# Credible Mechanisms*

Mohammad Akbarpour[†]    Shengwu Li[‡]

July 3, 2018

## 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.

**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.

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 communicate privately from a distance. 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".[1] 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".[2]

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.[3] Second, auctioneers want to prevent col-

---

[1] *Why auction rooms seem empty these days*, The Wall Street Journal, June 15 2014.

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

[3] Dworczak (2017) studies how post-auction strategic interactions affect what information auctioneers should publicly release. A participant in a spectrum auction in India reported, "Those in the war room

lusion. Thus, in some auctions, bidders are forbidden from conferring - they must submit their bids only to the auctioneer. For instance, the 2017 US wireless spectrum auction's rules state that bidders "are prohibited from communicating directly or indirectly any incentive auction applicant's bids or bidding strategies".[4] In auctions for art or wine, the auctioneer typically does not reveal the identity of the winner, since keeping this private gives bidders incentives to defect from collusive arrangements (Ashenfelter, 1989). 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

---

had to sign non-disclosure agreements to ensure we wouldn't talk about auction strategy and discussions to any one, during or after the auction." (*Auction action: How telcos fought the bruising battle for spectrum*, The Economic Times, March 30 2015.)

[4]Section 1.2205(b)

[5]*How SSPs use deceptive price floors to squeeze ad buyers*, Digiday, Sep 13 2017

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.

Having defined the framework, we now turn to our main application. 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).[7] The first-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.[8] 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). However, many real-world auctioneers persist in running first-price auctions and ascending auctions, despite

---

[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]The Dutch (descending) auction, in which the price falls until one bidder claims the object, is less prevalent (Krishna, 2010, p.2).

[8]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".

the invention of this (apparently) superior format (Rothkopf et al., 1990). Why is this so?

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. Thus, the second-price auction (with reserve) is the unique strategy-proof static optimal auction, by the Green-Laffont-Holmström theorem (Green and Laffont, 1977; Holmström, 1979).

The results that follow require us to bridge the discrete world of extensive game forms and the continuous world of optimal auctions. We suppress these technicalities in the introduction, but the reader should be aware that the first result holds 'in the limit', as a finite grid type space becomes arbitrarily fine.

Our first result is as follows: The first-price auction (with reserve) is the unique credible static 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, too soon. 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?

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 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. The credible static auctions are now *twin-bid* auctions. This is a larger class
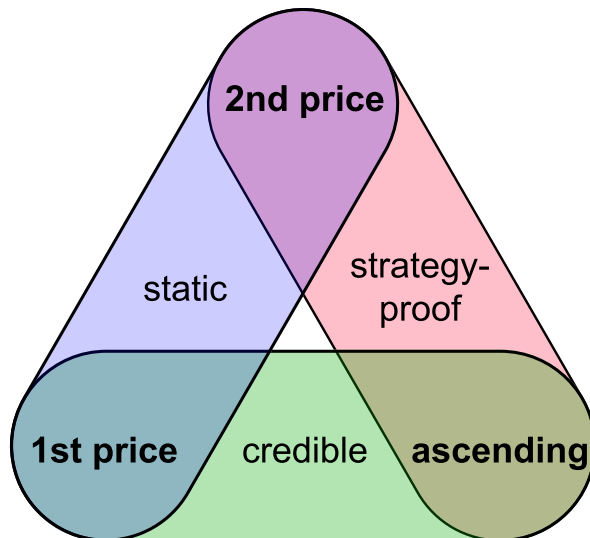
Figure 1: An auction trilemma: In the class of optimal auctions in which only the winner makes transfers, no auction is static, strategy-proof, and credible. Picking two out of three properties uniquely characterizes each standard format.

that includes all-pay auctions and first-price auctions with entry fees. In a twin-bid auction, each bid that an agent can place is associated with a weakly lower bid to be paid if that agent loses. Each agent pays his bid if he wins, and pays its lower 'twin' if he loses. If an agent wins the auction, then the difference between his bid and its twin is at least as large as the difference for any other agent. Under mild assumptions, twin-bid auctions are not strategy-proof.

We also relax the assumption that value distributions are symmetric. 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

The environment consists of:

1. A finite set of agents, $N$.

2. A set of outcomes, $X$.

3. A finite type space, $\Theta_N = \times_{i \in N} \Theta_i$.

4. A joint probability distribution $D : \Theta_N \to [0, 1]$.

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

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

The partition $\Omega_i$ 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.[9]

---

[9]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.

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. Let $\mathcal{G}$ denote the set of all extensive game forms with consequences in $X$ with finitely many histories and perfect recall.

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 $h$ | $\sigma(h)$ | |
| actions available at $I_i$ | $A(I_i)$ | |

$S_i$ denotes a (pure) strategy: For each information set where agent $i$ is called to play and each type of $i$ , $S_i$ chooses an action $S_i(I_i, \theta_i) \in A(I_i)$. $(S_i)_{i \in N} \equiv S_N$ denotes a strategy profile for the agents, for some $G \in \mathcal{G}$.

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 $x^G(S_N, \theta_N)$ denote the outcome in $G$, when agents play according to $S_N$ and the type profile is $\theta_N$. Let $u_i^G(S_N, \theta_N) \equiv u_i(x^G(S_N, \theta_N), \theta_N)$.

**Definition 1.** $(G, S_N)$ is ***Bayesian incentive-compatible (BIC)*** if, for all $i \in N$,

$$S_i \in \underset{S_i'}{\operatorname{argmax}} \, \mathbb{E}_{\theta_N}[u_i^G(S_i', S_{-i}, \theta_N)] \tag{1}$$

## 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 occur with zero probability under a BIC protocol $(G, S_N)$, then we can delete those histories from $G$ 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.[10] 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 $z(S_N, \theta_N)$ denote the terminal history that results from $(S_N, \theta_N)$. Formally:

**Definition 2.** $(G, S_N)$ is **pruned** if, for any history $h$:

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

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

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

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

Hence, from this point onwards we restrict attention to pruned $(G, S_N)$. Since every information set is reached with positive probability, any Bayes-Nash equilibrium in a pruned protocol survives equilibrium refinements that restrict off-path beliefs.

---

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

## 2.3 A messaging game

We now explicitly model the auctioneer[11] as a player (denoted $0$). The auctioneer has utility $u_0 : X \times \Theta_N \to \mathbb{R}$.

Our goal is to study surreptitious deviations by the auctioneer. To do so, we must first take a stand on what each player knows at each point in the mechanism. Consider the messaging game, defined thus:

1. The auctioneer chooses to:

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

    (b) Or: Go to step 2.

2. The auctioneer chooses some agent $i \in N$ and sends a message along with a set of acceptable replies $(m, R)$.

3. Agent $i$ privately observes $(m, R)$ and chooses $r \in R$.

4. The auctioneer privately observes $r$.

5. Go to step 1.

Assume the auctioneer has an arbitrarily rich message space $M$, from which she sends messages $m \in M$ and allows replies $R \in 2^M \setminus \emptyset$. At any prior round $k$, the auctioneer messaged $i^k \in N$ with $(m^k, R^k)$, and received reply $r^k \in R^k$.

Let $\mathcal{S}_0$ denote the set of auctioneer pure strategies. A strategy for the auctioneer specifies what to do next, as a function of the entire history of communications: $S_0((i^k, m^k, R^k, r^k)_{k=1}^t) \in (N \times M \times (2^M \setminus \{\emptyset\})) \cup X$. We restrict these to send finitely many messages and (for each message) to allow finitely many replies.

A strategy for agent $i$ specifies what reply to give at each point in the messaging game. Let $m_i^k$ denote the $k$th message that the auctioneer sent to agent $i$, and similarly for $R_i^k$ and $r_i^k$. Upon receiving a query, $i$ chooses a reply, which depends on the past communication between $i$ and the auctioneer $(m_i^\tau, R_i^\tau, r_i^\tau)_{\tau=1}^{k-1}$, the current query $(m_i^k, R_i^k)$, and $i$'s type; that is, $S_i((m_i^\tau, R_i^\tau, r_i^\tau)_{\tau=1}^{k-1}, m_i^k, R_i^k, \theta_i) \in R_i^k$.[12]

Let $T_i$ denote the total number of messages sent to $i$ at the end of the messaging game. For any $S_0 \in \mathcal{S}_0$ and any $S_N$, $(S_0, S_N)$ results in some sequence of communication between the auctioneer and agent $i$, $o_i^c \equiv (m_i^k, R_i^k, r_i^k)_{k=1}^{T_i}$ and some outcome $x$. Let $o_i^x$ denote $\omega_i \in \Omega_i \mid x \in \omega_i$; this is what $i$ directly observes about the outcome that the

---

[11]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.

[12]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.

auctioneer selected, as implied by the partition $\Omega_i$. $o_i = (o_i^c, o_i^x)$ denotes an *observation* for agent $i$. $o_i(S_0, S_N, \theta_N)$ denotes the unique observation resulting from $(S_0, S_N)$, when the type profile is $\theta_N$.

Given some extensive-form mechanism $G$, the auctioneer can 'carry out' $G$ in the messaging game. That is, she can, starting from the initial history in $h_\emptyset \in H$, contact the agent who is called to play $P(h_\emptyset)$, and send a message that corresponds to that information set, and allow replies that correspond to the actions. Upon receiving a reply, she can then contact the next agent, sending messages (information sets) and allowing replies (actions) as specified by $G$. Eventually, she will reach a terminal history $h \in Z \subseteq H$, whereupon she can choose the corresponding outcome and end the messaging game. We say that this 'rule-following' strategy **runs** $G$.

To define this formally, we use a one-to-one function $\lambda$, that takes as an input an information set in $G$ or an action in $G$, and outputs an element of the message space $M$.

**Definition 3.** *Take any $G = \langle H, \prec, P, A, \mathcal{A}, (\mathcal{I}_i)_{i \in N}, g \rangle$. $S_0$ **runs** $G$ if there exists a one-to-one function $\lambda : (\cup_{i \in N} \mathcal{I}_i) \cup A \to M$ such that $S_0$ is described by the following algorithm, where we initialize $h := h_\emptyset$.*

1. *If $h \in Z$, terminate and select $x = g(h)$.*

2. *Else:*

    (a) *Choose agent $P(h)$ and send $(m, R) = (\lambda(I_i), \lambda(A(I_i)))$ for $I_i$ such that $h \in I_i$.*

    (b) *Upon receiving $r \in R$, choose $h'$ such that $\mathcal{A}(h') = \lambda^{-1}(r)$ and $h' \in \sigma(h)$. Set $h := h'$ and go to step 1.*

*We use $S_0^G$ to denote an auctioneer strategy that runs $G$.*

Given $S_0^G$, for any $S_i$, we can define an equivalent strategy $\tilde{S}_i$ for agent $i$ in the messaging game. Given $(m, R)$ and $\theta_i$, $\tilde{S}_i$ selects reply $\lambda(S_i(\lambda^{-1}(m), \theta_i))$.[13] We abuse notation and use $S_i$ to denote both a strategy for agent $i$ in $G$, and the equivalent strategy for agent $i$ in the messaging game.

## 2.4 Credible mechanisms

In formulating the idea of incentive compatibility, Hurwicz (1972) writes:

> *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 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.*

---

[13] We specify $\tilde{S}_i$ arbitrarily for communication sequences that are never observed under $S_0^G$.
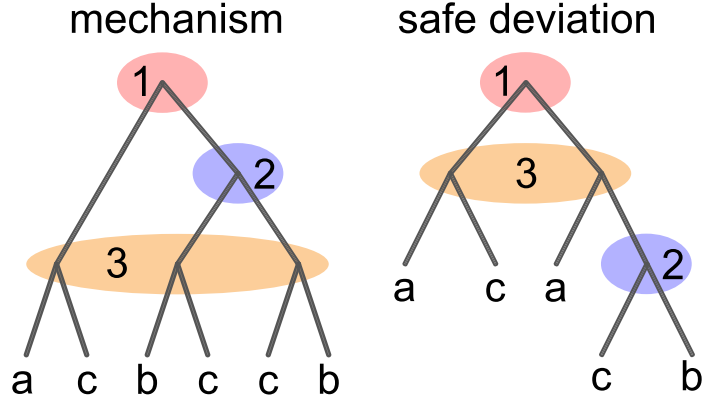
Figure 2: A mechanism and a deviation. If agent 1 cannot distinguish outcomes $a$ and $b$, then the deviation is safe.

Here we modify Hurwicz's seminal idea to include incentive compatibility for the auctioneer. First, we say that some observation $o_i$ has an "innocent explanation" (with respect to $S_0^G$) if there exist some types of other agents such that if they had those types and the auctioneer played $S_0^G$, $o_i$ would be the observation of agent $i$.

**Definition 4.** *Suppose the auctioneer promises to play $S_0^G$, the agents play $S_N$, and $i$'s type is $\theta_i$. Observation $o_i'$ has an **innocent explanation** if there exists $\theta_{-i}'$ such that $o_i(S_0^G, S_N, (\theta_i, \theta_{-i}')) = o_i'$.*

**Definition 5.** *Suppose the auctioneer promises to play $S_0^G$ and the agents play $S_N$. Then, an auctioneer strategy $S_0' \in \mathcal{S}_0$ is **safe** if for all agents $i \in N$ and all type profiles $\theta_N \in \Theta_N$, $o_i(S_0', S_N, \theta_N)$ has an innocent explanation.*

Let $\mathcal{S}_0^*(S_0^G, S_N) \equiv \{S_0' \mid S_0' \text{ is safe given promise } S_0^G \text{ and agent strategies } S_N\}$. The function $\mathcal{S}_0^*(\cdot, \cdot)$ takes a strategy for the auctioneer and a strategy profile for the agents as inputs and outputs the set of all strategies that the auctioneer can deviate to without being detected by a single agent.

Let $u_0(S_0, S_N, \theta_N)$ denote the auctioneer's utility in the messaging game, when the strategy profile is $(S_0, S_N)$ and the type profile is $\theta_N$. A protocol is credible if playing 'by the book' maximizes the auctioneer's expected utility.

**Definition 6.** $(G, S_N)$ *is **credible** if:*

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

Definition 6 permits the auctioneer to misrepresent agents' actions to each other midway through the mechanism. The following example illustrates.

**Example 1.** *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 corre-*

*sponding 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 plays 'by the book'. 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, playing by the book 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 6 takes the expectation of $\theta_N$ with respect to the *ex ante* distribution $D$, but it 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 book 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.** *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 $\theta_N$, $x^{G'}(S'_N, \theta_N) = x^G(S_N, \theta_N)$.*

**Observation 1.** *$(G, S_N)$ is credible and BIC if and only if $(S_0^G, S_N)$ is a Bayes-Nash equilibrium of the messaging game in which the auctioneer is constrained to play strategies in $\mathcal{S}_0^*(S_0^G, S_N)$.*

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.** *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 given promise $S_0$, then $(S_0', S_N)$ is a Bayes-Nash equilibrium given promise $S_0'$.*

Definition 6 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, and specify that $o_i'$ has an innocent explanation to $\theta_i$ if there exists $\theta_{-i}'$ such that $o_i'$ occurs with positive probability when the auctioneer plays $S_0^G$, the agents play $S_N$, and the type profile is $(\theta_i, \theta_{-i}')$.

Here we restrict attention to protocols in which the auctioneer does not randomize. If some randomized protocol $(G, S_N)$ is credible, then the deterministic protocol $(G', S_N)$ in which we simply fix a particular realization of the randomization is also credible. Since $(G, S_N)$ is credible, 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.

# 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.[14] Secondly, in the 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 $X = (N \cup \{0\}) \times \mathbb{R}^{|N|}$.

We assume that type spaces are discrete $\Theta_i = \{\theta_i^1, \ldots, \theta_i^{K_i}\}$. This allows us to bypass the known paradoxes of extensive game forms with continuous time or infinite actions (Simon and Stinchcombe, 1989; Myerson and Reny, 2016). Each type is associated with a real number $v(\theta_i^k)$. Assume $v(\theta_i^1) \le 0 < v(\theta_i^{K_i})$ and that $v(\theta_i^{k+1}) - v(\theta_i^k) > 0$. We will abuse notation slightly, and use $\theta_i^k$ to refer both to $i$'s $k$th 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 $F_i(\theta_i^k) = \sum_{l=1}^k f_i(\theta_i^l)$.

---

[14] 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."

Agents have private values, that is:

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

$\Omega_i$ 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[15]:

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

Let $\pi(G, S_N)$ denote the expected revenue of $(G, S_N)$.

The virtual values machinery in Myerson (1981) applies *mutatis mutandis* to the discrete setting. Suppose we choose $(G, S_N)$ to maximize expected revenue subject to incentive compatibility and voluntary participation.

**Definition 7.** $(G, S_N)$ *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 $S'_i$ that ensures that $i$ does not win and has a zero net transfer, regardless of $S'_{-i}$.[16]*

$(G, S_N)$ *is $\epsilon$-**optimal** if it satisfies the constraints and the difference between $\pi(G, S_N)$ and the optimal expected revenue is no more than $\epsilon$.*

**Definition 8.** *The **virtual value** of $\theta_i^k$ is:[17]*

$$\eta_i(\theta_i^k) \equiv \theta_i - \frac{1 - F_i(\theta_i)}{f_i(\theta_i)}(\theta_i^{k+1} - \theta_i^k) \qquad (5)$$

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

As is well-known, optimal auctions have a characterization in terms of virtual values when certain constraints bind. $\tilde{u}_i^{G, S_N}(k, k')$ denotes the expected utility of agent $i$ when

---

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

[16]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 14.

[17]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 11.

his type is $\theta_i^k$ and he plays as though his type is $\theta_i^{k'}$. $y^{G,S_N}(\theta_N)$ denotes the allocation at type profile $\theta_N$. $y_i^{G,S_N}(\theta_N)$ is an indicator variable equal to 1 if $i$ wins the object at $\theta_N$ and 0 otherwise. $t_i^{G,S_N}(\theta_N)$ is $i$'s transfer at $\theta_N$.

**Proposition 3.** *(Elkind, 2007) Assume $F_N$ is regular and $(G, S_N)$ satisfies the constraints in Definition 7. $(G, S_N)$ 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 $y^{G,S_N}(\theta_N) \in \mathrm{argmax}_i \, \eta_i(\theta_i)$.*

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

Even when the local incentive constraints are slack, the expected revenue is close to the expected virtual value of the winning bidder, provided that adjacent types are not far apart.

**Proposition 4.** *If $(G, S_N)$ is BIC, then:*

$$0 \leq \underbrace{\mathbf{E}_{\theta_N}\left[\sum_{i \in N} y_i^{G,S_N}(\theta_N)\eta_i(\theta_i)\right]}_{\text{expected virtual value}} - \pi(G, S_N) - \underbrace{\sum_{i \in N} \check{u}_i^{G,S_N}(1, 1)}_{\substack{\text{slack in lowest} \\ \text{participation constraints}}} \leq \max_i \max_{2 \leq k \leq K_i} \theta_i^k - \theta_i^{k-1}$$

(6)

**Definition 9.** *$F_N$ is **symmetric** if $\forall i, j : \forall k : K_i = K_j = K, \theta_i^k = \theta_j^k$, and $f_i(\theta_i^k) = f_j(\theta_j^k)$*

**Definition 10.** *$(G, S_N)$ is **winner-paying** if $t_i^{G,S_N}(\theta_N) \neq 0$ only if $y_i^{G,S_N}(\theta_N) = 1$.*

We start by studying the regular symmetric case. We also restrict attention to protocols which are winner-paying. These protocols respect reciprocity, in the sense that no party pays money unless he gets the object in return. We will later relax these restrictions.

For ease of exposition, we sometimes assume that the protocol breaks ties deterministically according to a fixed priority order.

**Definition 11.** *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 $\rho$ 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 $\rho$, $i$ wins the object if and only if $\theta_i \rhd \overset{\rhd}{\max}_{j \neq i} \theta_j$ and $\theta_i \rhd \rho$.*

## 3.1 Credible and static $\epsilon$-optimal auctions

We now study credible and static $\epsilon$-optimal auctions. We focus on $\epsilon$-optimality because the Revenue Equivalence Theorem breaks slightly for finite type spaces. By Proposition 3, two auctions that result in the same allocation might have different expected revenue when the local incentive constraints are slack. However, by Proposition 4, this does not matter much when the type space is fine.

**Definition 12.** $(G, S_N)$ *is* **static** *if every agent has exactly one information set and for every terminal history $z$, there exists $h \prec z$ such that $P(h) = i$.*

Consider the following mild generalization of first-price auctions: At most one agent can buy the object at a posted price. Each agent submits bids from some feasible set. If the special agent bids the posted price, then he wins for sure. Otherwise, the highest bidder wins and pays his bid.

**Definition 13.** $(G, S_N)$ *is a* **quasi-first-price auction** *if $(G, S_N)$ is static, and there exists for each agent a bid function $b_i : A(I_i) \to \mathbb{R}$ and at most one special agent $i^*$ with posted price $p^*$, such that:*

1. *Each agent pays his bid if he wins and pays nothing if he loses.*

2. *If some agent places a positive bid, then some agent wins the object.*

3. *If $i^*$ bids the posted price $p^* = \max_{a \in A(I_{i^*})} b_{i^*}(a)$, then $i^*$ wins the object. Otherwise, if $i$ wins the object then $i$'s bid is non-negative and at least as high as any other agent's bid.*

We represent a reserve price by restricting the feasible bids $b_i(A(I_i))$ for each agent. If $i$ never wins after playing action $a$, then we set $b_i(a)$ to be negative; such actions effectively decline to bid.

**Theorem 1.** *Assume $(G, S_N)$ is BIC and winner-paying. $(G, S_N)$ is credible and static if and only if $(G, S_N)$ is a quasi-first-price auction.*

*Proof.* Suppose $(G, S_N)$ is a quasi-first-price auction. $(G, S_N)$ is static by definition. If $i^*$ has bid the posted price, then the auctioneer has no discretion; every safe deviation sells the object to $i^*$ at his bid. Otherwise, no safe deviation can sell to a bidder with a negative bid, and every safe deviation that sells the object involves charging some bidder his bid, so it is optimal to sell the object to the highest bidder (breaking ties arbitrarily). Thus, $(G, S_N)$ is credible.

Suppose $(G, S_N)$ is credible and static. Recall that $(G, S_N)$ is pruned. If after some action $a$ by bidder $i$, there are two prices that $i$ might pay upon winning, the auctioneer can safely deviate to charge the higher price. Thus, for each action, if $i$ might win after

playing that action, then there is a unique price that $i$ might pay; this is $b_i(a)$. If $i$ never wins after playing some action, then we set $b_i(a) < 0$.

At most one agent can have bids that win the object for sure; that agent is $i^*$. Since $(G, S_N)$ is BIC and pruned, those actions are the set $\mathrm{argmax}_{a \in A(I_{i^*})} b_{i^*}(a)$ and we define $p^* = \max_{a \in A(I_{i^*})} b_{i^*}(a)$.

Consider any bid not placed $i^*$ or not equal to $\max_{a \in A(I_{i^*})} b_{i^*}(a)$; let $j$ be the identity of the agent who placed that bid. If that bid is positive, then the auctioneer does not keep the object; otherwise she could safely deviate to sell the object to $j$. There is an an innocent explanation if $j$ loses. Thus, if $j$ wins then $j$'s bid is non-negative, since the auctioneer can safely deviate to keep the object. Moreover, if $j$ wins then $j$'s bid is at least as high as any other bid, or the auctioneer can safely deviate to sell the object to the highest bidder for strictly more revenue. Thus, $(G, S_N)$ is a quasi-first-price auction. $\quad\square$

Clause 3 of Definition 13 is needed because posted prices are not ruled out by the requirement that the auction is credible and static. If the rules require that $i^*$ wins for sure when he bids $p^*$, then the auctioneer cannot safely deviate to sell to a higher bidder. Of course, at most one bidder can face such a price, because at most one bidder can have actions that win for sure.

When $F_N$ is regular and symmetric and the type space is fine, there exist quasi-first-price auctions that are almost optimal. Moreover, if the auction maximizes virtual value, then the posted price is no less than the equilibrium bid of any agent whose type is two steps below the highest possible type. Thus, the posted-price 'discount' vanishes as the type space gets fine. The next proposition states this formally:

**Proposition 5.** *Assume $F_N$ is regular and symmetric. There exists a quasi-first-price auction that is orderly and $\epsilon$-optimal, for $\epsilon = \max_{2 \leq k \leq K} \theta_i^k - \theta_i^{k-1}$. If a quasi-first-price auction is BIC and maximizes virtual value, then the posted price (if it exists) is at least $\max_{i \in N} b_i(S_i(I_i, \theta_i^{K-2}))$.*

*Proof.* We now construct feasible bids that result in an $\epsilon$-optimal first-price auction.

Set a reserve $\rho^* = \min_k \theta_i^k \mid \eta_i(\theta_i^k) > 0$. Consider a *second-price* auction with reserve $\rho^*$. We break ties between agents according to the order $\rhd$. It is always a best-response for type $\theta_i$ to bid $\theta_i$. Let $\tilde{b}_i(\theta_i)$ be $\theta_i$'s expected payment conditional on winning under this protocol, and set $\tilde{b}_i(\theta_i) = -1$ if $\theta_i$ never wins. Observe that if $\theta_i > \theta_j$, then $\tilde{b}_i(\theta_i) \geq \tilde{b}_j(\theta_j)$. Similarly, if $\theta_i = \theta_j$ and $i \rhd j$, then $\tilde{b}_i(\theta_i) \geq \tilde{b}_j(\theta_j)$.

Now consider a quasi-first-price auction in which $b_i(S_i(I_i, \theta_i)) = \tilde{b}_i(\theta_i)$ and we allocate the object in the same way. This new protocol is BIC and the participation constraints of the lowest types bind.[18] Since $F_N$ is regular and symmetric, the result-

---

[18] Recall our assumption that $\theta_i^1 \leq 0$, which implies that $\eta_i(\theta_i^1) < 0$. These agents never receive the object and always have zero transfers.

ing allocation maximizes virtual value. Thus, by Proposition 4, $(G, S_N)$ is $\epsilon$-optimal for $\epsilon = \max_{2 \leq k \leq K_i} \theta_i^k - \theta_i^{k-1}$.

Take any quasi-first-price price auction $(G, S_N)$ that is BIC and maximizes virtual value. By construction, $\theta_{i*}^K$ always wins and, since $(G, S_N)$ maximizes virtual value, $\theta_{i*}^{K-1}$ sometimes loses. Thus, the posted price is strictly greater than the bid placed by $\theta_{i*}^{K-1}$, by BIC. That bid, in turn, must be at least $\max_{i \in N} b_i(S_i(I_i, \theta_i^{K-2}))$, since $\theta_{i*}^{K-1}$ only wins when there is no strictly higher bid. $\qquad \square$

## 3.2 Credible and strategy-proof optimal auctions

We now characterize credible and strategy-proof optimal auctions.

**Definition 14.** $(G, S_N)$ is **strategy-proof** if, for all $i \in N$, for all $S'_{-i}$:

$$S_i \in \underset{S'_i}{\arg\max} \, \mathbb{E}_{\theta_N}[u_i^G(S'_i, S'_{-i}, \theta_N)] \tag{7}$$

**Definition 15.** $(G, S_N)$ is an **ascending auction** (with reserve price $\rho$) 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 $\rho$ (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 $\Theta_i$ and actions that **quit**.

   (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.[19]

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 $\rho$ and there is no high bidder, he places a bid.

---

[19]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.

(c) *If $i$'s type is above the current high bid (breaking ties with $\rhd$), he places a bid.*[20]

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 $\theta_i^K$, and no active bidder has higher tie-breaking priority. In that case, the object is sold to the high bidder at his last bid.*

We pause to note a mild indeterminacy: When an active bidder is called to play, it could be that the available bid is not yet enough to become the high bidder. For instance, bidder $i$ might choose whether to place a bid of 50 or quit, even though the current high bid is 100. In that case, types of $i$ between 50 and 100 could place the bid or could quit. However, $S_i$ is measurable with respect to $i$'s information sets, so it must be that bidder $i$ never quits when he might still win.

**Observation 3.** *If $F_N$ is 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, such an auction is optimal.*

The definition of extensive-form mechanisms permits the auctioneer to communicate with agents in any order, to convey any information (or no 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 2.** *Assume $F_N$ is 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. A key feature of ascending auctions is that, at each history, the types of $i$ that might win pool on the same action, unless every other agent has quit. This is stated precisely in Proposition 19, and is closely related to unconditional winner privacy as defined by Milgrom and Segal (2017). If at some history winning 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

---

[20]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.

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.) □

By Theorem 1, restricting attention to revelation mechanisms forces a sharp choice between incentives for the auctioneer and strategy-proofness for the agents. Theorem 2 shows that allowing other extensive forms relaxes this trade-off.

Unlike Theorem 1, the characterization in Theorem 2 assumes optimality. This is not just a feature of our proof technique: the ascending auction is credible because it is optimal. If the type distributions are asymmetric, then the auctioneer may profitably deviate by enforcing bidder-specific reserve prices.[21] We characterize the asymmetric case in Theorem 4.

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:

**Proposition 6** (Pay-as-bid). *Assume $(G, S_N)$ is credible. Suppose $i$ is called to play at information set $I_i$, takes some action $a$, and might win without being called to play again. Then there is a unique price $t_i(I_i, a)$ that $i$ will pay if he wins without being called to play again.*

*Proof.* Suppose $i$ might win without being called to play again, and there are two distinct prices $t_i < t_i'$ that $i$ might pay in that circumstance. 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. □

---

[21]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.

Proposition 6 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 Proposition 6 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.

## 3.3   Corollaries

Clearly, the quasi-first-price auction is not strategy-proof, except in the degenerate case that each agent has exactly one feasible bid. Thus, we have the following corollary.

**Corollary 1** (Auction Trilemma)**.** *Assume $F_N$ is regular and symmetric and $(G, S_N)$ is BIC and winner-paying. If there exist $\theta_N$ and $\theta'_N$ such that $t_i^{G,S_N}(\theta_N) > t_i^{G,S_N}(\theta'_N) > 0$, then $(G, S_N)$ is not static, strategy-proof, and credible. However, there exist $\epsilon$-optimal $(G, S_N)$, for $\epsilon = \max_i \max_{2 \le k \le K} \theta_i^k - \theta_i^{k-1}$, that are:*

1. *static and strategy-proof (the second-price auction),*

2. *static and credible (the first-price auction),*

3. *strategy-proof and credible (the ascending auction).*

Consider the messaging game restricted to safe deviations $\mathcal{S}_0^*(S_0^G, S_N)$. Under a quasi-first-price auction, the auctioneer's strategy $S_0^G$ is a best-response to any agent strategy profile $S'_N$. Under an ascending auction, each agent's strategy is a best-response to any opponent strategies $S'_{N \setminus i}$ and any safe auctioneer strategy $S'_0$. Theorem 2 implies that no protocol can provide dominant-strategy incentives to both sides at once.[22]

**Corollary 2** (Strategy-proofness for one side only)**.** *Assume $F_N$ is regular and symmetric and $(G, S_N)$ is orderly and optimal. Assume that for any optimal reserve $\rho^* < \theta_i^{K-2}$. In the messaging game restricted to $\mathcal{S}_0^*(S_0^G, S_N)$, either there exists $S'_N$ such that $S_0^G$ is not a best-response to $S'_N$, or there exists $i \in N$ and $S'_{N \setminus i}$ such that $S_i$ is not a best-response to $(S_0^G, S'_{N \setminus i})$.*

*Proof.* If $(G, S_N)$ is not credible, then $S_0^G$ is not a best-response to $S_N$. If $(G, S_N)$ is not strategy-proof, then there exists $i \in N$ and $S'_{N \setminus i}$ such that $S_i$ is not a best-response to $(S_0^G, S'_{N \setminus i})$. If $(G, S_N)$ is credible and strategy-proof, then $(G, S_N)$ is an ascending auction by Theorem 2. In that case, consider $S'_N$ such that $i$ bids until the price hits $\theta_i^{K_i}$, and every other bidder quits before the reserve is met. $S_0^G$ is not a best-response to $S'_N$.   □

---

## 3.4  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 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 $\theta_N$, 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

## 4.1  Transfers from losing bidders

In stating Theorem 1, we restricted attention to winner-paying protocols. We now relax that assumption: the quasi-first-price auctions of Theorem 1 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. It could be that real-world auctions must 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 a definition that generalizes quasi-first-price auctions.

**Definition 16.** $(G, S_N)$ *is a **twin-bid auction** if $(G, S_N)$ is static, and there exist for each agent two bid functions, $b_i^W : A(I_i) \to \mathbb{R}$ and $b_i^L : A(I_i) \to \mathbb{R}$, such that:*

1. *After playing $a_i$, $i$ pays $b_i^W(a_i)$ if he wins and $b_i^L(a_i)$ if he loses.*

2. *If some agent plays an action $a_i$ with $b_i^W(a_i) - b_i^L(a_i) > 0$, then some agent wins the object.*

3. *If $i$ wins the object after action profile $(a_j)_{j \in N}$, then $b_i^W(a_i) - b_i^L(a_i) \geq \max\{0, \max_{j \neq i} b_j^W(a_j) - b_j^L(a_j)\}$.*

**Theorem 3.** $(G, S_N)$ *is credible and static if and only if $(G, S_N)$ is a twin-bid auction.*

*Proof.* Suppose $(G, S_N)$ is a twin-bid auction. After action profile $(a_i)_{i \in N}$, every safe deviation charges $b_i^W(a_i)$ if agent $i$ wins and $b_i^L(a_i)$ 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. The same argument as in the proof of Theorem 1 pins down the uniqueness of $b_i^W(a_i)$ if $i$ might win after playing $a_i$, and the uniqueness of $b_i^L(a_i)$ if $i$ might lose. We set $b_i^W(a_i)$ to be strictly less than $b_i^L(a_i)$ if $i$ never wins after $a_i$. At most one agent has actions that guarantee he wins the object; for such an action we set $b_i^L(a_i)$ to be arbitrarily low, so that $b_i^W(a_i) - b_i^L(a_i) > \max\{0, \max_{j \neq i, a_j} b_j^W(a_j) - b_j^L(a_j)\}$.

Consider some agent $i$ who plays $a_i$, where $a_i$ does not guarantee that $i$ wins. The auctioneer does not keep the object when $b_i^W(a_i) - b_i^L(a_i) > 0$, since she could safely deviate to allocate the object to $i$. There is an innocent explanation if $i$ loses; thus if $i$ wins then $b_i^W(a_i) - b_i^L(a_i) \geq 0$, since the auctioneer could safely deviate to keep the object. Finally, if $i$ wins then $b_i^W(a_i) - b_i^L(a_i) \geq \max_{j \neq i} b_j^W(a_j) - b_j^L(a_j)$, otherwise she could safely deviate to sell the object to $\text{argmax}_{j \neq i} b_j^W(a_j) - b_j^L(a_j)$. Thus, $(G, S_N)$ is a twin-bid auction. $\qquad\square$

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, $b_i^W(a_i) - b_i^L(a_i)$ must be no less than the auctioneer's cost of provision, which rules out standard all-pay auctions.[23]) Twin-bid auctions also encompass first-price auctions with entry fees ($\forall a_i, a_i' : b_i^L(a_i) = b_i^L(a_i') > 0$), and first-price auctions in which losing bidders are paid fixed compensation ($\forall a_i, a_i' : b_i^L(a_i) = b_i^L(a_i') < 0$). 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 bidders with constant absolute risk aversion.[24]

Twin-bid auctions are not strategy-proof, except in degenerate cases.

**Proposition 7.** *Assume there exist* $\theta_i < \theta_i' < \theta_i'' < \theta_i'''$, $\theta_{-i}$, *and* $\theta_{-i}'$ *such that* $y^{G, S_N}(\theta_i, \theta_{-i}) \neq i = y^{G, S_N}(\theta_i', \theta_{-i})$ *and* $y^{G, S_N}(\theta_i'', \theta_{-i}') \neq i = y^{G, S_N}(\theta_i''', \theta_{-i}')$. *If* $(G, S_N)$ *is a twin-bid auction, then* $(G, S_N)$ *is not strategy-proof.*

## 4.2  Asymmetric distributions

Theorem 2 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 $\theta_j$ 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.

---

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

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

**Definition 17.** $F_N$ 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 18.** $(G, S_N)$ 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.

   (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 4.** *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 $\Theta_i$ and $\Theta_j$ 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 18 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 $\theta_i^k$ to $\theta_i^{k'}$, the auctioneer promises ahead of time to jump $i$'s price directly from $\theta_i^k$ 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 $\theta_i$ makes a bid equal to the expected payment of $\theta_i$ 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.[25]

## 4.3 Affiliated values

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 are not credible. Even using extensive forms does not generally permit credible full surplus extraction.

**Definition 19.** $(G, S_N)$ **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 8.** *The Cremer and McLean (1988) conditions are not sufficient for the existence of a credible protocol that extracts full surplus.*

Optimal auctions with correlation are better-behaved if we additionally require *ex post*

---

incentive compatibility and *ex post* individual rationality.[26] 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 $\theta_i^k = \theta_j^k$ for all $i, j, k$, as well as affiliated types (Milgrom and Weber, 1982).

**Definition 20.** $f_N$ is **symmetric** *if its value is equal under any permutation of its arguments.* $f_N$ is **affiliated** *if for all* $\theta_N, \theta'_N$:

$$f_N(\theta_N \vee \theta'_N) f_N(\theta_N \wedge \theta'_N) \geq f_N(\theta_N) f_N(\theta'_N) \tag{8}$$

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

**Definition 21.** $(G, S_N)$ *is* **optimal among ex post auctions** *if it maximizes* $\pi(G, S_N)$ *subject to the constraints:*

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

$$\theta_i y_i^{G,S_N}(\theta_i, \theta_{-i}) - t_i^{G,S_N}(\theta_i, \theta_{-i}) \geq \theta_i y_i^{G,S_N}(\theta'_i, \theta_{-i}) - t_i^{G,S_N}(\theta'_i, \theta_{-i}) \tag{9}$$

*2. Ex post individual rationality. For all* $i, \theta_i, \theta_{-i}$:

$$\theta_i y_i^{G,S_N}(\theta_i, \theta_{-i}) - t_i^{G,S_N}(\theta_i, \theta_{-i}) \geq 0 \tag{10}$$

**Definition 22.** *The* **conditional virtual value** *of* $\theta_i^k$ *given* $\theta_{-i}$ *is:*

$$\eta_i(\theta_i^k | \theta_{-i}) \equiv \theta_i^k - \frac{1 - F_i(\theta_i^k | \theta_{-i})}{f_i(\theta_i^k | \theta_{-i})} (\theta_i^{k+1} - \theta_i^k) \tag{11}$$

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

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.[27] That reserve depends on the final bids from the bidders who quit.

---

[26] *Ex post* incentive compatibility and *ex post* individual rationality are implied by strategy-proofness and voluntary participation (Definition 7). 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 $\theta_{-i}$.

[27] 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.

**Definition 23.** $(G, S_N)$ *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[28] 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 9.** *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.*

# 5 A 'Prior-free' Definition

The definition of credibility depends on the joint distribution of agent types (Definition 6). It may be useful to have a definition that is 'prior-free', for settings such as matching or maxmin mechanism design.

**Definition 24.** *Given $(G, S_N)$, $S_0 \in \mathcal{S}_0^*(S_0^G, S_N)$ is **always-profitable** if, for all $\theta_N$:*

$$u_0(S_0, S_N, \theta_N) \geq u_0(S_0^G, S_N, \theta_N) \tag{12}$$

*with strict inequality for some $\theta_N$.*

*$(G, S_N)$ 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.

What happens if we replace "credible" with "prior-free credible" in the statement of Theorems 1 and 2? Quasi-first-price auctions and ascending auctions are credible, so they are prior-free credible. However, the other direction of implication now starts from weaker premises. Prior-free credibility still suffices for the characterizations.

**Proposition 10.** *Theorems 1, 2, 3, and 4 remain true when "credible" is replaced by "prior-free credible".*

---

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

*Proof.* Since credible protocols are prior-free credible, each "if" direction is immediate. In the proof of Theorem 1, we show that if $(G, S_N)$ is static but not a quasi-first-price auction, then there exists a safe deviation that is always-profitable, so $(G, S_N)$ is not prior-free credible. In the proof of Theorem 2, 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 Theorems 3 and 4. $\qquad\square$

# 6 Other Applications

## 6.1 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 25.** $\mathcal{F}$ *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 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 18, 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.

## 6.2 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 $\Theta_i = \{0, 1, 2, \dots, K\}$. The efficient allocation is:

$$y^*(\theta_N) = \begin{cases} 1 & \text{if } \sum_i \theta_i - c \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

The planner wants to choose the efficient allocation, but also receives a small benefit from having higher transfers. Formally, for $\gamma \in (0, \frac{1}{|N|K})$:

$$u_0(y, t_N, \theta_N) = 1_{y=y^*(\theta_N)} + \gamma \sum_i t_i \tag{14}$$

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 key intuition is that, for a static protocol to be prior-free credible, $i$'s transfers must be measurable with respect to the allocation rule, which prevents the

use of threshold prices.

**Proposition 12.** *Assume there exist $\theta_i < \theta_i' < \theta_i'' < \theta_i'''$, $\theta_{N\setminus i}$, and $\theta_{N\setminus i}'$ 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.*

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.

**Proposition 13.** *The above protocol for two agents is efficient, strategy-proof, and prior-free credible.*

*Proof.* Efficiency and strategy-proofness follow by inspection. 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 $\gamma$ is small, the protocol is prior-free credible. $\square$

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.

# 7 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.[29] 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.[30]

When an auctioneer makes repeated sales, reputation could help enforce the full-commitment outcome. However, the force of reputation depends on the discount rate 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.[31] 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).[32] 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.

---

[29]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.

[30]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."

[31]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.

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

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.[33] 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.

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.

---

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

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*.

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.

Grant, D. (2014). Why auction rooms seem empty these days. *The Wall Street Journal*.

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.

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.

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. (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.

Myerson, R. B. and Reny, P. J. (2016). Broad sequential equilibria of multi-stage games with infinite sets of signals and actions. *Working paper*.

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[34].

    (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) $\sigma(h)$ 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 $\sigma(h)$.

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

$$A(h) \equiv \bigcup_{h' \in \sigma(h)} \mathcal{A}(h') \tag{15}$$

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 $h \in I_i$.

---

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

(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. $g : Z \to X$

# B  Proofs omitted from the main text

## B.1  Proposition 1

For each of the three clauses in Definition 2, we show that if $(G, S_N)$ does not satisfy this clause, we can transform $(G, S_N)$ to have strictly fewer histories, such that the transformed protocol is BIC and results in the same outcomes for each type profile. Since the set of histories in $(G, S_N)$ is finite, it follows that there exists a BIC $(G', S'_N)$ that cannot be reduced further, but results in the same outcomes as $(G, S_N)$.

**Clause 1:** Suppose there exists $h$ such that there is no $\theta_N$ such that $h \preceq z(S_N, \theta_N)$. Since the game tree is finite, 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 $i$ preferred his original $S_i$ to any strategy that played $a'$ at $I'_i$, his new strategy $S'_i$ remains incentive-compatible. Thus, the transformed $(G', S'_N)$ is BIC.

**Clause 2:** Suppose there exists $h \notin Z$ such that $|\sigma(h)| = 1$. 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 $H \setminus I_i$. $(G', S'_N)$ is BIC.

**Clause 3:** The above arguments prove that, starting from an arbitrary $(G, S_N)$, we can produce an outcome-equivalent $(G, S_N)$ that satisfies Clauses 1 and 2. We now take $(G, S_N)$ that satisfies Clauses 1 and 2, and show that if it does not satisfy Clause 3, then we can reduce the protocol further.

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 $i$

played the action that would be selected by some type $\theta_i$. This modified protocol contains strictly fewer histories.

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 $(\theta_i, \theta_{-i})$ such that $h \prec z(S_N, (\theta_i, \theta_{-i}))$. Upon reaching $h$, we can henceforth 'automate' play as though $i$ had type $\theta_i$. 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 $\theta_{-i}''$ such that $h' \prec z(S_N, (\theta_i, \theta_{-i}''))$; this has the effect of 'automating' play as though $i$ has type $\theta_i$. 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 $H'$.

By construction, for all $\theta_i'$, if $i$ is playing as though his type is $\theta_i'$ and we would have reached 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 $\theta_i$ under $(G, S_N)$ (which by hypothesis is the same as when $i$ is playing as though his type is $\theta_i'$ under $(G, S_N)$). Plainly, if we would not have reached history $h$ under $(G, S_N)$, then the outcomes under $(G, S_N)$ and $(G', S_N')$ are identical. Thus, $(G', S_N')$ is BIC.

This completes the proof of Proposition 1.

## B.2    Proposition 2

To prove Proposition 2, we show that each of the three transformations we used in the proof of Proposition 1 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 follow action $a$ at $I_i$, when $a$ is never chosen on the path of play). Suppose the auctioneer had a profitable safe deviation $S_0'$ from $S_0^{G'}$. The auctioneer could make that same deviation, but additionally offer the response $\lambda(a)$ whenever sending the message $\lambda(I_i)$. By hypothesis, agent $i$ never selects $\lambda(a)$ as a reply, so for any $\theta_N$ and any $j$, $j$'s resulting observation has an innocent explanation. Thus, the auctioneer also has a profitable safe deviation from $S_0^G$.

Consider the second transformation (deleting all histories in some information set with a singleton action set). 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 the

deleted information set $I_i$, the auctioneer delays sending $\lambda(I_i, A(I_i))$ until the last possible moment. That is consider $S_0$ that is the same as $S_0'$, except that, if the auctioneer has not yet sent $\lambda(I_i, A(I_i))$ to agent $i$, then:

1. If $S_0'$ specifies that the auctioneer sends $\lambda(I_i', A(I_i'))$ for $I_i' \succ I_i$, then $S_0$ specifies that she first sends $\lambda(I_i, A(I_i))$ and then (immediately thereafter) sends $\lambda(I_i', A(I_i'))$

2. If $S_0'$ specifies that the auctioneer chooses an outcome such that the resulting observation for $i$ does not have an innocent explanation under $S_0^G$, then $S_0$ specifies that she first sends $\lambda(I_i, A(I_i))$ before choosing that outcome.

$S_0$ is a profitable safe deviation from $S_0^G$.

Consider the third transformation (deleting histories where $i$ is called to play, following some 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 $S_0^{G'}$.

If the observation for $i$ that results from $S_0'$ does not have an innocent explanation under $S_0^G$, it must be that (given on all the communication $i$ has seen so far), the outcome $S_0'$ is about to select can only occur under $G$ at terminal histories that follow $h$. But by hypothesis, for any $\theta_i$ and $\theta_i'$ that are consistent with reaching $h$, and any $\theta_{-i}$ consistent with reaching $h$, the resulting outcome is the same. Thus, let $S_0$ be exactly as in $S_0'$, except that if $S_0'$ specifies that the auctioneer chooses an outcome such that the resulting observation for $i$ does not have an innocent explanation under $S_0^G$, then $S_0$ specifies that the auctioneer communicates with $i$ as though play started from $h$ and the opponent type profiles were $\theta_{-i}$, for some $\theta_{-i}$ consistent with reaching $h$.

Formally, if $S_0'$ would choose an outcome such that $i$'s observation has no innocent explanation, then fix some $\theta_N$ such that $h \prec z(S_N, \theta_N)$. Initialize $\hat{h} := h$.

1. If $\hat{h} \in Z$, then terminate and choose $x = g(\hat{h})$.

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

   (a) $\hat{h} := h' \mid h' \in \sigma(\hat{h})$ and $S_{P(\hat{h})}(I_{P(\hat{h})}, \theta_{P(\hat{h})}) = \mathcal{A}(h')$.
   (b) Go to step 1.

3. Else:

   (a) Send $(m, R) = \lambda(I_i, A(I_i))$ for $I_i$ such that $\hat{h} \in I_i$.
   (b) Upon receiving $r \in R$, choose $\hat{h} := h' \mid \mathcal{A}(h') = \lambda^{-1}(r)$ and $h' \in \sigma(\hat{h})$.
   (c) Go to step 1.

Since, under $S_i$, $i$'s play in this final stage makes no difference to the outcome, delaying communication with $i$ until the outcome is about to be selected results in a safe deviation. This completes the proof of Proposition 2..

41

## B.3 Proposition 4

To derive the lower bound, we use the downward incentive constraints. Since $(G, S_N)$ is BIC, type $\theta_i^k$ should not wish to imitate type $\theta_i^{k-1}$, *i.e.*:

$$\forall i : \forall k \geq 2 : \check{u}_i^{G,S_N}(k, k) \geq \check{u}_i^{G,S_N}(k, k-1) \tag{16}$$

Thus,

$$\check{u}_i^{G,S_N}(k, k) - \check{u}_i^{G,S_N}(1, 1) = \sum_{l=2}^{k} \check{u}_i^{G,S_N}(l, l) - \check{u}_i^{G,S_N}(l-1, l-1)$$
$$\geq \sum_{l=2}^{k} \check{u}_i^{G,S_N}(l, l-1) - \check{u}_i^{G,S_N}(l-1, l-1) = \mathbf{E}_{\theta_{-i}}[\sum_{l=2}^{k}(\theta_i^l - \theta_i^{l-1})y_i^{G,S_N}(\theta_i^{l-1}, \theta_{-i})] \tag{17}$$

Thus, $i$'s expected utility is at least:

$$\mathbf{E}_{\theta_{-i}}\left[\sum_{k=2}^{K_i} f_i(\theta_i^k)\sum_{l=1}^{k}(\theta_i^l - \theta_i^{l-1})y_i^{G,S_N}(\theta_i^{l-1}, \theta_{-i})\right] + \check{u}_i^{G,S_N}(1, 1)$$
$$= \mathbf{E}_{\theta_{-i}}\left[\sum_{k=1}^{K_i}(1 - F_i(\theta^k))(\theta_i^{k+1} - \theta_i^k)y_i^{G,S_N}(\theta_i^k, \theta_{-i})\right] + \check{u}_i^{G,S_N}(1, 1)$$
$$= \mathbf{E}_{\theta_{-i}}\left[\sum_{k=1}^{K_i}(\theta_i^{k+1} - \theta_i^k)f_i(\theta^k)\frac{1 - F_i(\theta^k)}{f_i(\theta^k)}y_i^{G,S_N}(\theta_i^k, \theta_{-i})\right] + \check{u}_i^{G,S_N}(1, 1) \tag{18}$$

Summing across agents yields the lower bound:

$$0 \leq \mathbf{E}_{\theta_N}\left[\sum_{i \in N} y_i^{G,S_N}(\theta_N)\eta_i(\theta_i)\right] - \pi(G, S_N) - \sum_{i \in N}\check{u}_i^{G,S_N}(1, 1) \tag{19}$$

To derive the upper bound, we use the upward incentive constraints. Since $(G, S_N)$ is BIC, $\theta_i^{k-1}$ should not wish to imitate type $\theta_i^k$, *i.e.*:

$$\forall i : \forall k \geq 2 : \check{u}_i^{G,S_N}(k-1, k-1) \geq \check{u}_i^{G,S_N}(k-1, k) \tag{20}$$

$$\check{u}_i^{G,S_N}(k, k) - \check{u}_i^{G,S_N}(1, 1) = \sum_{l=2}^{k} \check{u}_i^{G,S_N}(l, l) - \check{u}_i^{G,S_N}(l-1, l-1)$$
$$\leq \sum_{l=2}^{k} \check{u}_i^{G,S_N}(l, l) - \check{u}_i^{G,S_N}(l-1, l) = \sum_{l=2}^{k}(\theta_i^l - \theta_i^{l-1})\mathbf{E}_{\theta_{-i}}[y_i^{G,S_N}(\theta_i^l, \theta_{-i})] \tag{21}$$

Let $\epsilon = \max_j \max_{2 \le k \le K_j} \theta_j^k - \theta_j^{k-1}$.

$$\left( \sum_{k=1}^{K_i} f_i(\theta_i^k) \check{u}_i^{G,S_N}(k,k) \right) - \check{u}_i^{G,S_N}(1,1)$$

$$\le \mathbf{E}_{\theta_{-i}} \left[ \sum_{k=1}^{K_i} f_i(\theta_i^k) \sum_{l=2}^{k} (\theta_i^l - \theta_i^{l-1}) y_i^{G,S_N}(\theta_i^l, \theta_{-i}) \right]$$

$$\le \mathbf{E}_{\theta_{-i}} \left[ \sum_{k=1}^{K_i} (1 - F_i(\theta^k))(\theta_i^{k+1} - \theta_i^k) y_i^{G,S_N}(\theta_i^k, \theta_{-i}) \right]$$

$$+ \epsilon \mathbf{E}_{\theta_{-i}} \left[ \sum_{k=2}^{K_i} f_i(\theta_i^k)(y_i^{G,S_N}(\theta_i^k, \theta_{-i}) - y_i^{G,S_N}(\theta_i^1, \theta_{-i})) \right]$$

$$\le \mathbf{E}_{\theta_{-i}} \left[ \sum_{k=1}^{K_i} (\theta_i^{k+1} - \theta_i^k) f_i(\theta_i^k) \frac{1 - F_i(\theta_i^k)}{f_i(\theta_i^k)} y_i^{G,S_N}(\theta_i^k, \theta_{-i}) \right] + \epsilon \quad (22)$$

Summing across agents yields the upper bound:

$$\mathbf{E}_{\theta_N} \left[ \sum_{i \in N} y_i^{G,S_N}(\theta_N) \eta_i(\theta_i) \right] - \pi(G, S_N) - \sum_{i \in N} \check{u}_i^{G,S_N}(1,1) \le \epsilon \quad (23)$$

## B.4  Theorem 2

### B.4.1  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 $y^{G,S_N}$ and $t_i^{G,S_N}$ to reduce clutter.

**Proposition 14.** *If $(G, S_N)$ is optimal and strategy-proof, then $(G, S_N)$ is winner-paying.*

*Proof.* For all $(\theta_i, \theta_{-i})$, if $y(\theta_i, \theta_{-i}) \ne i$ then $t_i(\theta_i, \theta_{-i}) \le 0$. Suppose not. $(G, S_N)$ satisfies voluntary participation. When $i$'s opponent's imitate $\theta_{-i}$,[35] type $\theta_i$ can profitably deviate to non-participation if $t_i(\theta_i, \theta_{-i}) > 0$, contradicting strategy-proofness.

$\theta_i^1 \le 0$, so $\eta_i(\theta_i^1) < 0$. $(G, S_N)$ is optimal, so $\theta_i^1$ never wins (by Proposition 3). $\theta_i^1$'s participation constraint binds, so for all $\theta_{-i}$, $t_i(\theta_i^1, \theta_{-i}) = 0$.

Take any $(\theta_i, \theta_{-i})$. If $y(\theta_i, \theta_{-i}) \ne i$ and $t_i(\theta_i, \theta_{-i}) > 0$, then when $i$'s opponents imitate $\theta_{-i}$, $\theta_i^1$ can profitably imitate $\theta_i$, contradicting strategy-proofness. Thus, $(G, S_N)$ is winner-paying. $\qquad \square$

**Proposition 15.** *If $(G, S_N)$ is strategy-proof, then the allocation rule is monotone. That is, if $\theta_i < \theta_i'$ and $y(\theta_i, \theta_{-i}) = i$, then $y(\theta_i', \theta_{-i}) = i$.*

---

[35]Formally, define $S'_{-i}$ such that for all $j \ne i$, $I_j$, and $\theta_j'$, $S_j'(I_j, \theta_j') = S_j(I_j, \theta_j)$

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

**Definition 26.** $(G, S_N)$ *has* **threshold pricing** *if:*

$$t_i(\theta_N) = \begin{cases} \min_{\theta_i' \in \Theta_i} \theta_i' \mid y(\theta_i', \theta_{-i}) = i & \text{if } y(\theta_N) = i \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

**Proposition 16.** *If $(G, S_N)$ is optimal and strategy-proof, then $(G, S_N)$ has threshold pricing.*

*Proof.* Proposition 14 pins down the payments whenever $y(\theta_N) \neq i$.

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

Take any $\theta_{-i}$. There are three cases to consider.

If $y(\theta_i^k, \theta_{-i}) = i$, then strategy-proofness implies that $y(\theta_i^{k+1}, \theta_{-i}) = i$ and $t_i(\theta_i^{k+1}, \theta_{-i}) = t_i(\theta_i^k, \theta_{-i}) = \min_{\theta_i' \in \Theta_i} \theta_i' \mid y(\theta_i', \theta_{-i}) = i$.

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

Notice that, in the previous two cases, $\theta_i^{k+1}$ is exactly indifferent between $S_i$ and deviating to imitate type $\theta_i^k$. Finally, suppose $y(\theta_i^k, \theta_{-i}) \neq i$ and $y(\theta_i^{k+1}, \theta_{-i}) = i$. $t_i(\theta_i^{k+1}, \theta_{-i}) \leq \theta_i^{k+1}$, since $(G, S_N)$ is strategy-proof. If $t_i(\theta_i^{k+1}, \theta_{-i}) < \theta_i^{k+1}$, then $(G, S_N)$ is not optimal, since the incentive constraints do not bind locally downward (Proposition 3). Thus, $t_i(\theta_i^{k+1}, \theta_{-i}) = \theta_i^{k+1}$, and the inductive step is proved. $\square$

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

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

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

**Proposition 17.** *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 $\Theta_i^h$ 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 $\theta_j$ wins, then $i$ cannot (at this information set) rule out the possibility that $j$'s type is instead some $\theta_j' > \theta_j$.

**Proposition 18.** *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 $y(\theta_i, \theta_j, \theta_{N \setminus \{i,j\}}) = j$ and $y(\theta_i', \theta_j, \theta_{N \setminus \{i,j\}}) \neq j$, 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 $\theta_j'$ has a profitable deviation. For $l \in N \setminus \{i,j\}$, let $l$ imitate $\theta_l$; that is $\forall I_l : \forall \theta_l' : S_l'(I_l, \theta_l') = S_l(I_l, \theta_l)$. Let $i$ imitate $\theta_i'$ unless he encounters $I_i$, and let him imitate type $\theta_i$ 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 16, $(G, S_N)$ has threshold pricing. If type $\theta_j'$ deviates to imitate $\theta_j$, then (when facing $S_j'$), the path of play passes through $I_i$, so $j$ wins at price $\min_{\theta_j'' \in \Theta_j} \theta_j'' \mid y(\theta_i, \theta_j'', \theta_{N \setminus \{i,j\}}) = j$, for a positive surplus since $\theta_j' > \theta_j$. On the other hand, if type $\theta_j'$ 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 $\min_{\theta_j'' \in \Theta_j} \theta_j'' \mid y(\theta_i', \theta_j'', \theta_{N \setminus \{i,j\}}) = j$, or does not win and has zero surplus. Thus, $j$ has a profitable deviation, and $(G, S_N)$ is not strategy-proof, a contradiction. $\square$

Let $W_i^h$ denote the subset of $i$'s types that might reach $h$ and then win. Similarly, let $L_i^h$ denote the subset of $i$'s types that might reach $h$ and then lose.

$$W_i^h = \{\theta_i \in \Theta_i^h \mid \exists \theta_{-i} \in \Theta_{-i}^h : y(\theta_i, \theta_{-i}) = i\} \tag{29}$$

$$L_i^h = \{\theta_i \in \Theta_i^h \mid \exists \theta_{-i} \in \Theta_{-i}^h : y(\theta_i, \theta_{-i}) \neq i\} \tag{30}$$

**Definition 27.** $(G, S_N)$ is **winner-pooling** if for all $I_i$, $h \in I_i$:

1. Either: $\forall \theta_i, \theta_i' \in W_i^h : S_i(I_i, \theta_i) = S_i(I_i, \theta_i')$

2. Or: $W_i^h \cap L_i^h = \emptyset$

**Proposition 19.** *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 19, we highlight that this is the reason that we have assumed regularity and orderliness in the statement of Theorem 2. Together,

regularity and orderliness imply that, if there are two distinct types $\theta_i < \theta_i'$ in $W_i^h$ that do not pool on the same action, then there exists $\theta_{-i}$ such that $\theta_i$ loses when facing $\theta_{-i}$, but $\theta_i'$ wins. This enables us to construct profitable safe deviations for the auctioneer.[36]

*Proof.* Under the assumptions of Proposition 19, 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 orderly and the winner-pooling property held at all predecessors to $h^*$, it follows that for all $i$, either $W_i^{h^*} = \emptyset$ or $W_i^{h^*} = \{\theta_i \mid \theta_i \rhd \overset{\rhd}{\max_{j \neq i}} \underline{\theta}_{-j}^{h^*}$ and $\theta_i \rhd \rho\}$.

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, $\overset{\rhd}{\min} W_{i^*}^{h^*} \in W_{i^*}^{h^*} \cap L_{i^*}^{h^*}$. Define

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

We are going to squeeze extra revenue out of agent $i^*$ when his type is $\theta_{i^*}^*$: 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 $\theta_{i^*}^*$, that is

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

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:

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

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

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

where $\unrhd$ is the reflexive order implied by the strict order $\rhd$.

Given $S_0^G$ (with corresponding $\lambda$), we now exhibit a (partial) behavioral strategy that

---

[36]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.

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 $\theta_{-i}$, updating $\hat{h}$ as we go along. When we do this, we say that we **simulate** $\theta_{-i}$ 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 $\hat{h} := h \in \sigma(\hat{h}) \mid \theta_{-i} \in \Theta_{-i}^h$

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

    (a) Send $(\lambda(I_i), \lambda(A(I_i)))$ for $I_i \mid \hat{h} \in I_i$ to $i$.

    (b) Upon receiving $r$, set $\hat{h} := h \mid h \in \sigma(\hat{h})$ and $\mathcal{A}(h) = \lambda^{-1}(r)$.

    (c) Go to step 1.

The algorithm also calls the following subroutine: Given some history $h$ and some $\theta_{-i}$, where $i$ was called to play at $h$'s immediate predecessor, we may **find the cousin** of $h$ consistent with $\theta_{-i}$. 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 $\theta_{-i}$. Formally, let cousin$(h, \theta_{-i})$ be equal to $h'$ such that $\exists I_i : \exists h'', h''' :$

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

2. $h \in \sigma(h'')$

3. $h' \in \sigma(h''')$

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 cousin$(h, \theta_{-i})$ is well-defined when we invoke 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 $\beta$ to be the least type consistent with $i^*$'s responses, and go to Stage 2. Otherwise, we lower $\beta$ appropriately, and proceed to Stage 2. In Stage 2, we cycle through the bidders, updating $\beta$ 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 $\beta$.

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

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

4. Else, set $\hat{N} := \hat{N} \setminus \{i^*\}$, $\beta := \min\limits_{i \neq i^*, \theta_i}^{\rhd} \theta_i \mid \theta_i \in W_i^{h^*}$ and go to Stage 2.

**Stage 2**

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

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

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

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

6. If $\underline{\theta}_{\hat{i}}^{\hat{h}_{\hat{i}}} \unrhd \beta$, set $\beta := \underline{\theta}_{\hat{i}}^{\hat{h}_{\hat{i}}}$ 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 $\beta$.

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 $\theta_{\psi(\hat{i})} \rhd \theta_{\hat{i}}$, since $\theta_{\psi(\hat{i})}$ is $\hat{i}$-equivalent to $\beta$, the latter implies that $\beta \rhd \theta_{\hat{i}}$.[37] Moreover, since $(G, S_N)$ is orderly and has threshold pricing (by Proposition 16), 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 $(\theta^*_{i^*}, \underline{\theta}^{h^*}_{-i^*})$. Under $(G, S_N)$, $t_{i^*}(\theta^*_{i^*}, \underline{\theta}^{h^*}_{-i^*}) = \overset{\rhd}{\min} W^{h^*}_{i^*}$, 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 $W^{h^*}_{\psi(\hat{i})}$. (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 $\theta^{\text{old}}_{\psi(\hat{i})}$ and $\beta^{\text{old}}$ denote the simulated nemesis type and the best offer from the last time the algorithm communicated with $\hat{i}$. Let $\theta^{\text{new}}_{\psi(\hat{i})}$ and $\beta^{\text{new}}$ denote the current simulated nemesis type and best offer. Observe that we always revise the nemesis type upwards; $\beta^{\text{old}} \unlhd \beta^{\text{new}}$, so $\theta^{\text{old}}_{\psi(\hat{i})} \leq \theta^{\text{new}}_{\psi(\hat{i})}$. If $\theta^{\text{old}}_{\psi(\hat{i})} = \theta^{\text{new}}_{\psi(\hat{i})}$, we are done, since $(\theta^{\text{old}}_{\psi(\hat{i})}, \underline{\theta}^{h^*}_{N\setminus\{\hat{i},\psi(\hat{i})\}}) \in \Theta^{\hat{h}_{\hat{i}}}_{-\hat{i}}$. 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, $\theta^{\text{old}}_{\psi(\hat{i})}$ would win against $\underline{\theta}^{h'}_{\hat{i}}$, but would lose against $\underline{\theta}^{\hat{h}_{\hat{i}}}_{\hat{i}}$, i.e. $y(\underline{\theta}^{h'}_{\hat{i}}, \theta^{\text{old}}_{\psi(\hat{i})}, \underline{\theta}^{h^*}_{N\setminus\{\hat{i},\psi(\hat{i})\}}) = \psi(\hat{i}) \neq y(\underline{\theta}^{\hat{h}_{\hat{i}}}_{\hat{i}}, \theta^{\text{old}}_{\psi(\hat{i})}, \underline{\theta}^{h^*}_{N\setminus\{\hat{i},\psi(\hat{i})\}})$. By Proposition 18, there exists another history $h''$ in the same information set as $h'$, such that $(\theta^{\text{new}}_{\psi(\hat{i})}, \underline{\theta}^{h^*}_{N\setminus\{\hat{i},\psi(\hat{i})\}}) \in \Theta^{h''}_{-\hat{i}}$. 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 $(\theta_{\psi(\hat{i})}, \underline{\theta}^{h^*}_{N\setminus\{\hat{i},\psi(\hat{i})\}})$. Thus, the algorithm produces a profitable safe deviation. $\square$

---

[37]Since $(G, S_N)$ is orderly, we must eventually learn either that $\theta_{\hat{i}} \rhd \theta_{\psi(\hat{i})}$ 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 $\beta \rhd \theta_{\hat{i}}$.

We are now ready to show that, under the assumptions of Theorem 2, if $(G, S_N)$ is credible and strategy-proof, then $(G, S_N)$ is an ascending auction. With Propositions 16 and 19 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 inactive bidder is called to play.

3. An active bidder is called to play, and there are no other active bidders. (We leave this case till last - it can only happen when every other bidder has a type below the reserve.)

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 19 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 17, 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 orderly, if there is no high bidder, then all types weakly above the reserve $\rho$ place a bid. Similarly, all types above the current high bid place a bid.

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 $\triangleright$). 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 are called to play. Similarly, if $i$ is the current high bidder at history $h$ and there is another active bidder, then by Proposition 19, 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 - all the other bidders have types below the reserve. Let $I_i$ be such that $h \in I_i$.

In this case, we can define an action $a$ as **quitting** if there is no type above the reserve that plays $a$, that is:

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

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

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

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 17, each bid is weakly more than the last bid that $i$ placed.

The three conditions that specify what happens when the auction ends are similarly entailed by orderliness and threshold pricing (Proposition 16). 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 $\overset{\triangleright}{\min}\{\theta_i \mid \theta_i \triangleright \overset{\triangleright}{\max_{j \neq i}} \theta_j \text{ and } \theta_i \triangleright \rho\} = \overset{\triangleright}{\min}\{\theta_i \mid y(\theta_i, \theta_{-i}) = i\}$. 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 $\theta^K$ and no active bidder has higher tie-breaking priority, then $i$ must win and pay $\theta^K$, and since $(G, S_N)$ is pruned, $h$ is a terminal history.

This completes the proof that, under the assumptions of Theorem 2, if $(G, S_N)$ is credible and strategy-proof, then $(G, S_N)$ is an ascending auction.

### B.4.2 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 $\theta_i$ such that $\theta_i \in \Theta_i^{I_i}$. Recall that $S_i$ requires that $i$ quit if $\theta_i$ is strictly below the bid $b(I_i)$ at $I_i$, and that $i$ places the bid if $\theta_i$ 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 $\theta_i$ is

weakly above that bid. This is because $i$ can quit if the required bid ever rises strictly above $\theta_i$. Under any safe deviation, $i$ cannot be charged more than $\theta_i$ unless he (at some later point) bids more than $\theta_i$. 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. $b(I_i) \trianglelefteq \theta_i \lhd \min\{\rho, \overset{\triangleright}{\underset{h \in I_i, j \neq i}{\min}} \underline{\theta}_j^h\}$, 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 2.

## B.5  Proposition 7

Suppose $(G, S_N)$ is a twin-bid auction and strategy-proof. We drop the superscripts on $t_i^{G, S_N}$ and $y^{G, S_N}$ to reduce clutter. Strategy-proofness requires:

$$t_i(\theta_i', \theta_{-i}) - t_i(\theta_i, \theta_{-i}) \leq \theta_i' \tag{39}$$

$$\theta_i'' \leq t_i(\theta_i''', \theta_{-i}') - t_i(\theta_i'', \theta_{-i}') \tag{40}$$

$y^{G, S_N}$ is monotone, so $y(\theta_i, \theta_{-i}') \neq i$ and $y(\theta_i''', \theta_{-i}) = i$. It follows that:

$$t_i(\theta_i, \theta_{-i}) = t_i(\theta_i, \theta_{-i}') = t_i(\theta_i'', \theta_{-i}') \tag{41}$$

$$t_i(\theta_i''', \theta_{-i}') = t_i(\theta_i''', \theta_{-i}) = t_i(\theta_i', \theta_{-i}) \tag{42}$$

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 40 yields

$$\theta_i'' \leq t_i(\theta_i', \theta_{-i}) - t_i(\theta_i, \theta_{-i}) \tag{43}$$

which contradicts Equation 39.

## B.6 Theorem 4

### 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 2, 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 15, 16, 17, and 18 pin down some details even when $F_N$ is not symmetric. We start by proving an analogue to Proposition 19.

**Proposition 20.** *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 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 19 works here with the following modifications: First, we define

$$\psi(i) = \underset{j \in N \setminus \{i\}}{\operatorname{argmax}} \{\eta_j(\theta_j^{K_j}) \mid W_j^{h^*} \neq \emptyset\} \tag{44}$$

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

$$\{\theta_i \mid \eta_i(\theta_i) \geq \gamma\} = \{\theta_i \mid \eta_i(\theta_i) \geq \eta_j(\theta_{\psi(i)})\} \tag{45}$$

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 $\beta$.

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

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

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

$$\beta := \min_{i \neq i^*, \theta_i} \eta_i(\theta_i) \mid \theta_i \in W_i^{h^*} \tag{46}$$

and go to Stage 2.

**Stage 2**

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

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

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

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 $\eta_{\hat{i}}(\underline{\theta}^{\hat{h}_{\hat{i}}}_{\hat{i}}) \geq \beta$ 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 $\beta$.

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 $S_0^G$, 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 $\beta = \eta_{\psi(i^*)}(\theta_{\psi(i^*)}^{K_{\psi(i^*)}})$ or $\beta = \eta_{i^*}(\theta_{i^*}^*) < \eta_{\psi(i^*)}(\theta_{\psi(i^*)}^{K_{\psi(i^*)}})$. In the first case, $\theta_{\psi(i^*)}^{K_{\psi(i^*)}}$ will $i^*$-separate at $\beta$. In the second case, since $\eta_{i^*}(\theta_{i^*}^*) > \eta_{\psi(i^*)}(\theta_{\psi(i^*)}^1)$, by $F_N$ interleaved there exists $\theta_{\psi(i^*)}$ that will $i^*$-separate at $\beta$.

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

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 2 that invokes Proposition 18. Thus, the algorithm is well-defined, and produces a profitable safe deviation, which completes the proof. $\qquad\square$

With Proposition 20 in hand, we now complete the proof that, under the assumptions of Theorem 4, 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 and 16. At type profile $\theta_N$, 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 20 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 17, 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 20, 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 $h \in I_i$.

In this case, we can define an action $a$ as **quitting** if no type with a positive virtual value plays $a$, that is:

$$\neg \exists \theta_i \in \Theta_i^h \mid \eta_i(\theta_i) > 0 \text{ and } S_i(I_i, \theta_i) = a \tag{47}$$

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

$$\min\{\theta_i \mid \eta_i(\theta_i) > 0 \text{ or } [\theta_i \in \Theta_i^h \text{ and } S_i(I_i, \theta_i) = a]\} \tag{48}$$

In this case, $W_i^h \cap L_i^h = \emptyset$, so there can be multiple actions that place bids. Again, by Proposition 17, 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 16). Thus, under the assumptions of Theorem 4, if $(G, S_N)$ is credible and strategy-proof, then $(G, S_N)$ is a virtual ascending auction.

## B.7   Proposition 8

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.

Suppose $(G, S_N)$ is credible and extracts full surplus. By Propositions 1 and 2, 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 6). 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; $t_j^{G,S_N}(\theta_i, \theta_j) = t_j^{G,S_N}(\theta_i, \theta_j')$ and $t_j^{G,S_N}(\theta_i', \theta_j) = t_j^{G,S_N}(\theta_i', \theta_j')$.

Since $j$ always wins the object, the auctioneer can safely deviate to communicate with $j$ as though $i$'s type is $\theta_i$ or as though $i$'s type is $\theta_i'$. Since $(G, S_N)$ is credible, $j$'s transfer does not depend on $i$'s type; $t_j^{G,S_N}(\theta_i, \theta_j) = t_j^{G,S_N}(\theta_i', \theta_j)$. 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

56

extract full surplus, a contradiction.

## B.8 Proposition 9

To ease notation, we drop the superscripts on $t_i^{G,S_N}$ and $y_i^{G,S_N}$. Define $\nu(\theta_i, \theta_{-i}) = \theta_i y_i(\theta_i, \theta_{-i}) - t_i(\theta_i, \theta_{-i})$.

We can use the same method as Proposition 4 to derive an upper bound on $\nu(\theta_i, \theta_{-i})$ under *ex post* incentive compatibility and *ex post* individual rationality, namely:

$$\nu(\theta_i^k, \theta_{-i}) \geq \sum_{l=2}^{k} y_i(\theta_i^{l-1}, \theta_{-i})(\theta_i^l - \theta_i^{l-1}) \tag{49}$$

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

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

$$= \sum_{k=1}^{K} f_i(\theta_i^k | \theta_{-i}) \frac{1 - F_i(\theta_i^k | \theta_{-i})}{f_i(\theta_i^k | \theta_{-i})} (\theta_i^{k+1} - \theta_i^k) y_i(\theta_i^k, \theta_{-i}) \tag{50}$$

which gives an upper bound on expected revenue

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

$$= \sum_{i \in N} \mathbf{E}_{\theta_{-i}} \left[ \mathbf{E}_{\theta_i} \left[ \theta_i y_i(\theta_N) - \nu(\theta_i, \theta_{-i}) \mid \theta_{-i} \right] \right]$$

$$\leq \sum_{i \in N} \mathbf{E}_{\theta_{-i}} \left[ \mathbf{E}_{\theta_i} \left[ \eta_i(\theta_i | \theta_{-i}) y_i(\theta_N) \mid \theta_{-i} \right] \right] = \mathbf{E}_{\theta_N} \left[ \sum_{i \in N} \eta_i(\theta_i | \theta_{-i}) y_i(\theta_N) \right] \tag{51}$$

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 $\theta_{-i}$.

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.

**Proposition 21.** *If $f_N$ is affiliated and $\theta_j < \theta_j'$, then $\eta_i(\theta_i | \theta_j, \theta_{N \setminus \{i,j\}}) \geq \eta_i(\theta_i | \theta_j', \theta_{N \setminus \{i,j\}})$*

*Proof.* 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.

$$\frac{1 - F_i(\theta_i | \theta_j, \theta_{N \setminus \{i,j\}})}{f_i(\theta_i | \theta_j, \theta_{N \setminus \{i,j\}})} \leq \frac{1 - F_i(\theta_i | \theta_j', \theta_{N \setminus \{i,j\}})}{f_i(\theta_i | \theta_j', \theta_{N \setminus \{i,j\}})} \tag{52}$$

which implies $\eta_i(\theta_i|\theta_j, \theta_{N\setminus\{i,j\}}) \geq \eta_i(\theta_i|\theta_j', \theta_{N\setminus\{i,j\}})$. $\qquad\square$

**Proposition 22.** *Assume $f_N$ is symmetric, regular, and affiliated. For all $\theta_{N\setminus\{i,j\}}$, if $k \geq k'$, then $\eta_i(\theta_i^k|\theta_{N\setminus\{i,j\}}, \theta_j^{k'}) \geq \eta_j(\theta_j^{k'}|\theta_{N\setminus\{i,j\}}, \theta_i^k)$.*

*Proof.*

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

$$\geq \theta_i^{k'} - \frac{1 - F_i(\theta_i^{k'}|\theta_{N\setminus\{i,j\}}, \theta_j^{k})}{f_i(\theta_i^{k'}|\theta_{N\setminus\{i,j\}}, \theta_j^{k})}(\theta_i^{k'+1} - \theta_i^{k'}) = \eta_j(\theta_j^{k'}|\theta_{N\setminus\{i,j\}}, \theta_i^{k}) \quad (53)$$

where the first inequality follows from regularity, the second inequality follows from Proposition 21, and the equality follows from symmetry. $\qquad\square$

By Proposition 22, the right-hand side of Equation 51 is maximized by, at each $\theta_N$, 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 51 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 $\theta_{-i}$. 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.

## B.9 Proposition 11

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 $\theta_N$. 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 6.1 results in the same allocation.

Take the greedy algorithm's $j$th 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. Take any step at which $k$'s bid is equal to his type, and $k$'s score is minimal in $\hat{N}$. By construction, $\hat{N} \subseteq \{1, 2, \ldots, k\}$, since bidders with lower ironed virtual values have either been put in the allocation or quit.

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 $|Y| < j$.

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

## B.10  Proposition 12

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 $\theta_{N\setminus i}'$, type $\theta_i''$ can profitably imitate $\theta_i'''$, a contradiction.