 4 of Stage 1 or Step 7 of Stage 2 implies that $\beta \triangleright \theta_{\hat{i}}$.

has threshold pricing (by Proposition B.4), the resulting algorithm results in transfers that are always at least as high as the transfers under (G, S_N). The transfers are strictly higher for at least one type profile, namely $(\theta_{i^*}^*, \underline{\theta}_{-i^*}^{h^*})$. Under (G, S_N), $\tilde{t}_{i^*}(\theta_{i^*}^*, \underline{\theta}_{-i^*}^{h^*}) = \overset{\triangleright}{\min} W_{i^*}^{h^*}$, whereas under the deviation i^*'s transfer is $\theta_{i^*}^*$. Thus, the deviation is profitable.

It remains to prove that the deviation is safe. When we first start communicating with any agent $\hat{i}$ under the deviation, we are simulating opponent types that are consistent with h^*, because the winner-pooling property holds at all histories prior to h^*, and we have chosen the simulated nemesis type $\theta_{\psi(\hat{i})}$ to be in $W_{\psi(\hat{i})}^{h^*}$. (Thus, Step 4 of Stage 2 and Step 3 of Stage 3 are not triggered if this is the first time the deviating algorithm is communicating with that agent.)

Whenever the deviation communicates with some agent $\hat{i}$ for a second time, we have to prove that we can find cousins (in Step 4 of Stage 2 and Step 3 of Stage 3) in the way the algorithm requires. Let $\theta_{\psi(\hat{i})}^{\text{old}}$ and β^{old} denote the simulated nemesis type and the best offer from the last time the algorithm communicated with $\hat{i}$. Let $\theta_{\psi(\hat{i})}^{\text{new}}$ and β^{new} denote the current simulated nemesis type and best offer. Observe that we always revise the nemesis type upwards; $\beta^{\text{old}} \trianglelefteq \beta^{\text{new}}$, so $\theta_{\psi(\hat{i})}^{\text{old}} \leq \theta_{\psi(\hat{i})}^{\text{new}}$. If $\theta_{\psi(\hat{i})}^{\text{old}} = \theta_{\psi(\hat{i})}^{\text{new}}$, we are done, since $(\theta_{\psi(\hat{i})}^{\text{old}}, \underline{\theta}_{N \setminus \{\hat{i}, \psi(\hat{i})\}}^{h^*}) \in \Theta_{-\hat{i}}^{\hat{h}_{\hat{i}}}$. Otherwise, consider h', the immediate predecessor of $\hat{h}_{\hat{i}}$. At h', $\hat{i}$ is called to play, and it is not yet clear whether $\psi(\hat{i})$ wins. In particular, $\theta_{\psi(\hat{i})}^{\text{old}}$ would win against $\underline{\theta}_{\hat{i}}^{h'}$, but would lose against $\underline{\theta}_{\hat{i}}^{\hat{h}_{\hat{i}}}$, i.e. $\tilde{y}(\underline{\theta}_{\hat{i}}^{h'}, \theta_{\psi(\hat{i})}^{\text{old}}, \underline{\theta}_{N \setminus \{\hat{i}, \psi(\hat{i})\}}^{h^*}) = \psi(\hat{i}) \neq \tilde{y}(\underline{\theta}_{\hat{i}}^{\hat{h}_{\hat{i}}}, \theta_{\psi(\hat{i})}^{\text{old}}, \underline{\theta}_{N \setminus \{\hat{i}, \psi(\hat{i})\}}^{h^*})$. By Proposition B.6, there exists another history h'' in the same information set as h', such that $(\theta_{\psi(\hat{i})}^{\text{new}}, \underline{\theta}_{N \setminus \{\hat{i}, \psi(\hat{i})\}}^{h^*}) \in \Theta_{-\hat{i}}^{h''}$. Thus, we can find cousins in the way that the algorithm requires.

Observe that, whenever $\hat{i}$ is removed from $\hat{N}$, he has seen a communication sequence that is consistent with his reaching a terminal history with an opponent type profile such that $\hat{i}$ does not win and has a zero transfer, and the Stage 3 outcome respects that. At Stage 3, the final agent $\hat{i}$'s observation is consistent with $(\theta_{\psi(\hat{i})}, \underline{\theta}_{N \setminus \{\hat{i}, \psi(\hat{i})\}}^{h^*})$. Thus, the algorithm produces a profitable safe deviation. $\qquad \square$

We are now ready to show that, under the assumptions of Theorem 3.18, if (G, S_N) is credible and strategy-proof, then (G, S_N) is an ascending auction. With Propositions B.4 and B.8 in hand, what remains is mostly an exercise in labeling.

Bidder i is **active** at h if $W_i^h \neq \emptyset$. There are three cases to consider:

1. An active bidder is called to play, and there is more than one active bidder.

2. An active bidder is called to play, and there are no other active bidders. (This can only happen when every other bidder has a type below the reserve.)

3. An inactive bidder is called to play.

Case 1: Take any I_i and $h \in I_i$ such that an active bidder i is called to play. Suppose there exists another active bidder, so $W_i^h \cap L_i^h \neq \emptyset$. There is some action $S_i(I_i, \theta_i^K)$ that is taken by the highest type of i. Proposition B.8 implies that for all $\theta_i \in W_i^h$, $S_i(I_i, \theta_i) = S_i(I_i, \theta_i^K)$. Thus, any agent who does not play that action has **quit**. The **bid** at I_i is the least type of i consistent with playing $S_i(I_i, \theta_i^K)$, that is $\overset{\triangleright}{\min}\{\theta_i \in \Theta_i^h \mid S_i(I_i, \theta_i) = S_i(I_i, \theta_i^K)\}$. By Proposition B.5, each bid is weakly more than the last bid that i placed. This construction implies that the bid is the same at all histories in I_i, since it is equal to the least type that plays the pooling action. Thus, bidder i knows the bid associated with the pooling action.

By construction, all types strictly below the bid quit. Since (G, S_N) is orderly, if there is no high bidder, then all types weakly above the reserve ρ place a bid. Similarly, all types above the current high bid place a bid.

Moreover, if i is the current high bidder at history h and there is another active bidder, then by Proposition B.8, all i's types who reach h take the same action, and (by (G, S_N) pruned) i is not called to play at h. This implies that, if an active bidder i is called to play at h, he is not the current high bidder.

Case 2: Suppose an active bidder i is called to play at some h in some information set I_i, and is the unique active bidder at h. Since (G, S_N) is pruned, i is not the current high bidder, which implies that there is *no* high bidder - all the other bidders have types below the reserve.

If there exists $h' \in I_i$ such that i is not the unique active bidder at h', then we define bids and quitting actions at h as at h'. Otherwise, we define them as follows: We define an action a as **quitting** if there is no type above the reserve that plays a, that is:

$$\neg \exists \theta_i \in \Theta_i^h \mid \theta_i \rhd \rho \text{ and } S_i(I_i, \theta_i) = a \tag{35}$$

For any non-quitting action a, the associated **bid** is:

$$\overset{\triangleright}{\min}\{\theta_i \mid \theta_i \rhd \rho \text{ or } [\theta_i \in \Theta_i^h \text{ and } S_i(I_i, \theta_i) = a]\} \tag{36}$$

By construction, if i has a type strictly below the bid associated with a, then he does not play a. If i has a type above the reserve, then he places a bid. However, $W_i^h \cap L_i^h = \emptyset$, so there can be multiple actions that place bids. Again, by Proposition B.5, each bid is weakly more than the last bid that i placed. Again, this construction implies that i knows whether an action quits, and the bid associated with each non-quitting action.

Case 3: From Case 1 and 2, if bidder i quits, then he either has a type lower than the reserve, or we have identified another bidder whose type is greater than i's (according to the order $\rhd$). Thus, since (G, S_N) is orderly, once i is inactive, further information about his type no longer affects the outcome, so (since (G, S_N) is pruned) only active bidders

48

are called to play. This implies that no histories satisfy Case 3.

The three conditions that specify what happens when the auction ends are similarly entailed by orderliness and threshold pricing (Proposition B.4). If there are no active bidders at h, then for all i, $\rho \triangleright \overline{\theta}_i^h$. Thus, the object is not sold, and since (G, S_N) is pruned, h is a terminal history. If the high bidder i is the unique active bidder at h, then we know that no bidder in $N \setminus i$ has a higher type than i, and that i's current bid is equal to $\overset{\triangleright}{\min}\{\theta_i \mid \theta_i \triangleright \overset{\triangleright}{\max_{j \neq i}} \theta_j \text{ and } \theta_i \triangleright \rho\} = \overset{\triangleright}{\min}\{\theta_i \mid \tilde{y}(\theta_i, \theta_{-i}) = i\}$. Thus, i must win and pay his bid, and since (G, S_N) is pruned, h is a terminal history. Finally, if the high bidder has bid θ^K and no active bidder has higher tie-breaking priority, then i must win and pay θ^K, and since (G, S_N) is pruned, h is a terminal history.

This completes the proof that, under the assumptions of Theorem 3.18, if (G, S_N) is credible and strategy-proof, then (G, S_N) is an ascending auction.

B.5.2 ascending → credible, strategy-proof

Now we show that if (G, S_N) is orderly, optimal, and an ascending auction, it is credible and strategy-proof.

That (G, S_N) is strategy-proof is straightforward. It remains to show that (G, S_N) is credible. As a preliminary, we prove that for any safe deviation $S_0' \in \mathcal{S}_0^*(S_0^G, S_N)$ and for any S_{-i}', S_i is a best response to (S_0', S_{-i}') in the messaging game.

First, consider information sets at which there is a unique action that places a bid. Take any i, I_i, and θ_i such that $\theta_i \in \Theta_i^{I_i}$. Recall that S_i requires that i quit if θ_i is strictly below the bid $b(I_i)$ at I_i, and that i places the bid if θ_i is above the least high bid consistent with reaching I_i. The least high bid consistent with reaching I_i is, formally,

$$\overset{\triangleright}{\min}\{\rho, \overset{\triangleright}{\min_{h \in I_i, j \neq i}} \underline{\theta}_j^h\} \tag{37}$$

And, since (G, S_N) is optimal and has threshold pricing,

$$b(I_i) \trianglelefteq \overset{\triangleright}{\min}\{\theta_i' \mid \theta_i' \triangleright \overset{\triangleright}{\min}\{\rho, \overset{\triangleright}{\min_{h \in I_i, j \neq i}} \underline{\theta}_j^h\}\} \tag{38}$$

For any safe deviation S_0' and for any S_{-i}', it is optimal for i to quit (upon reaching information set I_i) if $\theta_i \triangleleft \overset{\triangleright}{\min}\{\rho, \overset{\triangleright}{\min_{h \in I_i, j \neq i}} \underline{\theta}_j^h\}$. In particular, note that under (G, S_N), if i wins after reaching I_i, he pays at least $\overset{\triangleright}{\min}\{\rho, \overset{\triangleright}{\min_{h \in I_i, j \neq i}} \underline{\theta}_j^h\}$. Thus, for any safe deviation, i's best possible payoff upon placing a bid is no more than zero, so it is optimal to quit (which yields zero payoff).

For any safe deviation S_0' and for any S_{-i}', it is optimal for i to place a bid if θ_i is weakly above that bid. This is because i can quit if the required bid ever rises strictly above θ_i. Under any safe deviation, i cannot be charged more than θ_i unless he (at some

later point) bids more than θ_i. Thus, the worst possible payoff from placing a bid is zero, and the best possible payoff from quitting is zero.

By the above arguments and Equation 38, there are three possibilities at each I_i and $\theta_i \in \Theta_i^{I_i}$:

1. $\min\{\rho, \overset{\triangleright}{\underset{h \in I_i, j \neq i}{\min}} \underline{\theta}_j^h\} \lhd \theta_i$, in which case S_i requires that i place a bid, and this is a best response to (S_0', S_{-i}').

2. $\theta_i \lhd b(I_i)$, in which case S_i requires i to quit, and this is a best response to (S_0', S_{-i}').

3. $b(I_i) \unlhd \theta_i \lhd \overset{\triangleright}{\min}\{\rho, \underset{h \in I_i, j \neq i}{\min} \underline{\theta}_j^h\}$, in which case S_i is underdetermined, and both quitting now or placing the bid and quitting later are best responses to (S_0', S_{-i}').

Finally, consider information sets at which there are multiple bid-placing actions. In this case, under any safe deviation, i is sure to win if and only if he eventually bids the reserve - this implies that S_i remains a best response to any safe deviation.

Suppose now that (G, S_N) is an orderly ascending auction but not credible, so the auctioneer has a profitable safe deviation S_0'. Consider a corresponding G' in which the auctioneer 'commits openly' to that deviation, that is to say, G' such that S_0' runs G'. For all i, S_i is a best response to (S_0', S_{-i}), so (G', S_N) is also BIC. (We abuse notation slightly to use S_N as a strategy profile for G and G'. Every information set in G' has a corresponding information set in G, so it is clear what is meant.) By hypothesis, S_0' is a profitable deviation, so $\pi(G', S_N) > \pi(G, S_N)$, so (G, S_N) is not optimal. Thus, if (G, S_N) is orderly, optimal, and an ascending auction, then (G, S_N) is credible. This completes the proof of Theorem 3.18.

B.6 Theorem 3.22

B.6.1 virtual ascending → credible, strategy-proof

Suppose (G, S_N) is a virtual ascending auction. By inspection, (G, S_N) is strategy-proof. Moreover, S_i is a best response to any (S_0', S_{-i}) for $S_0' \in \mathcal{S}_0^*(S_0^G, S_N)$. (This requires only small modifications to the proof of Theorem 3.18, which we omit to avoid repetition.) Thus, if (G, S_N) is not credible, then there exists (G', S_N) that yields strictly higher expected revenue for the auctioneer, which implies that (G, S_N) is not optimal. Thus, if (G, S_N) is optimal and a virtual ascending auction, then (G, S_N) is credible.

B.6.2 credible, strategy-proof → virtual ascending

Propositions B.2, B.4, B.5, and B.6 pin down some details even when F_N is not symmetric. We start by proving an analogue to Proposition B.8.

Proposition B.9. *Assume F_N is regular and interleaved, and (G, S_N) is optimal and strategy-proof. If (G, S_N) is credible, then (G, S_N) is winner-pooling.*

Proof. As before, we will show that if (G, S_N) is not winner-pooling, then the auctioneer has a profitable safe deviation, so (G, S_N) is not credible. Let h^* be some history at which the winner-pooling property does not hold; we pick h^* such that, for all $h \prec h^*$, h is not a counterexample to winner-pooling. Since (G, S_N) is regular and interleaved, and the winner-pooling property held at all predecessors to h^*, Proposition 3.14 implies that for all i, either $W_i^{h^*} = \emptyset$ or $W_i^{h^*} = \{\theta_i \mid \eta_i(\theta_i) > \max(0, \overset{\triangleright}{\underset{j \neq i}{\max}} \, \eta_j(\underline{\theta}_j^{h^*}))\}$. Let us define i^*, $\theta_{i^*}^*$ and h^{**} as before.

The proof of Proposition B.8 works here with the following modifications: First, we define

$$\psi(i) = \underset{j \in N \setminus \{i\}}{\operatorname{argmax}} \{\eta_j(\theta_j^{K_j}) \mid W_j^{h^*} \neq \emptyset\} \tag{39}$$

Second, we say $\theta_{\psi(i)}$ i**-separates** at $\gamma \in \mathbb{R}$ if

$$\{\theta_i \mid \eta_i(\theta_i) \geq \gamma\} = \{\theta_i \mid \eta_i(\theta_i) \geq \eta_j(\theta_{\psi(i)})\} \tag{40}$$

Thirdly, we initialize $\beta := \min\{\eta_{i^*}(\theta_{i^*}^*), \eta_{\psi(i^*)}(\theta_{\psi(i^*)}^{K_{\psi(i^*)}})\}$ and specify the algorithm as:

Stage 1

1. Pick $\theta_{\psi(i^*)}$ that i^*-separates at β.

2. Simulate $(\theta_{\psi(i^*)}, \underline{\theta}_{N \setminus \{i^*, \psi(i^*)\}}^{h^*})$ against i^* starting from $\hat{h}_{i^*}$, until either $\eta_{i^*}(\underline{\theta}_{i^*}^{\hat{h}_{i^*}}) \geq \beta$ or $\hat{h}_{i^*} \in Z$.

3. If $\eta_{i^*}(\underline{\theta}_{i^*}^{\hat{h}_{i^*}}) \geq \beta$, then set $\beta := \underline{\theta}_{i^*}^{\hat{h}_{i^*}}$ and go to Stage 2.

4. Else, set $\hat{N} := \hat{N} \setminus \{i^*\}$,

$$\beta := \min_{i \neq i^*, \theta_i} \eta_i(\theta_i) \mid \theta_i \in W_i^{h^*} \tag{41}$$

and go to Stage 2.

Stage 2

1. If $\hat{N} = 1$, go to Stage 3.

2. Set $\hat{i} :\in \{i \in \hat{N} \mid \eta_i(\underline{\theta}_i^{\hat{h}_i}) < \beta\}$.

3. Pick $\theta_{\psi(\hat{i})}$ that $\hat{i}$-separates at β.

4. If $(\theta_{\psi(\hat{i})}, \underline{\theta}_{N \setminus \{\hat{i}, \psi(\hat{i})\}}^{h^*}) \notin \Theta_{-\hat{i}}^{\hat{h}_{\hat{i}}}$, set $\hat{h}_{\hat{i}} := \text{cousin}(\hat{h}_{\hat{i}}, (\theta_{\psi(\hat{i})}, \underline{\theta}_{N \setminus \{\hat{i}, \psi(\hat{i})\}}^{h^*}))$.

5. Simulate $(\theta_{\psi(\hat{i})}, \underline{\theta}^{h^*}_{N \setminus \{\hat{i}, \psi(\hat{i})\}})$ against $\hat{i}$ starting from $\hat{h}_{\hat{i}}$, until either $\eta_{\hat{i}}(\underline{\theta}^{\hat{h}_{\hat{i}}}_{\hat{i}}) \geq \beta$ or $\hat{h}_{\hat{i}} \in Z$.

6. If $\eta_{\hat{i}}(\underline{\theta}^{\hat{h}_{\hat{i}}}_{\hat{i}}) \geq \beta$, set $\beta := \eta_{\hat{i}}(\underline{\theta}^{\hat{h}_{\hat{i}}}_{\hat{i}})$ and go to Step 1 of Stage 2.

7. Else, set $\hat{N} := \hat{N} \setminus \{\hat{i}\}$ and go to Step 1 of Stage 2.

Stage 3

1. Set $\hat{i} := i \mid i \in \hat{N}$.

2. Pick $\theta_{\psi(\hat{i})}$ that $\hat{i}$-separates at β.

3. If $(\theta_{\psi(\hat{i})}, \underline{\theta}^{h^*}_{N \setminus \{\hat{i}, \psi(\hat{i})\}}) \notin \Theta^{\hat{h}_{\hat{i}}}_{-\hat{i}}$, set $\hat{h}_{\hat{i}} := \text{cousin}(\hat{h}_{\hat{i}}, (\theta_{\psi(\hat{i})}, \underline{\theta}^{h^*}_{N \setminus \{\hat{i}, \psi(\hat{i})\}}))$.

4. Simulate $(\theta_{\psi(\hat{i})}, \underline{\theta}^{h^*}_{N \setminus \{\hat{i}, \psi(\hat{i})\}})$ against $\hat{i}$ starting from $\hat{h}_{\hat{i}}$, until $\hat{h}_{\hat{i}} \in Z$.

5. Choose the outcome that corresponds to that terminal history, $x = g(\hat{h}_{\hat{i}})$, and terminate.

This deviating algorithm does not change the allocation; the object is kept if $\max_i \eta_i(\theta_i) \leq 0$, and allocated to $\text{argmax}_i \eta_i(\theta_i)$ otherwise (where $\text{argmax}_i \eta_i(\theta_i)$ is singleton since F_N is interleaved). Revenue is at least as high as under S^G_0, and strictly higher when $\theta_N = (\theta^*_{i^*}, \underline{\theta}^{h^*}_{-i^*})$.

It remains to check that the various steps of the algorithm are well-defined. We can pick separating types in Step 1 of Stage 1, because either $\beta = \eta_{\psi(i^*)}(\theta^{K_{\psi(i^*)}}_{\psi(i^*)})$ or $\beta = \eta_{i^*}(\theta^*_{i^*}) < \eta_{\psi(i^*)}(\theta^{K_{\psi(i^*)}}_{\psi(i^*)})$. In the first case, $\theta^{K_{\psi(i^*)}}_{\psi(i^*)}$ will i^*-separate at β. In the second case, since $\eta_{i^*}(\theta^*_{i^*}) > \eta_{\psi(i^*)}(\theta^1_{\psi(i^*)})$, by F_N interleaved there exists $\theta_{\psi(i^*)}$ that will i^*-separate at β.

When we pick separating types in Step 3 of Stage 2 and Step 2 of Stage 3, β is equal to $\eta_j(\theta_j)$ for some agent j where $\theta_j \in W^{h^*}_j$. Consider $\theta'_{\hat{i}} = \min\{\theta_{\hat{i}} \mid \eta_{\hat{i}}(\theta_{\hat{i}}) \geq \beta\}$. Since $\theta_j \in W^{h^*}_j$, it follows (by F_N regular and interleaved) that $\eta_{\hat{i}}(\theta'_{\hat{i}}) > \eta_{\psi(\hat{i})}(\theta^1_{\psi(\hat{i})})$. If $\eta_{\hat{i}}(\theta'_{\hat{i}}) < \eta_{\psi(\hat{i})}(\theta^{K_{\psi(\hat{i})}}_{\psi(\hat{i})})$, then, by F_N interleaved, there exists $\theta_{\psi(\hat{i})}$ that will $\hat{i}$-separate at β. If $\eta_{\hat{i}}(\theta'_{\hat{i}}) \geq \eta_{\psi(\hat{i})}(\theta^{K_{\psi(\hat{i})}}_{\psi(\hat{i})})$ then since β never exceeds $\min\{\eta_{\hat{i}}(\theta_{\hat{i}}) \mid \eta_{\hat{i}}(\theta_{\hat{i}}) \geq \eta_{\psi(\hat{i})}(\theta^{K_{\psi(\hat{i})}}_{\psi(\hat{i})})\}$, it follows that $\theta^{K_{\psi(\hat{i})}}_{\psi(\hat{i})}$ will $\hat{i}$-separate at β.

We can choose cousins (in Step 4 of Stage 2 and Step 3 of Stage 3) because F_N is regular and (G, S_N) is strategy-proof and optimal, by the same argument as in the proof of Theorem 3.18 that invokes Proposition B.6. Thus, the algorithm is well-defined, and produces a profitable safe deviation, which completes the proof. $\qquad\square$

With Proposition B.9 in hand, we now complete the proof that, under the assumptions of Theorem 3.22, if (G, S_N) is credible and strategy-proof, then (G, S_N) is a virtual ascending auction. Since F_N is regular and interleaved, the allocation and payments are

entirely pinned down by Proposition 3.14 and B.4. At type profile θ_N, agent i wins if and only if $\eta_i(\theta_i) > \max\{0, \max_{j \neq i} \eta_j(\theta_j)\}$, and pays $\min \theta_i' \mid \eta_i(\theta_i') > \max\{0, \max_{j \neq i} \eta_j(\theta_j)\}$.

Bidder i is **active** at h if $W_i^h \neq \emptyset$. There are three cases to consider:

1. An active bidder is called to play, and there is more than one active bidder.

2. An inactive bidder is called to play.

3. An active bidder is called to play, and there are no other active bidders.

Take any I_i and $h \in I_i$ such that an active bidder i is called to play, and there exists another active bidder, so $W_i^h \cap L_i^h \neq \emptyset$. Proposition B.9 implies that for all $\theta_i \in W_i^h$, $S_i(I_i, \theta_i) = S_i(I_i, \theta_i^{K_i})$. Thus, if bidder i does not play that action, then he has **quit**. The **bid** at I_i is the least type of i consistent with playing $S_i(I_i, \theta_i^{K_i})$, that is $\min\{\theta_i \in \Theta_i^h \mid S_i(I_i, \theta_i) = S_i(I_i, \theta_i^{K_i})\}$. By Proposition B.5, each bid is weakly more than the last bid that i placed.

By construction, all types strictly below the bid quit. Since (G, S_N) is optimal, i places a bid if $\eta_i(\theta_i) > \max\{0, \max_{j \neq i} \eta_j(b_j)\}$.

If bidder i quits, then either his virtual value is negative, or we have identified another bidder with a strictly higher virtual value. Thus, since (G, S_N) is pruned, only active bidders are called to play. Similarly, if i is the current high bidder at history h and there is another active bidder, then by Proposition B.9, all i's types who reach h take the same action, and (by (G, S_N) pruned) i is not called to play at h. Thus, if i is called to play at h, he is an active bidder who is not the current high bidder.

Suppose an active bidder i is called to play at h and is the unique active bidder. Since (G, S_N) is pruned, i is not the current high bidder, which implies that there is *no* high bidder. Let I_i be such that $h \in I_i$.

If there is another $h' \in I_i$ such that there is more than one active bidder at h', then we define bids and quitting at h as at h'. Otherwise, we define an action a as **quitting** if no type with a positive virtual value plays a, that is:

$$\neg \exists \theta_i \in \Theta_i^h \mid \eta_i(\theta_i) > 0 \text{ and } S_i(I_i, \theta_i) = a \tag{42}$$

For any non-quitting action a, the associated **bid** is:

$$\min\{\theta_i \mid \eta_i(\theta_i) > 0 \text{ or } [\theta_i \in \Theta_i^h \text{ and } S_i(I_i, \theta_i) = a]\} \tag{43}$$

In this case, $W_i^h \cap L_i^h = \emptyset$, so there can be multiple actions that place bids. Again, by Proposition B.5, each bid is weakly more than the last bid that i placed. The three conditions that specify what happens when the auction ends are entailed by optimality and threshold pricing (Proposition B.4). Thus, under the assumptions of Theorem 3.22, if (G, S_N) is credible and strategy-proof, then (G, S_N) is a virtual ascending auction.

B.7 Proposition 5.5

By inspection, first-price auctions are prior-free credible and static.

Suppose (G, S_N) is prior-free credible and static. Suppose there exist $\theta_i, \theta_{-i}, \theta'_{-i}$ such that i wins the object at (θ_i, θ_{-i}) and at (θ_i, θ'_{-i}), but $t_i(\theta_i, \theta_{-i}) < t_i(\theta_i, \theta'_{-i})$. We now construct a deviation: If the action profile is consistent with (θ_i, θ_{-i}), award the object to i and instead charge $t_i(\theta_i, \theta'_{-i})$. This deviation is always-profitable.

Consequently, there exists a function $\tilde{b}_i : \Theta_i \to \mathbb{R}$ such that if type θ_i wins, then i pays $\tilde{b}_i(\theta_i)$. Notably, this property holds everywhere, and not just almost everywhere.

We now partition i's actions into bidding actions $B_i = \{\tilde{b}_i(\theta_i) \mid \theta_i \in \Theta_i \text{ and } \exists \theta_{-i} : \tilde{y}(\theta_i, \theta_{-i}) = i\}$, and actions that decline. The same steps as in the proof of Theorem 3.7 establish that (G, S_N) is a first-price auction.

B.8 Proposition 5.6

With finite type-spaces, credible protocols are prior-free credible, so each "if" direction is immediate. In the proof of Theorem 3.18, we show that if (G, S_N) is strategy-proof but not an ascending auction, then there exists a safe deviation that is always-profitable, so (G, S_N) is not prior-free credible. So too for Theorem 3.22.

C   Extensions and other applications

C.1   Affiliated values

Here we use a discrete model of single-object auctions, as in Section 3.2. As is well-known, relaxing the independence assumption even slightly results in auctions that extract all bidder surplus (Cremer and McLean, 1988). The standard (static) mechanisms for full surplus extraction make each agent's payment depend on the other agents' types. The auctioneer can increase revenue by misrepresenting the other agents' types, so these mechanisms are not credible. Even using extensive forms does not generally permit credible full surplus extraction.

Definition C.1. (G, S_N) ***extracts full surplus*** *if it is BIC, has voluntary participation, and* $\pi(G, S_N) = \mathbf{E}_{\theta_N}[\max\{0, \max_{i \in N} \theta_i\}]$.

Proposition C.2. *The Cremer and McLean (1988) conditions are not sufficient for the existence of a credible protocol that extracts full surplus.*

Proof. There are two bidders i and j, each with two possible values $0 < \theta_i < \theta'_i < \theta_j < \theta'_j$. The joint distribution of types is $f_N(\theta_i, \theta_j) = f_N(\theta'_i, \theta'_j) = 1/3$, $f_N(\theta_i, \theta'_j) = f_N(\theta'_i, \theta_j) = 1/6$, which satisfies the full rank condition of Cremer and McLean (1988) Theorem 2.

For a given protocol (G, S_N), consider the induced allocation rule $\tilde{y}$ and transfer rule $\tilde{t}_N$. Suppose (G, S_N) is credible and extracts full surplus. By Propositions 2.3 and 2.7, it is without loss of generality to restrict (G, S_N) so that after j is called to play once, he is never called to play again.

Take any information set I_j at which j is called to play. Since (G, S_N) is credible, for each action that j takes at I_j, there is a unique transfer from j if j wins (Proposition 3.19). Since (G, S_N) extracts full surplus, j wins no matter whether he plays $S_j(I_j, \theta_j)$ or $S_j(I_j, \theta'_j)$. Since (G, S_N) is BIC, j's transfer after playing $S_j(I_j, \theta_j)$ is the same as j's transfer after playing $S_j(I_j, \theta'_j)$.

This argument applies to every information set at which j is called to play, so j's transfer does not depend on his own type; $\tilde{t}_j(\theta_i, \theta_j) = \tilde{t}_j(\theta_i, \theta'_j)$ and $\tilde{t}_j(\theta'_i, \theta_j) = \tilde{t}_j(\theta'_i, \theta'_j)$.

Since j always wins the object, the auctioneer can safely deviate to communicate with j as though i's type is θ_i or as though i's type is θ'_i. Since (G, S_N) is credible, j's transfer does not depend on i's type; $\tilde{t}_j(\theta_i, \theta_j) = \tilde{t}_j(\theta'_i, \theta_j)$. Thus, j's transfer is some constant $\bar{t}_j$ across all type profiles. $\theta_j - \bar{t}_j = 0$, so $\theta'_j - \bar{t}_j > 0$, and (G, S_N) does not extract full surplus, a contradiction. $\qquad \square$

Optimal auctions with correlation are better-behaved if we additionally require *ex post* incentive compatibility and *ex post* individual rationality.[41] The virtual values machinery generalizes, and a modified ascending auction is optimal under some standard assumptions (Roughgarden and Talgam-Cohen, 2013). That modified ascending auction is credible. We now make the claim precisely.

Consider some probability mass function $f_N : \Theta_N \to [0, 1]$. We assume symmetric type spaces, $K_i = K_j = K$ and $\theta^k_i = \theta^k_j$ for all i, j, k, as well as affiliated types (Milgrom and Weber, 1982).

Definition C.3. f_N is **symmetric** if its value is equal under any permutation of its arguments. f_N is **affiliated** if for all θ_N, θ'_N:

$$f_N(\theta_N \vee \theta'_N) f_N(\theta_N \wedge \theta'_N) \geq f_N(\theta_N) f_N(\theta'_N) \qquad (44)$$

where $\vee$ *is the component-wise maximum and* $\wedge$ *the component-wise minimum.*

For a protocol (G, S_N), let $\tilde{y}_i^{G, S_N}(\theta_N)$ be an indicator variable equal to 1 if i wins the object at θ_N and 0 otherwise. (We suppress the independence on (G, S_N) to ease notation.)

---

[41]*Ex post* incentive compatibility and *ex post* individual rationality are implied by strategy-proofness and voluntary participation (Definition 3.1). For extensive forms, *ex post* incentive compatibility and strategy-proofness are not equivalent. An opponent strategy profile S_{-i} consists of complete contingent plans of action. *Ex post* incentive compatibility in effect considers only plans 'consistent with' some opponent type profile θ_{-i}.

Definition C.4. (G, S_N) *is* **optimal among ex post auctions** *if it maximizes expected revenue subject to the constraints:*

1. *Ex post incentive compatibility. For all $i, \theta_i, \theta_i', \theta_{-i}$:*

$$\theta_i \tilde{y}_i(\theta_i, \theta_{-i}) - \tilde{t}_i(\theta_i, \theta_{-i}) \geq \theta_i \tilde{y}_i(\theta_i', \theta_{-i}) - \tilde{t}_i(\theta_i', \theta_{-i}) \tag{45}$$

2. *Ex post individual rationality. For all i, θ_i, θ_{-i}:*

$$\theta_i \tilde{y}_i(\theta_i, \theta_{-i}) - \tilde{t}_i(\theta_i, \theta_{-i}) \geq 0 \tag{46}$$

Definition C.5. *The* **conditional virtual value** *of θ_i^k given θ_{-i} is:*

$$\eta_i(\theta_i^k | \theta_{-i}) \equiv \theta_i^k - \frac{1 - F_i(\theta_i^k | \theta_{-i})}{f_i(\theta_i^k | \theta_{-i})}(\theta_i^{k+1} - \theta_i^k) \tag{47}$$

where $f_i(\cdot | \theta_{-i})$ is the conditional distribution of θ_i given θ_{-i} and $F_i(\cdot | \theta_{-i})$ is the conditional cumulative distribution. f_N is **regular** *if, for all i and θ_{-i}, $\eta_i(\theta_i | \theta_{-i})$ is strictly increasing in θ_i.*

We now define a modified ascending auction. When there is only one bidder left, the auctioneer sets a reserve so that she only sells to types with a positive conditional virtual value.[42] That reserve depends on the final bids from the bidders who quit.

Definition C.6. (G, S_N) *is a* **quirky ascending auction** *if:*

1. *All bidders start as active, with initial bids $(b_i)_{i \in N} := (\theta_i^1)_{i \in N}$.*

2. *Whenever there is more than one active bidder, some active bidder i is called to play, where $b_i \leq \max_{j \neq i} b_j$.*

   (a) *i chooses between two actions; he can either raise b_i by one increment[43] or quit.*

   (b) *If i quits then he is no longer active.*

3. *When there is exactly one active bidder i, if $\eta_i(b_i | b_{-i}) \leq 0$, i chooses to either raise his bid to $\min b_i' \mid \eta_i(b_i' | b_{-i}) > 0$ or quit. Otherwise i wins and pays b_i.*

4. *Inactive bidders do not win the object, and have zero transfers.*

5. *S_i specifies that i bids b_i if and only if $\theta_i \geq b_i$.*

Proposition C.7. *Assume f_N is symmetric, affiliated, and regular. If (G, S_N) is a quirky ascending auction, then it is optimal among ex post auctions and is credible.*

---

[42]This definition is due to Roughgarden and Talgam-Cohen (2013), and differs only in that our construction is for finite type spaces to allow the use of extensive game forms.

[43]i.e. from θ_i^k to θ_i^{k+1}, where we set $\theta_i^{K+1} > \theta_i^K$.

Proof. Define $\nu(\theta_i, \theta_{-i}) = \theta_i \tilde{y}_i(\theta_i, \theta_{-i}) - \tilde{t}_i(\theta_i, \theta_{-i})$.

We can use the same method as in Elkind (2007) to derive an upper bound on $\nu(\theta_i, \theta_{-i})$ under *ex post* incentive compatibility and *ex post* individual rationality, namely:

$$\nu(\theta_i^k, \theta_{-i}) \geq \sum_{l=2}^{k} \tilde{y}_i(\theta_i^{l-1}, \theta_{-i})(\theta_i^l - \theta_i^{l-1}) \tag{48}$$

This implies a bound on i's expected utility conditional on θ_{-i}, namely

$$\mathbf{E}_{\theta_i}[\nu(\theta_i^k, \theta_{-i}) \mid \theta_{-i}] \geq \sum_{k=2}^{K} f_i(\theta_i^k) \sum_{l=1}^{k} \tilde{y}_i(\theta_i^{l-1}, \theta_{-i})(\theta_i^l - \theta_i^{l-1})$$

$$= \sum_{k=1}^{K} f_i(\theta_i^k | \theta_{-i}) \frac{1 - F_i(\theta_i^k | \theta_{-i})}{f_i(\theta_i^k | \theta_{-i})}(\theta_i^{k+1} - \theta_i^k)\tilde{y}_i(\theta_i^k, \theta_{-i}) \tag{49}$$

which gives an upper bound on expected revenue

$$\pi(G, S_N) = \sum_{i \in N} \mathbf{E}_{\theta_N}[\theta_i \tilde{y}_i(\theta_N) - \nu(\theta_i, \theta_{-i})]$$

$$= \sum_{i \in N} \mathbf{E}_{\theta_{-i}}[\mathbf{E}_{\theta_i}[\theta_i \tilde{y}_i(\theta_N) - \nu(\theta_i, \theta_{-i}) \mid \theta_{-i}]]$$

$$\leq \sum_{i \in N} \mathbf{E}_{\theta_{-i}}[\mathbf{E}_{\theta_i}[\eta_i(\theta_i | \theta_{-i})\tilde{y}_i(\theta_N) \mid \theta_{-i}]] = \mathbf{E}_{\theta_N}\left[\sum_{i \in N} \eta_i(\theta_i | \theta_{-i})\tilde{y}_i(\theta_N)\right] \tag{50}$$

Moreover, the above equation holds with equality if the local downward incentive constraints bind and the participation constraints bind for the lowest type, where these constraints are conditional on θ_{-i}.

We now apply the argument in Roughgarden and Talgam-Cohen (2013), which is written for continuous densities but works also for the discrete case. For the reader's convenience, we repeat it here.

Lemma C.8. *If f_N is affiliated and $\theta_j < \theta_j'$, then $\eta_i(\theta_i | \theta_j, \theta_{N \setminus \{i,j\}}) \geq \eta_i(\theta_i | \theta_j', \theta_{N \setminus \{i,j\}})$*

By affiliation, $F_i(\theta_i | \theta_j', \theta_{N \setminus \{i,j\}})$ dominates $F_i(\theta_i | \theta_j, \theta_{N \setminus \{i,j\}})$ in terms of hazard rate (Krishna, 2010, Appendix D), i.e.

$$\frac{1 - F_i(\theta_i | \theta_j, \theta_{N \setminus \{i,j\}})}{f_i(\theta_i | \theta_j, \theta_{N \setminus \{i,j\}})} \leq \frac{1 - F_i(\theta_i | \theta_j', \theta_{N \setminus \{i,j\}})}{f_i(\theta_i | \theta_j', \theta_{N \setminus \{i,j\}})} \tag{51}$$

which implies $\eta_i(\theta_i | \theta_j, \theta_{N \setminus \{i,j\}}) \geq \eta_i(\theta_i | \theta_j', \theta_{N \setminus \{i,j\}})$. This proves Lemma C.8.

Lemma C.9. *Assume f_N is symmetric, regular, and affiliated. For all $\theta_{N \setminus \{i,j\}}$, if $k \geq k'$, then $\eta_i(\theta_i^k | \theta_{N \setminus \{i,j\}}, \theta_j^{k'}) \geq \eta_j(\theta_j^{k'} | \theta_{N \setminus \{i,j\}}, \theta_i^k)$.*

$$\eta_i(\theta_i^k | \theta_{N \setminus \{i,j\}}, \theta_j^{k'}) \geq \theta_i^{k'} - \frac{1 - F_i(\theta_i^{k'} | \theta_{N \setminus \{i,j\}}, \theta_j^{k'})}{f_i(\theta_i^{k'} | \theta_{N \setminus \{i,j\}}, \theta_j^{k'})} (\theta_i^{k'+1} - \theta_i^{k'})$$

$$\geq \theta_i^{k'} - \frac{1 - F_i(\theta_i^{k'} | \theta_{N \setminus \{i,j\}}, \theta_j^{k})}{f_i(\theta_i^{k'} | \theta_{N \setminus \{i,j\}}, \theta_j^{k})} (\theta_i^{k'+1} - \theta_i^{k'}) = \eta_j(\theta_j^{k'} | \theta_{N \setminus \{i,j\}}, \theta_i^{k}) \quad (52)$$

where the first inequality follows from regularity, the second inequality follows from Lemma C.8, and the equality follows from symmetry. This proves Lemma C.9.

By Lemma C.9, the right-hand side of Equation 50 is maximized by, at each θ_N, selling to some agent in $\arg\max_i \theta_i$ if $\max_i \eta_i(\theta_i | \theta_{-i}) > 0$, and keeping the object otherwise. The quirky ascending auction does this, and additionally the local incentive constraints bind downward and the participation constraint of the lowest type binds, so the left-hand side of Equation 50 is equal to the right-hand side. Thus, any quirky ascending auction is optimal among *ex post* mechanisms.

It remains to prove that the quirky ascending auction is credible. Once more, note that S_i is a best response to any safe deviation by the auctioneer. Under any safe deviation, if $b_i \leq \theta_i$, then bidder i's utility is non-negative if he continues bidding according to S_i, and zero if he quits now. If $b_i > \theta_i$, then bidder i's utility is non-positive if he continues bidding, and zero if he quits now. Thus, S_i is a best-response to any safe deviation by the auctioneer, regardless of θ_{-i}. For any safe deviation S_0', the corresponding protocol (G', S_N) is *ex post* incentive compatible and *ex post* individually rational. Suppose that S_0' is profitable, so (G', S_N) yields strictly more expected revenue than (G, S_N). Since (G, S_N) is optimal among *ex post* mechanisms, we have the desired contradiction. $\qquad\square$

C.2 Auctions with matroid constraints

So far we have assumed that in each feasible allocation there is at most one winner. Suppose instead that multiple bidders can be satisfied at once; that is, the feasible sets of winners are a family $\mathcal{F} \subseteq 2^N$. Each bidder's type is independently distributed according to $f_i : \Theta_i \to (0, 1]$, where i's utility at allocation $Y \in \mathcal{F}$ is $\theta_i 1_{i \in Y} - t_i$. Each bidder observes whether or not he is in the allocation, and his own transfer.

Definition C.10. $\mathcal{F}$ *is a **matroid** if:*

  1. $\emptyset \in \mathcal{F}$

  2. If $Y' \subset Y$ *and* $Y \in \mathcal{F}$*, then* $Y' \in \mathcal{F}$*.*

  3. For any $Y, Y' \in \mathcal{F}$*, if* $|Y| > |Y'|$*, then there exists* $i \in Y \setminus Y'$ *such that* $Y' \cup \{i\} \in \mathcal{F}$*.*

Here are some examples of matroids:

1. The auctioneer can sell at most k items; that is, $Y \in \mathcal{F}$ if and only if $|Y| \leq k$.

2. There are incumbent bidders and new entrants. The auctioneer sells k licenses, and at most l licenses can be sold to incumbents.

3. The auctioneer is selling the edges of a graph. Each edge is demanded by exactly one bidder, and the auctioneer can sell any set of edges that is acyclic.

4. There are bands of spectrum $\{1, \ldots, K\}$, and each band k is acceptable to a subset of bidders N_k. Each bidder is indifferent between bands that he finds acceptable. At most one bidder can be assigned to each band.

Proposition C.11. *If $\mathcal{F}$ is a matroid, then there exists a credible strategy-proof optimal protocol.*

We describe this protocol informally, since the fine details parallel Definition 3.21, and our construction draws heavily on Bikhchandani et al. (2011) and Milgrom and Segal (2017). Each bidder's starting bid is equal to his lowest possible type. We score bids according to their ironed virtual values, and keep track of a set of active bidders $\hat{N}$.

Bidder i is **essential** at $\hat{N}$ if, for all $Y \subseteq \hat{N}$, if $Y \in \mathcal{F}$, then $Y \cup \{i\} \in \mathcal{F}$. At each step, we choose an active bidder i whose score is minimal in $\hat{N}$. If i's score is positive and i is essential at $\hat{N}$, then we guarantee that i is in the allocation and charge him his current bid, removing him from $\hat{N}$. Otherwise, i chooses to either raise his bid until his score is positive and no longer minimal, or quit (in which case he is also removed from $\hat{N}$). The auction ends when $\hat{N} = \emptyset$.

The above protocol outputs the same allocation as a greedy algorithm that starts with the empty set and at each step adds a bidder with the highest ironed virtual value among those that can be feasibly added, until no bidders with positive ironed virtual values can be added (we prove this in the Appendix). By a standard result in combinatorial optimization (Hartline, 2016, p.134), this greedy algorithm maximizes the ironed virtual value when $\mathcal{F}$ is a matroid. Given that the relevant participation constraints and incentive constraints bind, maximizing ironed virtual values implies that the protocol is optimal (Elkind, 2007).

The auction we described is credible, for the same reasons as before: Since truthful bidding is best response to any safe deviation, if the auctioneer could improve revenue by a safe deviation, she could have committed from the beginning to an alternative mechanism and increased revenue. Since the original protocol was optimal, we have a contradiction. The formal proof of Proposition C.11 follows.

Proof. Suppose we construct ironed virtual values for discrete type spaces as in Elkind (2007). Let the protocol break ties according to some fixed order on N, when two bids have the same ironed virtual value.

Fix some type profile θ_N. Let us label agents in decreasing order of ironed virtual values, $\{1, 2, \ldots, n\}$, breaking ties according to the fixed order. Let $\{i^1, i^2, \ldots, i^J\}$ be the set picked by the greedy algorithm, in order of selection (where the algorithm breaks ties using the same fixed order). We must show that the protocol described in Subsection C.2 results in the same allocation.

Take the greedy algorithm's jth pick, $i^j = k$. We will show that k is essential with respect to the set of active bidders $\hat{N}$ before k is asked to place a bid strictly above his type. Consider any step of the algorithm at which k, if not essential, would be asked to place a bid strictly above his type. At this step, $\hat{N} \subseteq \{1, 2, \ldots, k\}$, since bidders with lower ironed virtual values have either been put in the allocation or quit (and similarly bidders with equal ironed virtual values but who lose ties to k).

Take any $Y \subseteq \{1, 2, \ldots, k\}$ such that $Y \in \mathcal{F}$. We assert that $Y \cup \{k\} \in \mathcal{F}$. There are two cases: either $|Y| \geq j$ or $|Y| < j$.

Suppose $|Y| \geq j > |\{i^1, \ldots, i^{j-1}\}|$. Since $\mathcal{F}$ is a matroid, there exists $l \in Y \setminus \{i^1, \ldots, i^{j-1}\}$, such that $\{i^1, \ldots, i^{j-1}\} \cup \{l\} \in \mathcal{F}$. If $Y \cup \{k\} \notin \mathcal{F}$, then $k \notin Y$, so $k \neq l$. Thus, $i^j = k$ is not the greedy algorithm's jth pick, a contradiction.

If $|Y| < j = |\{i^1, \ldots, i^j\}|$, then since $\mathcal{F}$ is a matroid, there exists $l \in \{i^1, \ldots, i^j\} \setminus Y$ such that $Y \cup \{l\} \in \mathcal{F}$ and $Y \cup \{l\} \subseteq \{1, \ldots, k\}$. Thus, we can find $Y' \supset Y$ such that $|Y'| = j$, $Y' \subseteq \{1, \ldots, k\}$, and $Y' \in \mathcal{F}$. From the argument in the previous paragraph, $Y' \cup \{k\} \in \mathcal{F}$, and, since $\mathcal{F}$ is a matroid, $Y \cup \{k\} \in \mathcal{F}$.

We have now established that, since $\hat{N} \subseteq \{1, 2, \ldots, k\}$, k is essential with respect to $\hat{N}$. Thus the jth pick of the greedy algorithm is in the allocation produced by the protocol. This argument holds for all j, so the protocol's allocation is a superset of $\{i^1, \ldots, i^J\}$. But the protocol only sells to bidders with positive ironed virtual values, so its allocation is exactly $\{i^1, \ldots, i^J\}$, and the protocol is optimal.

Finally, note that for any safe deviation, each bidder's 'truth-telling' strategy is a best response. That is, each bidder should keep bidding so long as the price he faces is weakly below his value, and quit otherwise. Thus, if the auctioneer has a profitable safe deviation, then the original protocol is not optimal, a contradiction. $\qquad\square$

C.3 Public goods provision

A social planner chooses whether to provide a public good with integer cost $c > 0$. An outcome consists of an allocation $y \in \{0, 1\}$ and transfers from each agent $(t_i)_{i \in N}$. Agent i's utility is $\theta_i y - t_i$, where $\Theta_i = \{0, 1, 2, \ldots, K\}$. The efficient allocation is:

$$y^*(\theta_N) = \begin{cases} 1 & \text{if } \sum_i \theta_i - c \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{53}$$

The planner wants to choose the efficient allocation, but also receives a small benefit

from having higher transfers. Formally, for $\gamma \in (0, \frac{1}{|N|K})$:

$$u_0(y, t_N, \theta_N) = 1_{y=y^*(\theta_N)} + \gamma \sum_i t_i \qquad (54)$$

Each agent observes whether the public good is provided, as well as his own transfer. Under mild conditions, if a protocol is static, strategy-proof and efficient, then it is not prior-free credible. The conditions require that, when i's type is low, slightly raising i's type might affect the efficient allocation, and also that when i's type is high, slightly raising i's type might affect the efficient allocation. The key intuition is that when the planner has a preference for transfers, prior-free credibility implies that i's transfers are measurable with respect to the allocation rule, which prevents the use of threshold prices.

Proposition C.12. *Assume there exist $\theta_i < \theta_i' < \theta_i'' < \theta_i'''$, $\theta_{N\backslash i}$, and $\theta_{N\backslash i}'$ such that:*

1. *$\theta_i + \sum_{j\neq i} \theta_j < c < \theta_i' + \sum_{j\neq i} \theta_j$*

2. *$\theta_i'' + \sum_{j\neq i} \theta_j' < c < \theta_i''' + \sum_{j\neq i} \theta_j'$*

There does not exist (G, S_N) that is static, strategy-proof, efficient, and prior-free credible.

Proof. Suppose not. Since (G, S_N) is prior-free credible and efficient, there exist unique transfers $t_i^1(\theta_i'), t_i^1(\theta_i''), t_i^1(\theta_i''')$ that are paid if the public good is provided and i has the corresponding type. Since (G, S_N) is strategy-proof, these transfers are all equal $t_i^1(\theta_i') = t_i^1(\theta_i'') = t_i^1(\theta_i''') = t_i^1$. Similarly, there exist unique transfers $t_i^0(\theta_i) = t_i^0(\theta_i') = t_i^0(\theta_i'') = t_i^0$ that are paid if the public good is not provided and i has the corresponding type.

(G, S_N) is strategy-proof and efficient, so $\theta_i' - t_i^1 \geq -t_i^0$, which implies $\theta_i'' - t_i^1 > -t_i^0$. Thus, when i's opponents play as though their types are $\theta_{N\backslash i}'$, type θ_i'' can profitably imitate θ_i''', a contradiction. $\square$

If we allow non-static mechanisms, then there exist prior-free credible efficient protocols when $|N| = 2$. Our construction treats agents asymmetrically; i declares whether he is willing to buy the public good at a given price, and at each step the price rises. The public good is withheld if i quits. j declares whether he is willing to forgo the public good in return for payment, and at each step the payment offered to j falls. The public good is provided if j quits. We coordinate the price faced by i and the payment offered to j so that the public good is provided if and only if their values exceed the cost of provision. Formally, initialize $b_i := 0$, $b_j := K$.

1. If $b_i + b_j < c$, ask i to raise his bid to $c - b_j$ or quit.

    (a) i raises his bid if and only if $\theta_i \geq c - b_j$

    (b) If i quits, then the public good is not provided, $t_i = 0$ and $t_j = -b_j$.

2. If $b_i + b_j \geq c$, ask j to lower his bid to $c - b_i - 1$ or quit.

   (a) j lowers his bid if and only if $\theta_j \leq c - b_i - 1$

   (b) If j quits, then the public good is provided, $t_i = b_i$ and $t_j = 0$.

3. Go to step 1.

The above protocol for two agents is efficient, strategy-proof, and prior-free credible. Holding fixed the parameters c and K, at any point in the messaging game, for each agent there is at most one query that can be safely sent to him. Observe that, for any safe deviation, at any point in the messaging game, the planner knows only a lower bound for i's type $\underline{\theta}_i$ and an upper bound for j's type $\overline{\theta}_j$. If $\underline{\theta}_i + \overline{\theta}_j < c$, and the planner queries j, then j quits if his type is $\overline{\theta}_j$, causing the public good to be inefficiently provided when the type profile is $(\underline{\theta}_i, \overline{\theta}_j)$. If $\underline{\theta}_i + \overline{\theta}_j \geq c$, and the planner queries i, then i quits if his type is $\underline{\theta}_i$, causing the public good to be inefficiently withheld when the type profile is $(\underline{\theta}_i, \overline{\theta}_j)$. Any safe deviation can change revenue by no more than $2K$ so, since γ is small, the protocol is prior-free credible.

Since this protocol treats agents asymmetrically, there is no easy extension to three or more agents. For that case, it is an open question whether strategy-proofness, efficiency, and prior-free credibility are compatible.