Theoretical Economics 16 (2021), 1655–1714 1555-7561/20211655

# Bilateral trade with a benevolent intermediary

Ran Eilat
Department of Economics, Ben Gurion University of the Negev

Ady Pauzner
The Eitan Berglas School of Economics, Tel-Aviv University

We study intermediaries who seek to maximize gains from trade in bilateral negotiations. Intermediaries are players: they cannot commit to act against their objective function and deny, in some cases, trade they believe to be beneficial. This impairs their ability to assist the parties relative to conventional mechanisms. We analyze this limited commitment environment as a standard mechanism design problem with an additional "credibility" constraint, requiring that every outcome be interim-optimal conditional on available information. We investigate how such intermediaries communicate with the parties, analyze the tradeoffs they face, and study the bounds on what they can achieve.

Keywords. Intermediation, mechanism design, imperfect commitment, asymmetric information, bilateral trade.

JEL classification. C72, D82, D83.

## 1. Introduction

Bilateral negotiations are often facilitated by an intermediary whose goal is to bridge information gaps and lead the parties to a desired outcome. Examples include peace negotiations, divorce proceedings, and real estate transactions. While intermediaries may have various motives, such as maximizing their own profit or building up a reputation, helping the parties to materialize potential gains from trade is often the main one. In this paper, we abstract from selfish motives and study "benevolent" intermediaries—whose sole goal is to maximize social surplus—in environments of asymmetric information. We investigate how such intermediaries communicate with the parties and make decisions and to what extent they can help the parties realize the potential social surplus.

The related question, of how to design a mechanism that maximizes the social surplus, has been extensively studied in the literature. However, while some insights carry over, the problem of optimal intermediation is quite different from that of optimal mechanism design. In (conventional) mechanism design, the designer determines the communication and decision policies at the outset and commits to execute them even if they turn out to be suboptimal when information is revealed. Such commitment

Ran Eilat: eilatr@bgu.ac.il
Ady Pauzner: pauzner@tauex.tau.ac.il
This paper supersedes an earlier working paper entitled "Mediators and bilateral trade."

 https://doi.org/10.3982/TE3703

power can be achieved, for example, by delegating the execution to a separate institution or to a hard-wired machine. In contrast, an intermediary actively participates in the negotiations—he interacts with the parties and makes decisions based on the information they provide. Unlike a mechanism designer, he is bound by his preferences when communicating with the agents and when deciding on the outcome. In particular, a benevolent intermediary cannot credibly commit to sometimes act against the interests of the parties.

We study intermediation in the canonical bilateral trade environment of Myerson and Satterthwaite (1983). This simplified environment captures one of the main difficulties in realizing the potential social surplus in settings with asymmetric information. In this setup, a seller owns an object that a buyer potentially wants to purchase. Valuations are independently drawn, and each agent's valuation is known only to himself. The efficient outcome is attained if the parties trade whenever the buyer has the higher valuation. Myerson and Satterthwaite (1983) showed that unless the problem is trivial, this (first-best) outcome is unattainable due to the agents' incentives to misreport their valuations in order to obtain better trading terms. They also characterize the optimal (second-best) mechanism, which maximizes the gains from trade among all the feasible ones. This mechanism denies trade when the buyer's valuation is only slightly above the seller's. The commitment to do so weakens the agents' incentives to misreport their preferences and increases the social surplus on average.

An intermediary in the bilateral trade problem also strives to maximize the gains from trade. Unlike a mechanism, however, he cannot commit to deny trade in cases where trade is *known* to be beneficial. To capture that idea, we model him as player in a game. In this game, the intermediary moves first and decides on the message sets for the buyer and the seller. The buyer and seller then simultaneously choose messages or decide to quit. If no agent quits, the intermediary updates his beliefs given the reported messages. He then makes a decision—whether the object is traded and at what price—in a way that maximizes the interim expected social surplus given his beliefs.

In order to investigate intermediation, we convert this limited-commitment design game into a standard (full-commitment) mechanism-design problem, in which the set of permissible mechanisms is restricted to those satisfying a property we call credibility, according to which every outcome of the mechanism is interim-optimal conditional on the available information.[1] This allows us to take advantage of some (though not all) of the powerful tools and techniques of standard mechanism design.

We show that in order to credibly deny beneficial trade, the intermediary restricts the precision of the information he collects from the agents. He does so by communicating in a coarse language that pools many types of each agent in the same report. By coupling realizations in which trade is beneficial with realizations in which it is not within the same information set, the intermediary can sometimes deny beneficial trade and thereby incentivize the agents to reveal their private information. The downside is,

[1]This should not be confused with ex post efficiency, which requires the outcome to be optimal with respect to the agents' true types, rather than with respect to the posterior beliefs that are generated from their messages.

however, that doing so limits the intermediary's ability to make the trading decision dependent on the parties' exact preferences. Thus, in contrast to a mechanism designer, for whom information is always (weakly) helpful, an intermediary faces a trade-off in which possessing finer information also reduces the set of outcomes to which he can credibly commit.[2]

We find that even though the type space is continuous, the intermediary employs a finite language. We also study the bounds on the surplus that intermediation can generate. This surplus is *strictly* below that of the (second-best) optimal mechanism with full commitment power and weakly above that of the (optimal) posted-price outcome. We also show that, in the case that types are uniformly distributed, the prospects for effective intermediation are grim: the number of messages per agent reduces to only two, and the intermediary does no better than a posted price. Finally, we show that when the designer's resource constraint is relaxed (i.e., trade is subsidized), the negative consequences of his limited commitment are less severe.

The communication phase in our model is one-shot, such that each agent sends only one message before a decision is made. While this assumption is not without loss of generality, it allows us to focus the analysis on the core issue of the intermediary's use of coarse language. Adding stages of communication may indeed expand the set of achievable outcomes since this would allow the intermediary more freedom in controlling the flow of information, but would not change the basic intuitions delivered by the simpler model.

Beyond studying intermediation per se, the paper contributes to the theoretical literature on mechanism design with limited commitment. The environment we study—one with multiple privately informed agents—has been largely understudied. One notable complication of this type of environment is the failure of the revelation principle. Indeed, the "classic" revelation principle fails already in single agent limited-commitment environments, but several results have recovered weaker versions of it for the single-agent case. These remedies, however, are not valid for more than one privately informed agent.[3] In this paper, we develop a coarse revelation result that is applicable in our environment. We show that it is without loss of generality to assume that each agent's message set partitions his type space into intervals (and singletons) and that in equilibrium each agent truthfully reports the element to which his type belongs. The key feature that drives this result in our multisender environment is a form of monotonicity, and explained in Section 3.

Another special feature of the multiple agent environment relates to the information trade-off described earlier. While the use of a coarse language is common in limited commitment environments, in the multiagent environment the languages that the intermediary uses to communicate with the buyer and the seller are related. In our setting,

[2]The fact that avoidance of information can be used as a commitment device is well known in economics. See, for example, Cremer (1995) in the context of repeated contracting, Carrillo and Mariotti (2000) in the context of self-control, and Bénabou and Tirole (2002) in the context of self-confidence and motivation. Golman, Hagmann, and Loewenstein (2017) provided a survey of the literature.

[3]See, for instance, Bester and Strausz (2001, 2007), Doval and Skreta (2020), and Krishna and Morgan (2008) for variants of the revelation principle in the single-agent case and Bester and Strausz (2001) and Evans and Reiche (2008) for examples of the failure in the multiagent case.

we show that if in some range of valuations the intermediary collects fine information from one agent, then he must do so also for the other agent. While this result is specific to our model, it highlights the fact that considerations regarding the information trade-off are more intricate in the multiagent case.

The paper is organized as follows. In Section 2, we present the game-theoretic model and show the equivalence to a constrained mechanism-design problem. In Section 3, we derive the coarse revelation result. Section 4 explores properties of optimal intermediation mechanisms. In Section 5, we analyze optimal intermediation, its form and the bounds on the surplus that it can achieve. Section 6 concludes. Proofs are relegated to the Appendices.

*Related work* Communication in our model is cheap talk, i.e., the information transmitted by the agents is costless and nonverifiable. In their seminal paper on cheap talk, Crawford and Sobel (1982) showed that the misalignment of interests between the sender and the receiver necessarily leads to a coarse information structure in which a continuum of types pool into intervals. Krishna and Morgan (2008) enriched the Crawford and Sobel (1982) model by allowing for contractual monetary transfers between the principal and a single informed agent. These transfers allow the principal to fully separate the types in one end of the type space, whereas the other types pool like in Crawford and Sobel (1982). Our model also allows for transfers (between the buyer and the seller) that the intermediary controls. Thus, like in Krishna and Morgan (2008), the intermediary is able to fully separate the agents' types in some region. However, he never chooses to do so and the optimal information structure he employs is always coarse.

An extensive body of literature that followed Crawford and Sobel (1982) considered the case of strategic information transmission with multiple senders who are not symmetrically informed, as in our model. Notable examples are Austen-Smith (1993) and Wolinsky (2002) who study how different communication protocols affect equilibrium outcomes and Gerardi, McLean, and Postlewaite (2008) who consider a receiver who can commit to distorting the outcome, relative to his optimal rule, in order to incentivize the senders to reveal information.[4] See Sobel (2013) for a detailed discussion of this strand of the literature.

The intermediary's objective in our model is to assist the parties in overcoming their information asymmetry and realizing potential gains from trade. As such, our model is related to the literature on information mediators as settlement facilitators. In early contributions to this literature, Forges (1986) and Myerson (1986) showed that third parties (or communication devices) can act as information mediators and can expand the set of equilibria in games. In more recent contributions, Goltsman et al. (2009) compared various dispute resolution institutions in a Crawford–Sobel framework while Hörner, Morelli, and Squintani (2015) compare the performance of third parties as settlement facilitators with and without the ability to enforce their recommendations. Other notable examples include Blume, Board, and Kawamura (2007), Fey and Ramsey (2009), and Ivanov (2010).

[4]For the case of symmetrically informed senders see, for example, Gilligan and Krehbiel (1989), Krishna and Morgan (2001), and Battaglini (2002).

A number of papers have studied the problem of multiperiod auction design with imperfect commitment. Notable examples are McAfee and Vincent (1997) and Skreta (2006, 2015). In these papers, the seller can commit to the mechanism offered in the current period but cannot commit not to offer a new mechanism if the item remains unsold.[5] In our paper, the focus is on one-period mechanisms, but the designer can change the rules within that period after the agents send their reports. Akbarpour and Li (2020) considered a different type of imperfect commitment model in which an auctioneer who communicates sequentially and privately with the buyers can deviate from the pre-determined rules only if the deviation is undetectable. In contrast, the designer in our model is not concerned with whether or not his deviations can be detected. Other notions of credibility requirements were studied in a context of Nash implementation (rather than Bayesian mechanism design) by Baliga, Corchon, and Sjostrom (1997) and by Chakravorty, Corchon, and Wilkie (2006).

Other papers endow the seller with even less power to commit. In McAdams and Schwartz (2007), a seller sequentially offers an item to multiple buyers and cannot commit not to conduct another round of bids, but in contrast to our model the seller cannot affect the buyers' strategy space. Vartiainen (2013) considered a sequential auctioning model in which a seller can use a communication device to extract information from the buyers and can change the rules of the game as long as the physical transaction has not yet taken place. One key distinction between our paper and the above mentioned body of literature is that we assume that the designer's objective is to maximize welfare whereas most of the existing literature focuses on revenue maximization.

The fact that the conventional revelation principle fails to hold when the designer's commitment power is imperfect—as in our model—is well known in the literature. In a setting with one agent and finite type space, Bester and Strausz (2001) show that a weaker version of the revelation principle applies.[6] If, in addition, the designer has access to a communication device that may add noise to the agent's report, Bester and Strausz (2007) show that it is without loss of generality to assume that the agent reveals his type truthfully to the communication device. Doval and Skreta (2020) provided a revelation principle for dynamic settings in which the designer has access to a communication device and can commit only to short-term contracts. In this paper, we do not allow the designer to use a communication device, and the (coarse) revelation result we develop arises from the credibility property of the mechanism. We discuss some implications of introducing a communication device into multiagent environments in Section 6.

Finally, our model is also related to the literature on renegotiation proofness in settings with asymmetric information (see, e.g., Tirole (1986), Laffont and Tirole (1988),

[5]Zheng (2002) considered a model in which the lack of commitment is on the buyers' side, such that they cannot commit not to resell the object, and proposes a seller-optimal auction.

[6]Bester and Strausz (2001) showed that the designer may optimally use a direct mechanism under which truthful revelation is an optimal strategy for the agent, but unlike in the case of the conventional revelation principle, the agent cannot use this strategy with probability one. This result extends to multiagent environments when only one agent is privately informed (see Evan and Reiche (2008)) but not to the case in which several agents have private information, like the one in our model (see Bester and Strausz (2000)).

Dewatripont and Maskin (1990), and more recently Neeman and Pavlov (2012)). While there are various definitions of renegotiation proofness, the idea is that after the agents have played, there is no alternative mechanism that can improve on the realized outcome for at least some of the types. Our model differs in that the intermediary's decision is final, i.e., he does commit not to launch another round of communication. Thus, the outcome maximizes social surplus only with respect to the intermediary's limited knowledge. In particular, the outcome need not be ex post efficient, even though utility is transferable.

## 2. The intermediation game and mechanism

We begin by defining intermediation as a three-player game between the buyer, the seller, and the intermediary. We then introduce the concept of intermediation mechanisms, which are standard bilateral trade mechanisms à la Myerson and Satterthwaite (1983), augmented by a credibility constraint that requires all outcomes to be interim-optimal given the equilibrium beliefs. Finally, we provide an equivalence result.[7]

### 2.1 *The intermediation game*

Intermediation in the bilateral trade problem is represented by a game of three players: a seller (agent $s$), a buyer (agent $b$), and an intermediary. The intermediary communicates with the agents to decide on an *outcome* $x = (p, t)$ where $p \in [0, 1]$ is the probability that the object is transferred and $t \in \mathbb{R}$ is the monetary transfer from the buyer to the seller (which may be nonzero even if the object is not transferred). The value of the object to agent $i \in \{s, b\}$ is $v_i$ and the agents are risk neutral. Thus, if the outcome $(p, t)$ is chosen, the seller's payoff is $-p \cdot v_s + t$, and the buyer's payoff is $p \cdot v_b - t$. Each agent has an option to quit the game, in which case there is no trade or transfer. The intermediary's utility is the sum of the agents' utilities, i.e., $v_b - v_s$ if the object is transferred to the buyer and 0 otherwise.

For each agent $i$, the valuation $v_i$ is drawn independently from a distribution $F_i$ over $V_i = [\underline{v}_i, \bar{v}_i]$ and is privately known to the agent. We assume that $F_i$ admits a density $f_i$ that is strictly positive, continuously differentiable and bounded over the interval $V_i$. For nontriviality, we assume that the intersection $V_s \cap V_b$ is nonempty.

The timing of the game is as follows.

- Stage 1 The intermediary specifies a set $M_i$ of possible messages for each agent.[8] Each $M_i$ must contain a message labeled "out."
- Stage 2 The agents simultaneously choose messages $m_i \in M_i$ (mixing is allowed). If one of the messages is "out," then the game terminates without trade or payments.

[7] The idea of characterizing the game's outcomes using a mechanism design approach has precedents in the literature. See, for instance, Myerson (1979) and Ausubel and Deneckere (1989).

[8] Throughout the paper, message sets are restricted to be Polish spaces. We endow each Polish space with its Borel $\sigma$-algebra. Product spaces are endowed with the product $\sigma$-algebra. For a Polish space $Y$, we let $\Delta Y$ denote the set of all Borel probability measures over $Y$, endowed with the weak* topology.

Stage 3 The intermediary decides on the *outcome* $x = (p, t)$, and payoffs are realized.

We refer to the part of the game starting at stage 2 as the *reporting subgame*.

The solution concept is Perfect Bayesian Equilibrium (PBE). We impose the refinement that the equilibrium chosen in the reporting subgame is the best for the intermediary, i.e., the one that maximizes the expected social surplus.[9] This refinement rules out unreasonable equilibria in which the intermediary is "forced" to choose undesirable message sets simply because the agents would otherwise coordinate on a bad equilibrium in the reporting subgame (e.g., the babbling equilibrium).

Note that agents are allowed to opt out after the intermediary announces the set of messages,[10] but not after he declares the final outcome, i.e., his decision is binding. We discuss the case of nonbinding recommendations in Section 6.

## 2.2 *Intermediation mechanisms*

A *trade mechanism* $\Gamma$ consists of two sets of messages, $M_s$ for the seller and $M_b$ for the buyer, an allocation rule $p$, and two transfer rules $t_s$ and $t_b$. For each message pair $m = (m_s, m_b) \in M \equiv M_s \times M_b$, the allocation rule $p : M \to [0, 1]$ determines the probability that the object is traded while the transfer rule $t_i : M \to \mathbb{R}$ determines the monetary transfer (positive or negative) to agent $i \in \{b, s\}$. The mechanism is required to satisfy *ex post budget balance*:

$$t_b(m) + t_s(m) = 0 \tag{BB}$$

for every $m \in M$.

The utility of the seller of type $v_s$ when message pair $m = (m_s, m_b)$ is reported is $u_s(v_s; m) = -p(m)v_s + t_s(m)$ while the utility of the buyer of type $v_b$ is $u_b(v_b; m) = p(m)v_b + t_b(m)$. The social surplus is $W((v_s, v_b), m) = (v_b - v_s)p(m)$.

A strategy for agent $i$ is a measurable function $\sigma_i : V_i \to \Delta(M_i)$ that maps each of the agent's types to a distribution over messages. A Bayesian Nash Equilibrium (BNE) is a pair of strategies $\sigma = (\sigma_s, \sigma_b)$ such that each is a best response to the other. For convenience, we slightly abuse terminology by referring to $\Gamma$, together with its equilibrium $\sigma$, as a "trade mechanism." For simplicity, we assume that all the messages in $M_i$ are "on path," i.e., each message $m_i \in M_i$ is in the support of $\sigma_i(v_i)$ for some $v_i \in V_i$. This assumption does not change the set of implementable social choice functions (see the definition below and proof of Proposition 1).

[9]This refinement is equivalent to allowing the intermediary to recommend an equilibrium in stage 1, along with the refinement that the agents follow his recommendation.

[10]Allowing the agents to opt out also before message sets are chosen would not change the implemented social choice function. One can verify that an equilibrium in our model can be transformed into an equilibrium in a modified game in which agents can opt out at stage 0. In this equilibrium, all types opt in at stage 0 and then follow their original strategy. Conversely, an equilibrium in the modified game can be transformed into an equilibrium in our original game by simply adjusting the strategies of all types who opt out at stage 0 (in the modified game) so that they opt out at stage 2 (in the original game).

Given the trade mechanism $(\Gamma, \sigma)$, we denote the expected utility of agent $i$ of type $v_i$ who reports the message $m_i$ by $\bar{u}_i(v_i, m_i) = \mathbb{E}_{m_{-i}} u_i(v_i; m_i, m_{-i})$, where $\mathbb{E}_{m_{-i}}$ is evaluated according to the distribution over $M_{-i}$ induced by agent $-i$'s equilibrium strategy $\sigma_{-i}$. His expected probability of trade is $\bar{p}_i(m_i) = \mathbb{E}_{m_{-i}} p(m_i, m_{-i})$. We require trade mechanisms to satisfy *individual rationality*:

$$\bar{u}_i(v_i, m_i) \geq 0 \tag{IR}$$

for every $v_i \in V_i$, and every $m_i$ in the support of $\sigma_i(v_i)$, and for each agent $i$.

For every message pair $m = (m_s, m_b)$, the *interim surplus* induced by $m$ is given by $W_I(m) = \mathbb{E}_{v_s, v_b}[v_b - v_s|m] \cdot p(m)$, where $\mathbb{E}_{v_i}(v_i|m)$ is the posterior mean type of agent $i$ given that he reports $m_i$ in equilibrium. The *ex ante surplus* is given by $W_{\text{EA}} = \mathbb{E}_m W_I(m)$ where $\mathbb{E}_m$ is evaluated according to the distribution over $M$ induced by the equilibrium strategies. In the case of a pure-strategy equilibrium, in which the agents do not randomize, we have that $W_{\text{EA}} = \mathbb{E}_{v_s, v_b}[(v_b - v_s) \cdot p(\sigma_s(v_s), \sigma_b(v_b))]$.

We say that message $m_i$ is an *opt-out message* for agent $i$ in the mechanism $\Gamma$ if $p(m_i, m_{-i}) = t(m_i, m_{-i}) = 0$ for all $m_{-i} \in M_{-i}$. That is, by sending an opt-out message (if such a message exists), an agent can secure a payoff of zero.

A trade mechanism $(\Gamma, \sigma)$ is said to be *credible* if the allocation and transfer rules maximize the interim social surplus. As we show later, this captures the fact that in the intermediation game the allocation and transfer are chosen only after the agents play, i.e., that the intermediary cannot commit to the outcome. However, since in our quasi-linear setup the interim surplus $W_I(m)$ does not depend on the transfers, the credibility restriction involves only the allocation rule. Thus, a trade mechanism is credible if

$$p(m_s, m_b) \in \arg \max_{p' \in [0,1]} \mathbb{E}_{v_s, v_b}[v_b - v_s|m_s, m_b] \cdot p' \tag{CRED}$$

for every profile of messages $(m_s, m_b) \in M$, unless either $m_s$ or $m_b$ is an opt-out message. For a discussion of the necessity of the opt-out messages in our model and their relation to credibility and individual rationality, see Section 6. An *intermediation mechanism* is a credible trade mechanism. An *optimal intermediation mechanism* is one that maximizes the ex ante surplus over all intermediation mechanisms.

The designer is tasked with devising the optimal intermediation mechanism. His problem can be formulated as follows: Find a mechanism $\Gamma = \langle M, p, t_b, t_s \rangle$ and an equilibrium $(\sigma_s, \sigma_b)$ that maximize the ex ante surplus $W_{\text{EA}} = \mathbb{E}_m \mathbb{E}_{v_s, v_b}[v_b - v_s|m] \cdot p(m)$, subject to (BB), (IR), and (CRED).[11]

### 2.3 *An equivalence result*

Each equilibrium in the intermediation game, or in an intermediation mechanism, induces a social choice function $scf : V \to \Delta X$ from the set of types $V = V_s \times V_b$ to the set

[11]While the domain of mechanisms over which the designer maximizes is very large—in fact, a proper class in the terms of NBG set theory—the designer's problem is well-defined since maximization is well defined over classes.

of distributions over outcomes $X = [0, 1] \times \mathbb{R}$. We say that a social choice function $scf$ is *implementable* by the game (the reporting subgame) if there exists an equilibrium in the game (the reporting subgame) such that for any type-pair $v \in V$ the probability distribution it induces over outcomes is $scf(v)$. Similarly, we say that $scf$ is implementable by an intermediation mechanism if, for any type-pair $v \in V$, the probability distribution that its equilibrium induces over outcomes is $scf(v)$.

The following proposition provides an equivalence result between intermediation games and intermediation mechanisms.

Proposition 1. *For any social choice function $scf$:*

*(i) $scf$ is implementable by the reporting subgame starting with message set $M$ if and only if it is implementable by an intermediation mechanism with message set $M$.*[12]

*(ii) $scf$ is implementable by the intermediation game if and only if it is implementable by an optimal intermediation mechanism.*

Intuitively, the equivalence in the first part of the proposition is a result of the credibility restriction on the mechanism, which binds the outcome to be the interim-surplus maximizer and thus matches the decision of the (surplus-maximizing) intermediary in the third stage of the game. The equivalence in the second part follows from the equilibrium refinement that selects the intermediary-optimal equilibrium in any subgame.

## 3. Partition-direct representation

In standard mechanism design, the revelation principle implies that it is without loss of generality to restrict attention to direct-revelation mechanisms, in which each agent truthfully reports his type. This is no longer the case for intermediation mechanisms: if each agent truthfully reported his type, then equilibrium beliefs would become degenerate (with a single atom on the exact type of each agent), and credibility would then dictate fully efficient trade, which is infeasible. This result comes at no surprise in light of the existing literature on mechanism design with limited commitment. In particular, Bester and Strausz (2000) and Evans and Reiche (2008) highlighted the particular difficulty in restoring variants of the revelation principle in environments where more than one agent has private information, which is the case here.

In this section, we develop a coarse version of the revelation principle that applies to intermediation mechanisms. We show that it is without loss of generality to assume that each agent's message set forms a partition of his type space and that the equilibrium is truthful: each agent reports the element of the partition to which his type belongs. Moreover, each message is either an interval or a singleton. We refer to such a mechanism as a *partition-direct* mechanism.

[12]To be fully rigorous, since we exclude intermediation mechanisms with messages that are not sent on the equilibrium path, a slight adaptation of the statement is needed if the game has such messages. In this case, the mechanism's message set is simply defined by excluding those messages from $M$.

To clarify the scope of this contribution, note that the standard revelation principle has two features. First, that without loss communication is one-shot; second, that each agent's message set can be his set of types and that agents report truthfully. Our model assumes one round of communication at the outset. Thus, the contribution of our partial revelation result lies only within the domain of the second aspect.

### 3.1 *Monotonicity, minimality and pure strategies*

Consider a trade mechanism $(\Gamma, \sigma)$ where $\Gamma = \langle M, p, t_b, t_s \rangle$. We say that $(\Gamma, \sigma)$ is message-monotone if, given two messages of agent $i$ that induce different posterior mean types for agent $i$, if one of them leads to a higher trade probability for *some* message of agent $-i$, then it must lead to a (weakly) higher trade probability for *any* message of agent $-i$:

Definition 1 (Message monotonicity). $(\Gamma, \sigma)$ is message-monotone if $p(m_i, m_{-i}) > p(m_i', m_{-i})$ for some $m_{-i} \in M_{-i}$ implies $p(m_i, m_{-i}') \geq p(m_i', m_{-i}')$ for all $m_{-i}' \in M_{-i}$, for any $m_i, m_i' \in M_i$ such that $\mathbb{E}_{v_i}[v_i|m_i] \neq \mathbb{E}_{v_i}[v_i|m_i']$.

Message monotonicity is the key property that generates our coarse-revelation result in Section 3.2 (see the discussion following Proposition 2 below). The credibility of intermediation mechanisms guarantees that they are message monotone. Intuitively, since types are independent, agent $i$'s posterior mean type depends only on the message that he sent. Thus, if two messages $m_i$ and $m_i'$ induce different posterior mean types, i.e., $\mathbb{E}_{v_i}[v_i|m_i] \neq \mathbb{E}_{v_i}[v_i|m_i']$, then $p(m_i, m_{-i}) > p(m_i', m_{-i})$ is consistent with credibility only if $\mathbb{E}_{v_i}[v_i|m_i] > \mathbb{E}_{v_i}[v_i|m_i']$. But then credibility also implies that $p(m_i, m_{-i}') \geq p(m_i', m_{-i}')$ for all other $m_{-i}'$. Formally, we have the following.

Lemma 1. *An intermediation mechanism is message monotone.*

We say that two intermediation mechanisms are *payoff-equivalent* if each type of each agent obtains the same expected payoff under both. This also implies that the ex ante social surplus is identical. We say that an intermediation mechanism is *minimal* if for each agent $i$ there are no two messages that induce the same expected probability of trade, i.e.,

$$\bar{p}_i(m_i) \neq \bar{p}_i(m_i') \quad \text{for every two messages } m_i \neq m_i'. \tag{MIN}$$

We then have the following.

Lemma 2. *For any intermediation mechanism, there exists a payoff-equivalent minimal intermediation mechanism.*

Thus, any level of ex ante social surplus that can be attained by an intermediation mechanism can also be attained by a minimal one. Consequently, when seeking for the optimal intermediation mechanism we can restrict attention to minimal intermediation mechanisms. The key argument in the proof is that, due to message monotonicity, when $\bar{p}_i(m_i) = \bar{p}_i(m_i')$ it must be that either $\mathbb{E}_{v_i}[v_i|m_i] = \mathbb{E}_{v_i}[v_i|m_i']$ or $m_i$ and $m_i'$ lead to the same trade probabilities for *every* message of agent $-i$, i.e., $p(m_i, m_{-i}) = p(m_i', m_{-i})$ for every $m_{-i}$. We can therefore merge all the messages with the same $\bar{p}_i$ and update agent $i$'s strategy so that all types who sent any of these messages now send the merged message. We show that credibility is satisfied and adjust the transfer rules to support the equilibrium. The expected payoff of each type of each agent remains unchanged.

The minimality of the intermediation mechanism guarantees that each agent's messages can be identified with the expected probabilities of trade that they induce. This, along with the single-crossing property of the agents' preferences, leads to our next result.

Lemma 3. *In a minimal intermediation mechanism, almost all types of each agent do not randomize. Moreover, for any minimal intermediation mechanism there exists a payoff-equivalent minimal intermediation mechanism in which* all *types do not randomize.*

Note that the result that agents use pure strategies in equilibrium is not obvious in models of mechanism design with imperfect commitment. In fact, it is well known that in some settings the opposite is true. For example, Bester and Strausz (2000) showed that in a contracting problem, when there are multiple agents and the designer cannot fully commit to an allocation function, the optimal contract is sometimes achieved when the set of messages is strictly greater than the set of types and some types randomize.

### 3.2 *A coarse-revelation result*

Since it is without loss of generality to assume that the agents play pure strategies, we can now partition each agent's set of types according to the messages they send in equilibrium and rename each message to be the set of types that send it.[13] Due to the single-crossing property of the preferences, the set of types that send each message in the original mechanism is convex. Therefore, each element $m_i \in M_i$ in the modified mechanism (i.e., after messages are renamed) is either a singleton or an interval of types. Clearly, it is a best response for each type of each agent to report "truthfully," i.e., to report the message to which it belongs. We can therefore restrict attention to mechanisms in which messages are intervals (or singletons) that partition the type space of each agent and agents report truthfully. We now state this observation formally.

Given message set $M_i$ whose elements form a partition of $V_i$, we say that agent $i$'s strategy is *truthful* if every type $v_i$ reports the message $m_i \in M_i$ such that $v_i \in m_i$. An equilibrium that consists of truthful strategies is a *truth-telling* equilibrium.

Definition 2 (Partition-Direct). An intermediation mechanism is *partition-direct* if: (i) each agent's message set is a partition of his type-space, (ii) for each agent, each message is a set containing one type or an interval of types, and (iii) truth-telling is an equilibrium.

[13]A message can always be identified with the posterior belief it induces in equilibrium, but when strategies are pure (i.e., when types do not randomize) the supports of the beliefs induced by the various messages are disjoint and, therefore, each message can be identified with the support itself.

Proposition 2. *For any intermediation mechanism, there exists a partition-direct intermediation mechanism that is payoff-equivalent.*

It is worthwhile to highlight the features of our model that drive this result, which is not obvious in cheap-talk environments with more than one privately-informed agent. To that end, consider first the single-agent models of Crawford and Sobel (1982) and Krishna and Morgan (2008). In these models, there is no loss of generality in assuming that different sender messages induce different receiver actions in equilibrium. This is because messages that induce the same action can simply be merged (with all the types that reported any of those messages now reporting the merged one). Since different messages induce different actions, a standard argument using the single-crossing property of the agent's preferences guarantees that the equilibrium is essentially pure.

In our multiple agent setting, things are more subtle. Here, the action of the intermediary (the receiver) depends on the reports of both the buyer and the seller (the senders). However, each agent is only concerned with one aspect of the intermediary's action—the expected probability of trade given the agent's report.[14] It is thus possible, a priori, that there are distinct messages that induce the same expected probability of trade, while for the intermediary these messages are not the same as they involve different actions following different messages of the other agent. Thus, even though the messages are the same from the agent's perspective, it is not obvious that they can be merged (while maintaining credibility).

The property of our model that allows us to merge the messages is message monotonicity (Lemma 1). It guarantees that if an agent has different messages with the same expected probability of trade, then it is either the case that for each of these messages the intermediary has the same trading decision for any message of the other agent, or that these messages are sent by two sets of types with the same mean. In both cases, the messages can be merged without violating credibility (Lemma 2). The partial revelation result then follows from the usual single-crossing arguments.

## 4. Toward optimal intermediation mechanisms

With the coarse revelation result from Section 3 in hand, we can now seek the optimal intermediation mechanism within the class of partition-direct intermediation mechanisms. Since equilibrium strategies of partition-direct mechanisms are fully defined by the message sets, from now on we simplify notation by omitting the explicit reference to the equilibrium $\sigma$ from the definition of an intermediation mechanism.

### 4.1 *Credible minimality and the shape of trading rules*

Given a partition-direct intermediation mechanism $\Gamma = \langle M, p, t_b, t_s \rangle$, credibility implies that the allocation rule $p(m_s, m_b)$ maximizes the interim surplus for any pair of messages $(m_s, m_b)$ (unless either $m_s$ or $m_b$ is an opt-out message). Thus, if the buyer's mean

[14]In fact, the agent is also concerned with his transfer, but in equilibrium messages with the same expected probability of trade must have the same expected transfers.

type in $m_b$ is strictly larger than the seller's mean type in $m_s$, then $p(m_s, m_b) = 1$; if it is strictly smaller, then $p(m_s, m_b) = 0$. However, credibility does not pin down the probability of trade when the means of $m_s$ and $m_b$ are equal, because the interim surplus from trade is zero and, therefore, any value of $p(m_s, m_b)$ is consistent with interim-surplus maximization.

We now define a stronger restriction—credible minimality—which requires that $p$ be zero when the interim surplus is zero. In other words, a credible-minimal allocation rule $p$ dictates the least possible trade subject to the credibility constraint.

Definition 3 (Credible minimality). An allocation rule $p$ is credible-minimal with respect to message set $M$ if

$$p(m_s, m_b) = \begin{cases} 1 & \text{if } \mathbb{E}[v_b | v_b \in m_b] > \mathbb{E}[v_s | v_s \in m_s], \\ 0 & \text{if } \mathbb{E}[v_b | v_b \in m_b] \leq \mathbb{E}[v_s | v_s \in m_s], \end{cases}$$

for every $(m_s, m_b) \in (M_s, M_b)$ except for (opt-out) messages $m_i$ satisfying $p(m_i, m_{-i}) = 0$ for all $m_{-i}$.

Note that once the designer chooses the message set $M$, the credible-minimal allocation rule $p$ is pinned down by the expected types of the buyer and seller for each message pair. The designer has only one degree of freedom, such that he can pick a message $m_i$ and set $p(m_i, m_{-i}) = 0$ for all $m_{-i}$, thus making it an opt-out message. Note, however, that only the buyer's lowest message, or the seller's highest message, can serve as an opt-out message. This is because incentive compatibility requires the expected probability of trade to be monotone in an agent's type.

In what follows, we show that the allocation rule in an *optimal* intermediation mechanism must be credible-minimal. Otherwise, the mechanism would allow for zero-surplus trade that can be eliminated, thereby saving on information rents which can then be used to add beneficial trade while maintaining incentive-compatibility. While such a result is immediate in standard mechanism design, it is challenging in the case of intermediation mechanisms since adding beneficial trade is not straightforward, as can be seen in the intuition presented for Proposition 3 below. Thus, in order to show that nonbeneficial trade can be eliminated we need to establish some additional results.

When the allocation rule is credible-minimal, the number of buyer and seller messages in any interval in their type space is roughly the same.

Lemma 4. *Suppose that the allocation rule $p$ is credible-minimal with respect to message set $M$. Let $k_i$ be the number of agent $i$'s messages that are fully contained in some interval $\hat{V} \subset V_b \cap V_s$. If $k_i$ is finite, then $k_i + 3 \geq k_{-i} \geq k_i - 3$.*

Corollary. *If $v$ is an accumulation point of agent $i$'s messages, then it is also is an accumulation point of agent $-i$'s messages. If agent $i$'s types in some interval $\hat{V}$ are fully separated, then so are agent $-i$'s types in $\hat{V}$.*[15]

[15]We say that $v \in V_i$ is an accumulation point in $M_i$ if in any neighborhood of $v$ there are infinitely many messages. Types in an interval $\hat{V}$ are fully separated if each of them reveals himself, i.e., $\{v\} \in M_i$ for any $v \in \hat{V}$.

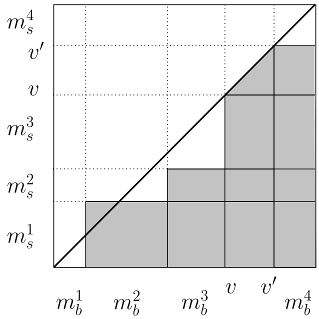


FIGURE 1. An allocation rule plotted in the agents' type space.

Figure 1 illustrates a credible-minimal allocation rule. The buyer types appear on the horizontal axis. They are partitioned into messages $M_b = \{m_b^1, m_b^2, m_b^3, m_b^4\} \cup [v, v']$ (i.e., four messages are intervals, and types within $[v, v']$ are each a singleton message). The seller types appear on the vertical axis (partitioned according to $M_s = \{m_s^1, m_s^2, m_s^3, m_s^4\} \cup [v, v']$). The diagonal is the *equi-type line*, i.e., the line along which the buyer and seller types are identical. For any message pair, trade occurs whenever the expected type of the buyer is higher than that of the seller (the gray area). Note that since intermediation mechanisms are minimal (see property (MIN) above), the trading rule is monotone, namely as we move to a higher buyer message, trade occurs up to a *strictly* higher seller message. Thus, when the messages are intervals, the trading area takes on a "step form." The example in Figure 1 also has an interval $(v, v')$ in which all the buyer and seller types are fully separated. Note that whenever that is the case, trade is ex post efficient.

### 4.2 *Ex ante budget balance*

An intermediation mechanism, according to our definition, is ex-post budget-balanced, namely $t_s(m_s, m_b) + t_b(m_s, m_b) = 0$ for every $m \in M$. As a step toward solving the optimization problem, we relax this requirement and replace it with an ex ante one. Thus, we will allow the designer to create a deficit following some reports by the agents, but require the expected deficit to be zero. Intermediation mechanism $\Gamma$ is said to be *ex ante budget-balanced* if

$$\mathbb{E}_{m_s, m_b}\big[t_s(m_s, m_b) + t_b(m_s, m_b)\big] = 0. \tag{1}$$

A well-known result in mechanism design (with full commitment) states that when types are independent the notions of ex post and ex ante budget balance are equivalent under interim IR and IC (see, e.g., Borgers and Norman (2009)). This equivalence holds also in the case of intermediation mechanisms, but requires an adjustment in the proof.

LEMMA 5. *For any* ex ante budget-balanced *partition-direct intermediation mechanism, there exists a payoff-equivalent* ex post budget-balanced *partition-direct intermediation mechanism with the same message sets and allocation rule.*

The key intuition of this result is the same as in the case of standard mechanisms. Since the agents are risk-neutral, they are willing to insure the mechanism designer at fair premiums. The designer can therefore find transfer rules $t_b$, $t_s$ that sum up to zero for every possible profile of reports (see Borgers and Norman (2009) for a detailed discussion). Since the translation from one mechanism to the other does not affect the message set or the allocation rule, credibility (and minimal credibility) is maintained. A modification of the proof is required to ensure that an opt-out message in the former remains so in the latter (i.e., if $m_i$ entails no transfers for agent $i$, then this remains the case in the new intermediation mechanism).

### 4.3 *Minimal budget and optimal intermediation mechanisms*

Given message set $M$ and a credible-minimal allocation rule $p$, let $\omega_s(v_b)$ denote the supremum of seller types that trade with buyer type $v_b$, and let $\omega_b(v_s)$ be the infimum of buyer types that trade with seller type $v_s$:[16]

$$\omega_s(v_b) = \sup\bigl(v_s : p\bigl(\sigma_s(v_s), \sigma_b(v_b)\bigr) = 1\bigr),$$
$$\omega_b(v_s) = \inf\bigl(v_b : p\bigl(\sigma_s(v_s), \sigma_b(v_b)\bigr) = 1\bigr).$$

Using $\omega_s$ and $\omega_b$, we define

$$\psi(p) = \int_{(v_b, v_s): p(\sigma_s(v_s), \sigma_b(v_b))=1} \bigl(\omega_s(v_b) - \omega_b(v_s)\bigr) \cdot dF(v_b) \cdot F(v_s). \tag{2}$$

Consider a standard mechanism design problem that involves implementing some credible-minimal trading rule $p$. The minimal net expected payment to the agents, required to sustain incentive compatibility and individual rationality, is $\psi(p)$ (we show this in the proof of the next proposition).[17] Thus, a budget-balanced trading mechanism exists if and only if $\psi(p) \leq 0$—after all, if the designer ends up with some money ($\psi < 0$) he can always give it to the agents as lump sums without affecting their incentives. In the case of intermediation mechanisms, lump sums might affect the reluctant type's decision to opt out and, therefore, $\psi < 0$ does not guarantee existence (see the remark following Lemma 6). Thus, a weaker result holds.

Lemma 6. *Suppose that the allocation rule* $p$ *is credible-minimal with respect to* $M$.

(i) *If* $\psi(p) = 0$, *then there are transfer rules* $t_b, t_s : M \to \mathbb{R}$ *such that* $\Gamma = \langle M, p, t_b, t_s\rangle$ *is an intermediation mechanism.*

(ii) *If* $\psi(p) > 0$, *then there exists no* $t_b, t_s : M \to \mathbb{R}$ *such that* $\Gamma = \langle M, p, t_b, t_s\rangle$ *is an intermediation mechanism.*

[16] Note that $\omega_s(\cdot)$ and $\omega_b(\cdot)$ are equivalent to the transfer functions of a canonical trade mechanism, as defined by Borgers (2015).

[17] In fact, this assertion is true for any allocation rule $p(v_b, v_s)$ that takes values 0 or 1 and increases (decreases) in the type of the buyer (seller).

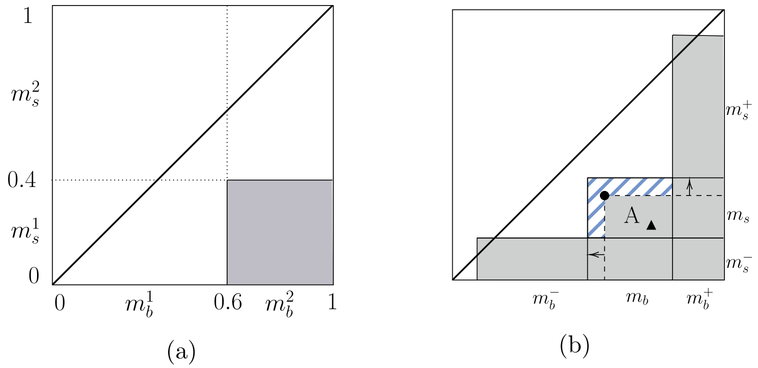


FIGURE 2. (a) An allocation rule with a negative $\psi$ that cannot be implemented by an intermediation mechanism. (b) A modification that increases the ex ante surplus and can be implemented by an intermediation mechanism.

REMARK. The lemma does not guarantee that an intermediation mechanism exists when $\psi(p) < 0$. To see this, consider the example depicted in Figure 2(a) with type distributions uniform over $[0, 1]$. Here, $\psi(p) = \int_{v_b \geq 0.6, v_s \leq 0.4}(0.4 - 0.6) < 0$. Note that the buyer's expected type is higher than the seller's for message pairs $(m_b^1, m_s^1)$ and $(m_b^2, m_s^2)$. Thus, the allocation rule can be credible only if $m_b^1$ and $m_s^2$ are opt-out messages, i.e., if transfers are 0 following any message pair other than $(m_b^2, m_s^1)$. But then incentive compatibility dictates $t_b(m_b^2, m_s^1) = -0.6$ and $t_s(m_b^2, m_s^1) = 0.4$, contradicting budget balance.

We conclude this section with two key properties of the allocation rule in an *optimal* intermediation mechanism.

PROPOSITION 3. *If* $\Gamma = \langle M, p, t_b, t_s \rangle$ *is an optimal intermediation mechanism, then:*

*(i) The allocation rule* $p$ *is credible-minimal with respect to* $M$*.*

*(ii)* $\psi(p) = 0$.

Given an intermediation mechanism $\Gamma$, the proof shows that if $\psi < 0$, or if $p$ allows for trade when the interim surplus is zero, then we can generate a temporary slack in the budget without affecting the surplus. On the other hand, since $\Gamma$ is budget-balanced, there must be type-pairs that do not trade even though trade would create a positive surplus. This follows from Myerson and Satterthwaite's (1983) impossibility theorem, according to which a trade mechanism cannot achieve ex post efficiency (let alone an intermediation mechanism that is bound also by the credibility constraint). Had the problem been a standard mechanism design one, the designer could directly change the allocation rule and add a *small* amount of beneficial trade while using the budget slack to cover the *small* amount of implied additional information rents.[18] This would be a contradiction to the optimality of the original mechanism. But, for an intermediation mechanism, trade can be added only by modifying the agents' message sets, while

[18]The only constraint would be to maintain the monotonicity of the expected trade probabilities.

the allocation rule is dictated by credibility. Moreover, when conducting this modification one has to make sure that the (small) change in messages does not result in large additions to the set of trading type-pairs and thereby to a large increase in the implied information rents.

We thus show that there must exist a message pair, in the frontier of the trading area, which is fully below the equi-type line, e.g., the message pair $(m_b, m_s)$ that corresponds to the rectangle A in Figure 2(b). Then, to add beneficial trade we slightly decrease the lower bound of $m_b$ and increase the upper bound of $m_s$. This modification increases the ex ante surplus without a large increase in the information rent. To see why, it is convenient to conduct the modification in two steps. In the first, the message bounds are changed but the allocation rule $p(m)$ is held fixed (perhaps violating credibility). As can be seen in the figure, this small modification only adds beneficial trade in the hatched area so that the ex ante surplus and the implied information rents slightly increase. In the second step, the allocation rule is readjusted to satisfy credibility. By definition, this can only further increase the surplus. Moreover, the implied information rents do not increase. This is because, for *any* message pair, the change in the bounds of $m_b$ and $m_s$ can only reduce the interim surplus from trade,[19] and thus the readjustment of the allocation rule can only eliminate (nonbeneficial) trade.[20]

## 5. The optimal intermediation mechanism

In this section, we apply the tools developed in the previous sections in order to prove the existence of an optimal intermediation mechanism and the finiteness of its message sets. We study the bounds on the surplus that intermediaries can achieve. We show that (under mild conditions) the intermediary is *strictly* less efficient than a mechanism with full commitment power, but does at least as well as a posted-price mechanism. For the case of uniform prior distributions, we solve for the optimal mechanism and show that imperfect commitment is extremely harmful, as the intermediary cannot do better than the optimal posted price. We conclude with an example showing that when the intermediary's resource constraint is relaxed (i.e., subsidizing trade is allowed) the consequences of limited commitment are less severe, and he can implement an outcome that is closer to that of the full-commitment case.

### 5.1 *Existence and finiteness of the optimal mechanism*

Since the type space in our model is continuous, one might have thought that the intermediary could benefit by separating the agents' types very finely. Indeed, it is possible to construct an intermediation mechanism with infinitely many messages and in that respect our setting differs from other limited commitment environments such as

[19]To see why, denote the message below $m_b$ by $m_b^-$. The decrease of the lower bound of $m_b$ moves the higher types of $m_b^-$ into $m_b$, thereby *decreasing* the mean type in both messages. A similar type of argument works for the seller.

[20]The proof handles the case in which large slacks in budget are generated by the modification, to guarantee that the resulting mechanism is an intermediation mechanism.

Crawford and Sobel (1982) in which every equilibrium is finite. However, separating the agents' types in some interval too finely implies an allocation rule that closely follows the equi-type line, thus allowing trade that generates very little surplus, not justifying the information rents that it necessitates. The next proposition thus asserts that the *optimal* intermediation mechanism never partitions the type spaces too finely.

Proposition 4. *For any type distribution $F = \{F_s, F_b\}$ an optimal intermediation mechanism exists. Moreover, there exists $\bar{K}_F$, such that in any optimal intermediation mechanism, each message set consists of fewer than $\bar{K}_F$ elements.*

The proof strategy is the following. We divide each agent's type space into many identical intervals (whose size depends solely on the distributions). We show that, given any intermediation mechanism, if an agent has more than some fixed number of messages in a small interval, then there exists an intermediation mechanism with fewer messages that generates a higher surplus. This, together with a simple compactness argument, proves the existence of an optimal intermediation mechanism and that the number of messages of each agent is below some finite $\bar{K}_F$.

To see the intuition why an agent cannot have many messages in some small interval, assume to the contrary that he does. By Lemma 4, the other agent also has many messages in the same small interval. Thus, there is a small box in the type space (the product of the two intervals) in which the differences in the agents' valuations are small while some type-pairs trade (see the proof of Lemma 4). While such trade creates only a negligible expected surplus, it requires the payment of significant information rents. At the same time, and since the mechanism is budget-balanced, there are type-pairs that do not trade even though trade would create a relatively large surplus (i.e., $v_b - v_s$ is relatively large). The idea is to modify the mechanism in a way that shifts trade from low to high surplus-to-rent-ratio type-pairs, reduces the number of messages and keeps the budget balanced. This modification, which is straightforward in standard mechanism design, needs to be carefully implemented in the case of intermediation mechanisms since the modified mechanism has to satisfy credibility.

To reduce low surplus-to-rent trade in the small box, we merge messages. To illustrate, consider the simple case depicted in Figure 3(a), in which the agents' types within the (very small) box $[v, v'] \times [v, v']$ are fully separated (i.e., each agent has a continuum of singleton messages). If the type-distributions are uniform, we merge all the (singleton) messages in $[v, v']$ into one message for each agent, i.e., $m_b = m_s = [v, v']$. Since the mean types in $m_b$ and $m_s$ are equal, we credibly set $p(m_s, m_b)$ to 0, thereby eliminating trade (see Figure 3(b)). With nonuniform distributions, the mean types are not exactly equal. However, the fact that the box is very small guarantees that the *conditional* type distributions are very close to uniform. Thus, a slight modification of the above construction allows us to eliminate *most* of the trade in the box in a credible way. This procedure for reducing trade, which is relatively simple when there is a continuum of messages, requires a more subtle argument when there is a countably infinite number of messages (see the proof).

The procedure for adding trade while maintaining credibility resembles the one described in the intuition of the proof of Proposition 3 above and illustrated in Figure 2(b).

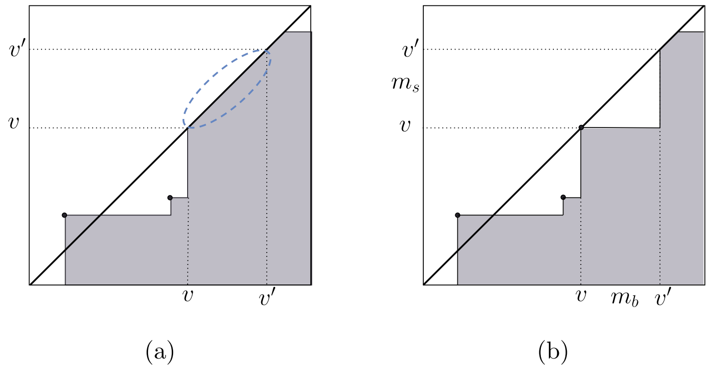


FIGURE 3. Merging messages to reduce low surplus-to-rent trade.

Extra care is taken to identify type-pairs whose surplus-to-rent ratios exceeds some lower bound that depends only on the type distributions.

### 5.2 *Infeasibility of the second-best outcome: An upper bound*

It is obvious that the social surplus attained by the optimal intermediation mechanism is bounded from above by that of the optimal (full-commitment) mechanism. But is this bound achievable? The next result, which is a direct implication of Proposition 4, asserts that for regular distributions the answer is negative, i.e., an intermediation mechanism does *strictly* worse than a full-commitment one.

COROLLARY (Upper bound). *Suppose that $F_s$ and $F_b$ are regular.*[21] *Then the ex ante social surplus attained by the intermediary is strictly smaller than that attained by the optimal trade mechanism.*

The proof follows immediately from Myerson and Satterthwaite's (1983) characterization of the (second-best) optimal mechanism. They show that, with regular type distributions, the boundary between the type pairs that trade and those that do not is strictly increasing.[22] But since implementing this allocation rule requires full separation of the agents' types (at least in some intervals), it cannot be achieved with finite message sets. Note also that since an optimal intermediation mechanism exists, it is not the case that a sequence of intermediation mechanisms can arbitrarily approach the surplus attained by the optimal trade mechanism.

### 5.3 *Feasibility of the optimal posted price: A lower bound*

Fix any price $x \in V_s \cap V_b$. We say that the intermediation mechanism $(\Gamma, \sigma)$ *implements the posted price $x$* if, in equilibrium, the agents trade whenever the seller's type is below

[21] $F_s$ and $F_b$ are regular if the virtual valuation of the buyer, $v_b - (1 - F_b(v_b))/f_b(v_b)$, is increasing in $v_b$, and the virtual valuation of the seller, $v_s + F_s(v_s)/f_s(v_s)$, is increasing in $v_s$.

[22] The boundary is the solution to $(v_s + \alpha(F_s(v_s)/f_s(v_s))) - (v_b - \alpha((1 - F_b(v_b))/f_b(v_b))) = 0$, where $\alpha > 0$ is the lowest value for which budget balance is satisfied.

$x$ and the buyer's type is above $x$, and the price that the buyer pays the seller in the case of trade is $x$. We then have the following.

Proposition 5. *For any price $x \in V_s \cap V_b$, there is an intermediation mechanism that implements the posted price $x$.*

The proof is immediate and goes by construction. Consider the message sets $M_s$ and $M_b$ where $M_i$ partitions agent $i$'s type space into two messages: $M_i = \{[\underline{v}_i, x], [x, \bar{v}_i]\}$. The allocation rule $p(m)$ imposes trade whenever the seller reports $[\underline{v}_s, x]$ and the buyer reports $[x, \bar{v}_b]$. The transfer rules are $t_s(m) = -t_b(m) = x$ if there is trade and 0 otherwise. Thus, $[x, \bar{v}_s]$ and $[\underline{v}_b, x]$ are opt-out messages, implying that $p$ satisfies credibility.

A posted price $x^*$ is said to be optimal if $x^*$ maximizes $\int_{\underline{v}_s}^{x} \int_{x}^{\bar{v}_b} (v_b - v_s)\, dF_b(v_b)\, dF_s(v_s)$ among all $x \in V_s \cap V_b$. We thus have the following.

Corollary (Lower bound). *The ex ante social surplus attained by the intermediary is weakly larger than that attained by the optimal posted price.*

### 5.4 *The case of uniform distributions*

We now characterize the optimal intermediation mechanism for the case of uniform distributions. Although the analysis of the uniform case relies on specific calculations, it is insightful as it gives a vivid illustration for the forces at work in the general case. In particular, it illustrates the challenge in constructing an intermediation mechanism, which arises from the need to satisfy credibility and budget balance at the same time.

With uniform distributions, the intermediary can do no better than the lower bound of a posted-price mechanism.[23]

Proposition 6. *If the type distributions are uniform, then the intermediary implements the optimal posted price $x^*$, where*

$$x^* = \begin{cases} \underline{v}_b & \text{if } \dfrac{\underline{v}_s + \bar{v}_b}{2} < \underline{v}_b, \\ \dfrac{\underline{v}_s + \bar{v}_b}{2} & \text{if } \underline{v}_b \leq \dfrac{\underline{v}_s + \bar{v}_b}{2} \leq \bar{v}_s, \\ \bar{v}_s & \text{if } \bar{v}_s < \dfrac{\underline{v}_s + \bar{v}_b}{2}. \end{cases}$$

*Moreover, if the distributions are close to uniform, then the intermediary implements an outcome that is close to that of the optimal posted price $x^*$.*[24]

[23]Note that intermediation mechanisms with more than two messages exist also with uniform type distributions, but they are not optimal. The following is an example of such an intermediation mechanism: The buyer's types are uniformly distributed over [0, 1] and the seller's over [0.44, 1.44]. Message sets are $M_b = \{m_b^1, m_b^2, m_b^3\} = \{[0, 0.9], [0.9, 1-\varepsilon], [1-\varepsilon, 1]\}$, and $M_s = \{m_s^1, m_s^2, m_s^3\} = \{[0.44, 0.44+\varepsilon], [0.44+\varepsilon, 0.54], [0.54, 1.44]\}$. The allocation rule is $p(m_b^k, m_s^l) = 1$ if $k \geq l$ and $p = 0$ otherwise. One can verify that it is credible for $\varepsilon \leq 0.02$. As $\varepsilon$ decreases from 0.02 to 0, $\psi(p)$ decreases from positive to negative; at $\varepsilon = 0.00367$, we have $\psi(p) = 0$, and hence this is an intermediation mechanism (and is clearly dominated by the posted price of 0.72).

[24]Formally, for any $\delta > 0$ there exists $\varepsilon > 0$ such that, if the maximal pointwise difference between the probability density $f_i$ and that of the uniform distribution is less then $\varepsilon$ for both $i \in \{b, s\}$, then in the optimal

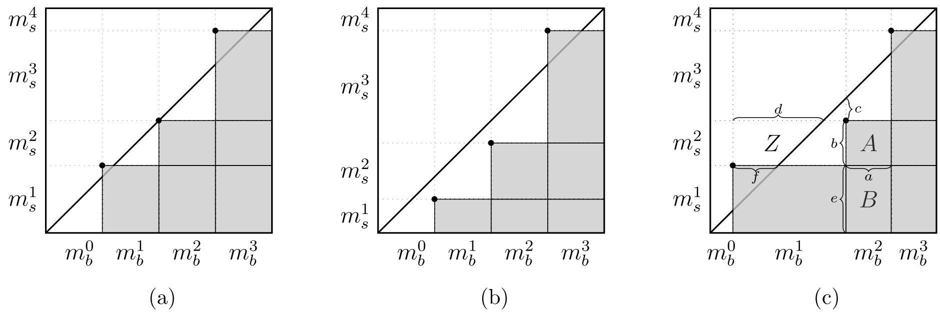


FIGURE 4. With uniform type distributions, the illustrated allocation rules cannot be part of a (budget-balanced) intermediation mechanism.

We sketch the intuition of the core argument using an example in which the (uniform) type distributions have identical supports. Assume to the contrary that the intermediary can do better than the optimal posted price. Then it must be that the message set of at least one agent contains more than two intervals. Consider then the three representative examples depicted in Figure 4.[25] We will show that there is no ex ante budget-balanced intermediation mechanism with an allocation rule that corresponds to them. In what follows, the term *extreme vertex* will refer to an outer corner of the frontier of trade, marked by small black dots in Figure 4.

Consider first the case in which $M$ and $p$ are such that all the extreme vertices are on or above the equi-type line, as depicted in Figure 4(a). In this case, the sign of $\omega_s(v_b) - \omega_b(v_s)$ is positive for every type pair $(v_b, v_s)$ that trades. Therefore, $\psi(p) > 0$ and, by Lemma 6, $M$ and $p$ cannot be part of a (budget-balanced) intermediation mechanism.

Next, consider the case in which there are two consecutive extreme vertices below the equi-type line, as depicted in Figure 4(b).[26] In this case, the fact that $p(m_s^2, m_b^1) = 0$ is inconsistent with credibility. This is because, since the types are uniformly distributed, the mean buyer type in $m_b^1$ is strictly larger than the mean seller type in $m_s^2$.

Finally, consider the case in which there is an extreme vertex above the equi-type line followed by an extreme vertex below it, as depicted in Figure 4(c). Note that the value of $\omega_s(v_b) - \omega_b(v_s)$ (which is constant within any rectangle with trade) is negative in $A$ and positive for all other rectangles with trade. Denote by $D_A$ (resp., $D_B$) the probability mass of rectangle $A$ (resp., $B$) multiplied by the value of $\omega_s(v_b) - \omega_b(v_s)$ in it. We will show that $D_A + D_B > 0$ and, therefore, $\psi(p) > 0$.

To see this, note that $D_A = (b/(\bar{v}_s - \underline{v}_s)) \cdot (a/(\bar{v}_b - \underline{v}_b)) \cdot (-c)$ whereas $D_B = d \cdot (e/(\bar{v}_s - \underline{v}_s)) \cdot (a/(\bar{v}_b - \underline{v}_b))$, where $a$, $b$, $c$, $d$, and $e$ are the lengths of the segments as marked in Figure 4(c). By credibility, the fact that there is no trade when the agents report $(m_s^2, m_b^1)$

intermediation mechanism, all types $[\underline{v}_i + \delta, x^* - \delta]$ are pooled into the same message $m_i^L$ and all the types $[x^* + \delta, \bar{v}_i - \delta]$ are pooled into the same message $m_i^H$, and the trading rule $p$ assigns trade for message pair $(m_b^H, m_s^L)$ and no trade for $(m_b^L, m_s^L)$, $(m_b^H, m_s^H)$, and $(m_b^L, m_s^H)$.

[25]There is an additional case not represented in these examples which requires the distributions to have nonidentical supports. We discuss this case separately in the proof.

[26]The same argument applies if one of the vertices is on the equi-type line and the other is strictly below it.

implies that the rectangle $Z$ has more mass above the equi-type line than below it and, therefore, it must be that $f \geq c$ and $d \geq b + c > b$. Since $e > f$, it follows that $D_B = d \cdot (e/(\bar{v}_s - \underline{v}_s)) \cdot (a/(\bar{v}_b - \underline{v}_b)) > c \cdot (b/(\bar{v}_s - \underline{v}_s)) \cdot (a/(\bar{v}_b - \underline{v}_b)) = -D_A$.

The above argument is generalized in the proof. In particular, we also show that, while with nonidentical supports intermediation mechanisms other than posted price may exist, they are dominated by the optimal posted-price one.

The second part of the proposition implies that the optimality of the posted price under uniform distributions is not a knife-edge result. Intuitively, the credibility and budget balance conditions are governed by weak inequalities, over functions that are continuous in the distribution and in the coordinates of the extreme vertices of the mechanism (which determine the bounds of the agents' messages). Thus, given a sequence of distributions that converge to the uniform, and a corresponding convergent sequence of optimal intermediation mechanisms (a mechanism for each distribution), the limit mechanism is an intermediation mechanism under the uniform distribution. Moreover, the surplus of this limit mechanism is the limit of the surpluses along the sequence (because the ex ante surplus is also continuous in the same variables) and it is equal to the surplus of the optimal posted price under the uniform distribution.[27] Finally, since the posted price mechanism is the unique optimal intermediation mechanism under the uniform distribution (as implied by the proof of the first part of the proposition), then the limit mechanism *is* the optimal posted price mechanism.

We conclude this section with the observation that it is not the general case that the intermediary cannot do better than a posted price. This is shown in the following example:

Example (Distributions for which intermediation outperforms the posted price). Suppose that the buyer and seller types are distributed on $[0.4, 1]$ and $[0, 0.6]$, respectively, according to the distribution functions $F_b(x) = 20t^3 - 42t^2 + \frac{439}{15}t - \frac{94}{15}$ and $F_s(x) = 20t^3 - 18t^2 + \frac{79}{15}t$. One can verify that the optimal posted price in this environment is 0.4 or 0.6. Now consider an intermediation mechanism with messages $[0.4, 0.7]$ and $[0.7, 1]$ for the buyer and messages $[0, 0.3]$ and $[0.3, 0.6]$ for the seller. When the buyer reports $[0.4, 0.7]$ and the seller reports $[0, 0.3]$, they trade at price 0.4; when they report $[0.7, 1]$ and $[0, 0.3]$, they trade at price 0.5; and when they report $[0.7, 1]$ and $[0.3, 0.6]$, they trade at price 0.6. For the reports $[0.4, 0.7]$ and $[0.3, 0.6]$, there is no trade or transfer. It is easy to verify that the mechanism satisfies individual rationality and incentive compatibility (note that each agent sends each of his two messages with equal probability, as $F_b(0.7) - F_b(0.4) = 0.5$ and $F_s(0.3) - F_s(0) = 0.5$). Credibility is satisfied because the buyer's type is always higher than the seller's for any message pair that trades. When there is no trade, the seller has a higher mean type: $\mathbb{E}[v_b|v_b \in [0.4, 0.7]] = 0.469 < 0.531 = \mathbb{E}[v_s|v_s \in [0.3, 0.6]]$. This intermediation mechanism generates a surplus of $\sim 0.42$ as

[27] It cannot be higher because, by the first part of the proposition, the surplus of the (optimal) posted price is the highest attainable surplus under the uniform distribution. It cannot be lower because, by Proposition 5, every mechanism along the sequence is at least as good as the optimal posted price under the corresponding distribution, and the surpluses of the optimal posted prices converge to the surplus of the optimal posted price under the uniform distribution.

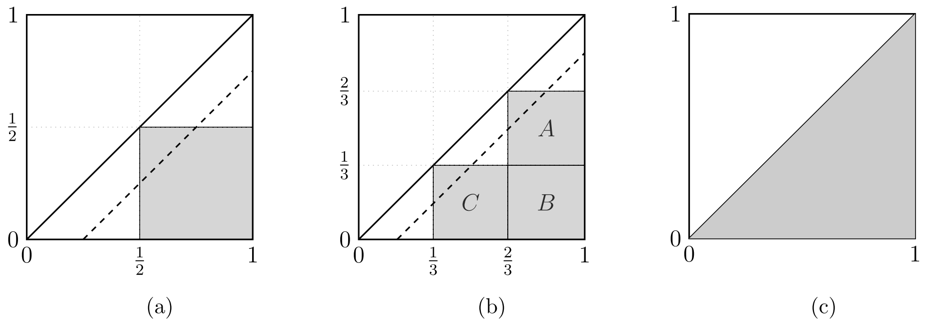


FIGURE 5. Intermediation mechanisms with a subsidy.

compared to $\sim 0.32$ for the optimal posted-price mechanism, which is therefore is not optimal. ◊

### 5.5 *Subsidizing intermediaries*

In this section, we expand the model to allow for an external subsidy. We compare (by means of an example) the benefit of the subsidy for intermediaries versus for full-commitment mechanisms. While the subsidy obviously increases the achievable surplus for both, the effect on the former is stronger. This is because the subsidy reduces the volume of beneficial trade that must be denied, allowing the intermediary to employ a finer language and thereby to better approach the (full-commitment) second-best allocation.

Assume that both agents' types are distributed uniformly over $[0, 1]$. Consider first the case of no subsidy, which is illustrated in Figure 5(a). The dashed line corresponds to the boundary of trade in the optimal full-commitment mechanism, in which trade takes place whenever the buyer's valuation is higher than the seller's by more than 0.25. The optimal intermediation mechanism implements the posted price $\frac{1}{2}$ so that trade takes place in the grey area. Calculations shows that the intermediation mechanism achieves 89% of the surplus generated by the full-commitment mechanism.[28]

Now consider the case in which there is an external (expected) subsidy of $\frac{1}{27}$, which corresponds to Figure 5(b). With full commitment, trade takes place whenever the valuation differ by more than $\sim 0.17$ (the dashed line).[29] It is not difficult to verify that there exists an intermediation mechanism with a three-message language—finer than that of

[28]The expected social surplus for the intermediation mechanism is $\int_{0.5}^{1}\int_{0}^{0.5}(v_b - v_s)\, dv_s\, dv_b = \frac{1}{8}$ and for the conventional mechanism is $\int_{0.25}^{1}\int_{0}^{v_b - 0.25}(v_b - v_s)\, dv_s\, dv_b = \frac{9}{64}$.

[29]From the equation in footnote 22, we know that trade takes place whenever $v_b - v_s > \beta$ for some $\beta \in (0, 1)$. We then solve $\int_{\beta}^{1}\int_{0}^{v_b-\beta}(v_b - v_s)\, dv_s\, dv_b + \frac{1}{27} = \int_{\beta}^{1}(\int_{\beta}^{v_b}(x-\beta)\, dx)\, dv_b + \int_{0}^{1-\beta}(\int_{v_s}^{1-\beta}(1-\beta-x)\, dx)\, dv_s$, where the left-hand side corresponds to the expected social surplus plus the subsidy and the right-hand side equals the expected information rents that are required to maintain incentive-compatibility. The solution is $\beta = \tilde{}0.17$.

TABLE 1. Subsidizing trade for intermediaries and mechanisms.

| | Surplus (×100) | | | Surplus Gain due to Subsidy (×100) | | Ratio of Surplus Gain to Subsidy | |
|---|---|---|---|---|---|---|---|
| Subsidy | Intermed. | Mech. | Ratio | Intermed. | Mech. | Intermed. | Mech. |
| 0 | 12.5 | 14.06 | 89% | | | | |
| 1/27 | 14.81 | 15.39 | 96% | 2.31 | 1.33 | 62% | 36% |
| 1/6 | 16.67 | 16.67 | 100% | 4.17 | 2.6 | 25% | 16% |

the no subsidy case (although note that this is not the optimal intermediation mechanism but rather just a feasible one).[30] The intermediation mechanism now achieves 96% of the surplus generated by the full-commitment mechanism.

Finally, consider a subsidy of $\frac{1}{6}$, which is sufficiently large that both the full-commitment mechanism and the intermediation mechanism can attain the first-best efficient outcome (Figure 5(c)). In this case, the ratio between the surpluses generated by the mechanisms is exactly 1.

Table 1 summarizes the surplus calculations for the three cases and compares the surplus gain due to a trade subsidy in an intermediation mechanism vs. a full-commitment mechanism. It shows that subsidizing an intermediary in more effective. A subsidy of $\frac{1}{27}$ increases the surplus of the optimal mechanism by $1.33 \times 10^{-2}$ (surplus gain to subsidy ratio of 36%) and that of the intermediation mechanism increases by $2.31 \times 10^{-2}$ (ratio of 62%). A subsidy of $\frac{1}{6}$ increases the surplus of the optimal mechanism by $2.6 \times 10^{-2}$ (surplus gain to subsidy ratio of 16%) and that of the intermediation mechanism by $4.17 \times 10^{-2}$ (ratio of 25%). This implies that if the subsidy has a social cost (such as a deadweight loss due to taxation), then there is a range in which subsidizing an intermediary is cost effective while subsidizing a conventional mechanism is not.

## 6. Concluding remarks

*Binding vs. nonbinding intermediary decisions* In our model, the intermediary's choice of outcome is binding. Alternatively, one could consider a nonbinding model in which each agent can reject the intermediary's decision and force the default outcome of no trade and payments.[31] While the analysis of that model is different, the main insights carry over.

[30] Transfers from the buyer to the seller that support the allocation rule depicted in Figure 5(b) such that the agents report truthfully are $\frac{5}{9}$ in rectangle $A$, $\frac{4}{9}$ in rectangle $B$, and $\frac{3}{9}$ in rectangle $C$. In addition, if there is trade, then the seller gets an additional amount of $\frac{1}{9}$ from the designer (which means that in expectation the designer subsidizes trade by $\frac{1}{27}$).

[31] In a different setting, Green and Laffont (1987) studied the notion of posterior implementability, which requires that each agent be content with the message he chose after learning the other agent's choice. In contrast, in the nonbinding variation of our model, having observed each other's messages, the agents can only change their action to opt-out and, moreover, in equilibrium some types actually do so.

As in our model, if the intermediary *knows* (given the reported messages) that the buyer's type is higher than the seller's, he recommends trade (and sets a price that satisfies both agents). Trade then occurs with probability 1. Likewise, if he knows that the buyer's type is lower, then there is no trade. The difference between binding and nonbinding cases is when the messages sent by the agents have a nonempty intersection. Then, instead of making a binding decision whether or not there will be trade, the intermediary only recommends trade at some price, knowing that types unsatisfied with that price will reject it. He thus picks the optimal posted price given his posterior beliefs, and trade occurs accordingly.

Thus, the main tension of our model remains: if the intermediary partitions the type spaces too finely, then trade occurs for type-pairs for which the buyer values the object only slightly more than the seller, thus violating the budget constraint. In particular, our result that the (full-commitment) second-best efficient outcome is infeasible continues to hold.

*Voluntary participation and credibility* In many "real world" institution-design problems, agents have the option to opt out and secure a default payoff. It is a useful practice in mechanism design to convert this voluntary participation property into an equivalent individual rationality constraint, i.e., restricting the set of permissible mechanisms to those in which, in equilibrium, the expected payoff to each type is at least the default. We also follow this practice when we transform the intermediation game into an intermediation mechanism in Section 2.

In our limited-commitment environment there is an additional "real world" feature—the intermediary is a player who decides the outcome only after the buyer and seller pick their messages. Our approach is to convert this game into a standard mechanism design problem by imposing the credibility constraint, thus further restricting the set of permissible mechanisms to those satisfying interim optimality.

However, imposing the two restrictions together requires some caution. In our case doing so requires weakening the credibility constraint so that opt-out messages are exempt from it. To see why, consider the following example in which the buyer and seller types are uniformly distributed over $[0.4, 1]$ and $[0, 0.6]$, respectively. In the intermediation game, there is an equilibrium that implements a posted price of 0.5. In this equilibrium, the intermediary allows each agent only one message in addition to his opt-out one, and if both agents opt in he sets a price of 0.5. All buyer types below 0.5 and all seller types above 0.5 then opt out; the other types opt in and accept the price of 0.5. Consider now the corresponding partition-direct intermediation mechanism, with messages $[0.4, 0.5]$ and $[0.5, 1]$ for the buyer and $[0, 0.5]$ and $[0.5, 0.6]$ for the seller. If the buyer chooses the message $[0.4, 0.5]$ and seller chooses the message $[0, 0.5]$, then trade is beneficial on average. A naive credibility constraint would require trade in this case, whereas in the game the fact that types $[0.4, 0.5]$ of the buyer opted out dictated no trade. Instead, our definition of credibility exempts opt out messages from interim optimality. The designer can then credibly set no trade and transfers for the buyer message $[0.4, 0.5]$ (whatever the seller message is), as this defines it to be an opt-out message. Consequently, the intermediation mechanism replicates the outcome of the game.

Thus, even though the credibility and individual rationality constraints are applied "at the same time," the order of moves in the game is respected.

*Communication devices* The literature on limited commitment mechanism design with a single informed agent has highlighted the potential benefit of using a communication device that can add noise to the agent's reports (see, e.g., Bester and Strausz (2007) and Doval and Skreta (2020)). This insight carries over to our multiagent framework: allowing the intermediary to commit ex ante to employ a hard-wired communication device, that garbles the message of each agent, can improve the outcome. However, in the multiagent setup, communication devices are, in fact, even more powerful. Since the communication device receives the messages of both agents (after all, if a communication device exists it is natural to assume that the same device can communicate with both agents), it can be used to generate signals that induce beliefs that are not independent across agents' types. For example, in the uniform $[0, 1] \times [0, 1]$ case analyzed above, the intermediary can implement the second-best outcome using a communication device that "mimics" the double auction (see Chatterjee and Samuelson (1983)) and recommends either no trade (with no further information) or trade (together with the appropriate price). It is then an equilibrium for the agents to follow their double-auction strategy and for the intermediary to follow the recommendation. However, that the second-best outcome is attainable is not a general result and, therefore, the role of communication devices with multiple informed agents remains an open research question.

## Appendix A: Proofs for Sections 2, 3, and 4

### *Proof of Proposition 1*

*Part I*: Suppose that some social choice function *scf* is implementable in the reporting subgame that starts with the message set $M$. Denote the agents' equilibrium strategies (in stage 2) by $\sigma' = (\sigma'_s, \sigma'_b)$ and the intermediary's strategy (in stage 3) by $(p', t')$, where $p'$ and $t'$ are functions that map each pair of reports to an outcome. We then have that:

(i) $(p', t')$ is optimal given $\sigma'$ and

(ii) $\sigma'_i$ is optimal given $\sigma'_{-i}$ and $(p', t')$ for each agent $i$.

Assume first that all the messages in $M = (M_b, M_s)$ are on-path, i.e., each message $m_i \in M_i$ is in the support of $\sigma_i(v_i)$ for some $v_i \in V_i$. Consider the trade mechanism $(\Gamma, \sigma')$ where $\Gamma = \langle M, p, t_b, t_s\rangle$ satisfies $p(m) = p'(m)$ and $-t_b(m) = t_s(m) = t'(m)$ whenever $(p', t')$ are defined (i.e., if neither $m_s$ nor $m_b$ in the report $m = (m_s, m_b)$ is the message labeled "out"), and $p(m) = t_b(m) = t_s(m) = 0$ otherwise. Note that $\sigma'$ is an equilibrium of the trade mechanism by (ii) above. Note also that $\Gamma$ satisfies credibility, since by (i) above we have that for any $m = (m_1, m_2)$, unless either $m_1$ or $m_2$ is an opt-out message (implying $p(m) = t_b(m) = t_s(m) = 0$), the allocation decision $p(m)$ maximizes $W_I(m) = E_{v_s,v_b}[v_b - v_s|m] \cdot \hat{p}(m)$ over all functions $\hat{p}(m)$, where the expectations are computed according to $\sigma'$. Thus, $\Gamma$ is an intermediation mechanism that implements *scf*. If the message sets $(M_b, M_s)$ of the game do contain messages that no type send in equilibrium, then in the construction of the intermediation mechanism we simply omit them and the argument remains intact.

Conversely, suppose that *scf* is implemented by an intermediation mechanism $(\langle M, p, t_b, t_s\rangle, \sigma')$. Label every opt-out message in the mechanism as "out". Then, having the agents play $\sigma'_1, \sigma'_2$ in stage 2 and the intermediary play $(p, t_s)$ in stage 3 is an equilibrium of the subgame that starts with $M$: For any $m$, the intermediary's strategy is optimal since $\Gamma$ satisfies credibility and $\sigma'_1, \sigma'_2$ constitute an equilibrium in stage 2. Thus, *scf* is implemented by the reporting subgame.

*Part II*: By part I, the set of social choice functions implementable by all intermediation mechanisms is the same as the set of social choice functions implementable by all possible reporting subgames. Note that: (1) the optimal intermediation mechanism solves $\max_{(\langle M,p,t\rangle,\sigma)} W_{\mathrm{EA}}$, where $(\langle M, p, -t, t\rangle, \sigma)$ is any intermediation mechanism; and (2) by the refinement that the intermediary-optimal equilibrium is played in any subgame, the outcome of the entire game is the one that solves $\max_M \max_{(\sigma,p,t)_M} W_{\mathrm{EA}}$, where $(\sigma, p, t)_M$ is any equilibrium in the reporting subgame starting with $M$. Clearly, both the direct maximization and the two-step maximization yield the same *scf*.

## *Proof of Lemma 1*

Suppose that $(\Gamma, \sigma)$ is an intermediation mechanism where $\Gamma = \langle M, p, t_b, t_s\rangle$. Suppose further that $m_s, m'_s \in M_s$ are two of the seller's messages such that $\mathbb{E}_{v_s}[v_s|m_s] \neq \mathbb{E}_{v_s}[v_s|m'_s]$ and $p(m_s, m_b) > p(m'_s, m_b)$ for some $m_b \in M_b$. If $m'_s$ is an opt-out message, then $p(m'_s, m'_b) = 0$ for all $m'_b \in M_b$ and the proof is complete. Otherwise, credibility implies that $\mathbb{E}_{v_s,v_b}[(v_b - v_s)|(m_s, m_b)] > \mathbb{E}_{v_s,v_b}[(v_b - v_s)|(m'_s, m_b)]$ and since types are independent then $\mathbb{E}_{v_s}[v_s|m_s] < \mathbb{E}_{v_s}[v_s|m'_s]$. Thus, $\mathbb{E}_{v_s,v_b}[(v_b - v_s)|(m_s, m'_b)] > \mathbb{E}_{v_s,v_b}[(v_b - v_s)|(m'_s, m'_b)]$ for any $m'_b \neq m_b$ implying $p(m_s, m'_b) \geq p(m'_s, m'_b)$. The proof for the buyer's side is similar.

## *Proof of Lemma 2*

Suppose that $(\Gamma, \sigma)$ is a nonminimal intermediation mechanism, where $\Gamma = \langle M, p, t_b, t_s\rangle$. Split the messages of each agent $i$ into equivalence classes, such that all the messages in the same equivalence class $\tilde{M}_i \subseteq M_i$ have the same expected probability of trade $\tilde{p}$ (i.e., $\bar{p}_i(m_i) = \tilde{p} \in [0, 1]$ for all $m_i \in \tilde{M}_i$). Nonminimality of $(\Gamma, \sigma)$ implies that (at least) one such equivalence class has more than one message.

We now construct a new intermediation mechanism $(\hat{\Gamma}, \hat{\sigma})$ in which all the messages in each equivalence class with more than one message are merged, and all types of both agents expect the same payoff in both mechanisms. The resulting mechanism is then a minimal one. To avoid confusion in evaluating conditional expectations, we add a superscript to the expected value operator to indicate the equilibrium ($\sigma$ or $\hat{\sigma}$) according to which expectations are evaluated (e.g., $\mathbb{E}^{\sigma}_{v_i}[v_i|m_i]$ is the mean type of agent $i$, conditional on message $m_i$ being sent by agent $i$ in the equilibrium $\sigma$).

Suppose that the equivalence class $\tilde{M}_i$ contains more than one message. Suppose further that each message in $\tilde{M}$ is sent by at least one type of agent $i$ in equilibrium (i.e., $\{v_i \mid m_i \in \operatorname{supp}[\sigma_i(v_i)]\}$ is nonempty for all $m_i \in \tilde{M}_i$), since otherwise we just drop the messages that are not sent by any type. We begin by making two useful observations:

(i) Either agent $i$'s mean type is the same for all messages $m_i \in \tilde{M}_i$ or, for any $m_{-i} \in M_{-i}$, the probability of trade $p(m_i, m_{-i})$ is the same for all $m_i \in \tilde{M}$. (This is because if there are two messages $m_i', m_i'' \in \tilde{M}_i$ with different agent $i$ mean types, and since $\bar{p}_i(m_i') = \bar{p}_i(m_i'')$, then $p(m_i', m_{-i}) = p(m_i'', m_{-i}) \equiv q(m_{-i})$ for all $m_{-i}$ by message-monotonicity of $\Gamma$. But then message monotonicity also implies $p(m_i, m_{-i}) = q(m_{-i})$ for all $m_{-i}$ for any $m_i$ with a mean type different than that of $m_i'$ or $m_i''$, i.e., for all $m_i$.)

(ii) Denote by $\bar{t}_i(m_i) = \mathbb{E}^{\sigma}_{m_{-i}} t(m_i, m_{-i})$ the expected monetary transfer to agent $i$ when he sends the message $m_i$. Since all messages $m_i \in \tilde{M}_i$ are being used in equilibrium, and since $\bar{p}_i(m_i) = \tilde{p}$ for all $m_i \in \tilde{M}_i$, then it must be the case that $\bar{t}_i(m_i) = \tilde{t}$ for all $m_i \in \tilde{M}_i$ for some $\tilde{t} \in \mathbb{R}$.

Consider a new mechanism, denoted as $(\hat{\Gamma}, \hat{\sigma})$, where $\hat{\Gamma} = \langle \hat{M}, \hat{p}, \hat{t}_b, \hat{t}_s \rangle$, which is identical to $(\Gamma, \sigma)$ (the "original mechanism") with the following modifications that are performed for each equivalence class $\tilde{M}_i$ with more than one message:

(i) All the messages in $\tilde{M}_i$ in the original mechanism are replaced by a single message $\hat{m}_i$ in the new one.

(ii) When agent $i$ sends the message $\hat{m}_i$ in the new mechanism $\hat{\Gamma}$, the monetary transfers to the agents and the probability of trade are set to be equal to their expected values conditional on messages in $\tilde{M}_i$ being sent in the original mechanism:

$$\hat{p}(\hat{m}_i, m_{-i}) = \mathbb{E}^{\sigma}_{m_i \in \tilde{M}_i} p(m_i, m_{-i}),$$

$$\hat{t}_j(\hat{m}_j, m_{-j}) = \mathbb{E}^{\sigma}_{m_i \in \tilde{M}_i} t_j(m_j, m_{-j}), \quad j = b, s,$$

for all $m_{-i} \in M_{-i}$, where $\mathbb{E}^{\sigma}_{m_i \in \tilde{M}_i}$ is evaluated according to the conditional distribution over $\tilde{M}_i$ in the original equilibrium $\sigma$.

(iii) All types who sent a message in $\tilde{M}_i$ under $\sigma_i$ in the original mechanism send $\hat{m}_i$ under $\hat{\sigma}_i$ in the new mechanism.

Given that agent $i$ plays according to $\hat{\sigma}_i$ in the new mechanism, it is a best response for agent $-i$ to play according to $\hat{\sigma}_{-i}$ (which is identical to $\sigma_{-i}$). This is because the monetary transfer and the probability of trade that agent $-i$ expects following every message $m_{-i} \in M_{-i}$ are by construction the same in both mechanisms. Similarly, given that agent $-i$ plays according to $\hat{\sigma}_{-i}$, it is a best response for agent $i$ to play according to $\hat{\sigma}_i$. This is because, if message $\hat{m}_i$ was merged from messages in an equivalence class $\tilde{M}_i$ for which the expected probability of trade is $\tilde{p}$, then $\overline{\hat{p}}(\hat{m}_i) = \tilde{p}$ and $\overline{\hat{t}}(\hat{m}_i) = \tilde{t}$. Thus, the expected probability of trade and expected payment are the same for any message $m_i \in \tilde{M}_i$ that agent $i$ sends in $\sigma$ and the merged message $\hat{m}_i$ he sends in $\hat{\sigma}$. Therefore, $\hat{\sigma} = (\hat{\sigma}_1, \hat{\sigma}_2)$ is an equilibrium in $\hat{\Gamma}$. Since each type of each agent expects the same probability of trade and the same monetary transfer in the new and the original equilibria, then $(\Gamma, \sigma)$ and $(\hat{\Gamma}, \hat{\sigma})$ are payoff-equivalent: $\bar{u}_i^{(\Gamma,\sigma)}(v_i, m_i) = \bar{u}_i^{(\hat{\Gamma},\hat{\sigma})}(v_i, \hat{m}_i)$ for every $v_i \in V_i$, $m_i \in \text{supp}[\sigma_i(v_i)]$, $\hat{m}_i \in \text{supp}[\hat{\sigma}_i(v_i)]$ and agent $i$.

Finally, it remains to verify that the new mechanism is credible. Note first that for message pairs $(\hat{m}_b, \hat{m}_s) \in \hat{M}_b \times \hat{M}_s$, in which neither $\hat{m}_b$ nor $\hat{m}_s$ is a merged message, the expected gains from trade (conditional on $(\hat{m}_b, \hat{m}_s)$ being reported) and the allocation decision are identical in the original and the new mechanisms. Next, fix a message $\hat{m}_i$ that was merged from an equivalence class $\tilde{M}_i$ in which all messages have the same expected trade probability $\tilde{p}$. We will show that $\hat{p}(\hat{m}_i, m_{-i})$ is consistent with credibility for all $m_{-i}$. If $\tilde{p} = 0$ and $\tilde{t} = 0$, then $\overline{\hat{p}}(\hat{m}_i) = 0$ and $\overline{\hat{t}}(\hat{m}_i) = 0$ and, therefore, $\hat{m}_i$ is an opt-out message.[32] Otherwise, either $\tilde{p} > 0$ or $\tilde{t} \neq 0$ and therefore the messages $m_i \in \tilde{M}_i$ are not opt-out messages. It then suffices to verify that $\hat{p}(\hat{m}_i, m_{-i})$ equals 1 (0) when $\mathbb{E}^{\hat{\sigma}}_{v_s,v_b}(v_b - v_s|\hat{m}_i, m_{-i})$ is positive (negative).

For any $m_{-i}$, denote $\underline{a} \equiv \inf_{m_i \in \tilde{M}_i} \mathbb{E}^{\sigma}_{v_s,v_b}[v_b - v_s|(m_i, m_{-i})]$ and $\bar{a} \equiv \sup_{m_i \in \tilde{M}_i} \mathbb{E}^{\sigma}_{v_s,v_b}[v_b - v_s|(m_i, m_{-i})]$. Note that $\mathbb{E}^{\hat{\sigma}}_{v_s,v_b}(v_b - v_s|\hat{m}_i, m_{-i}) \in [\underline{a}, \bar{a}]$, since all the types who sent messages in $\tilde{M}_i$ in the original mechanism $\Gamma$ send $\hat{m}_i$ in the new mechanism $\hat{\Gamma}$.

If the mean type of $i$ is the same for all $m_i \in \tilde{M}_i$, then $\underline{a} = \bar{a}$. If $\underline{a} = \bar{a} > 0$, then by the credibility of $\Gamma$ we have that $p(m_i, m_{-i}) = 1$ for all $m_i \in \tilde{M}_i$ and, therefore, $\hat{p}(\hat{m}_i, m_{-i}) = 1$, consistent with the credibility of $\hat{\Gamma}$. Similarly, if $\underline{a} = \bar{a} < 0$, then $p(m_i, m_{-i}) = 0$ for all $m_i \in \tilde{M}_i$ and, therefore, $\hat{p}(\hat{m}_i, m_{-i}) = 0$ as required. Finally, if $\underline{a} = \bar{a} = 0$, then any value of $\hat{p}(\hat{m}_i, m_{-i})$ in consistent with credibility.

If $i$'s mean types are not all the same, then $\bar{a} > \underline{a}$. By observation 1 above, for any $m_{-i} \in M_{-i}$ the probability of trade $p(m_i, m_{-i})$ is the same for all $m_i \in \tilde{M}$, and thus equal to $\hat{p}(\hat{m}_i, m_{-i})$. Note that it cannot be the case that $\bar{a} > 0 > \underline{a}$ since then there would have been two messages in $\tilde{M}_i$, say $m_i'$ and $m_i''$, for which $p(m_i', m_{-i}) = 1$ and $p(m_i'', m_{-i}) = 0$. Thus, either $\bar{a} > \underline{a} \geq 0$, in which case $p(m_i, m_{-i}) = 1$ for all $m_i \in \tilde{M}_i$ implying that $\hat{p}(\hat{m}_i, m_{-i}) = 1$, which is consistent with credibility, or $\underline{a} < \bar{a} \leq 0$ in which case $p(m_i, m_{-i}) = 0$ for all $m_i \in \tilde{M}_i$ implying that $\hat{p}(\hat{m}_i, m_{-i}) = 0$, which is again consistent with credibility.

## *Proof of Lemma 3*

Given a minimal intermediation mechanism $(\Gamma, \sigma)$, with $\Gamma = (M, p, t_s, t_b)$, we will first show that for each agent, the set of types who randomize in equilibrium can be partitioned into two disjoint subsets, each of measure zero, thus proving that almost all types of both agents do not randomize.

We say that message $m_i \in M_i$ is a *revealing message* in $(\Gamma, \sigma)$ if $m_i$ is sent by exactly one type of agent $i$ (and, therefore, it reveals the type). Then, for each agent $i$ we partition the set of types who randomize in equilibrium into two disjoint subsets: (i) the subset

[32] The fact that $\overline{\hat{p}}(\hat{m}_i) = \tilde{p} = 0$ implies that $\hat{p}(\hat{m}_i, m_{-i}) = 0$ for all $m_{-i} \in M_{-i}$. The fact that $\tilde{t} = 0$ implies that there is no loss in assuming that $\hat{t}(\hat{m}_i, m_{-i}) = 0$ for all $m_{-i} \in M_{-i}$. If this is not the case, the fact that $\tilde{t} = 0$ implies that it is possible to modify the transfer rules so that $\hat{t}(\hat{m}_i, m_{-i}) = 0$ for all $m_{-i} \in M_{-i}$ without affecting the interim utilities of both agents (e.g., by increasing $\hat{t}(\hat{m}_i', m_{-i})$ by $(F(\hat{m}_i)/(1 - F(\hat{m}_i))) \cdot \hat{t}(\hat{m}_i, m_{-i})$ for all $\hat{m}_i' \neq \hat{m}_i$, where $F(\hat{m}_i)$ is the probability measure of all types who send $\hat{m}_i$).

$V_i^1 \subseteq V_i$ that contains all types who randomize only between revealing messages, and (ii) the subset $V_i^2 \subseteq V_i$ that contains all types who randomize such that at least one message in their strategy's support is nonrevealing.

Consider first the set $V_i^1$. Suppose that type $v_i \in V_i^1$ randomizes between the messages $m_i$ and $m_i'$, and assume without loss that $\bar{p}_i(m_i) < \bar{p}_i(m_i')$ (recall that $\bar{p}_i(m_i) \neq \bar{p}_i(m_i')$ because the mechanism is minimal). By assumption, there is no other type in $V_i^1$ who sends the messages $m_i$ and $m_i'$. Moreover, there is no other type in $V_i^1$ who sends any other message that induces an expected trade probability between $\bar{p}_i(m_i)$ and $\bar{p}_i(m_i')$. This is because, by the single crossing property of the agents' preferences, any type $v_i'$ that is larger (smaller) than $v_i$ strictly prefers sending the message $m_i'$ $(m_i)$ over any message that induces an expected trade probability between $\bar{p}_i(m_i)$ and $\bar{p}_i(m_i')$. Thus, each type in $V_i^1$ can be associated with an interval $[\bar{p}_i(m_i), \bar{p}_i(m_i')]$ where the intervals are disjoint across types. Since $\bar{p}(\cdot)$ is bounded, the subset of types $V_i^1$ must be of zero measure.

Consider now the subset $V_i^2$. By assumption, for any type $v_i \in V_i^2$ there exists another type $v_i'$ (which may or may not randomize), such that both $v_i$ and $v_i'$ send some message $m_i \in M_i$ with positive probability in equilibrium. Assume without loss that $v_i' > v_i$. Since $m_i$ is in the best-response set of both $v_i$ and $v_i'$ then $m_i$ must be the *unique* optimal message for all types in $(v_i, v_i')$. This is again a consequence of the single crossing property of the preferences: if some type in $(v, v')$ finds a message $m_i' \neq m_i$ to be optimal for him, then since $\bar{p}_i(m_i) \neq \bar{p}_i(m_i')$ it must be the case that $m_i'$ is strictly better than $m_i$ for either $v_i$ or $v_i'$. Therefore, every nonrevealing message that is sent by some type in $V_i^2$ can be associated with an interval of types who send only this message in equilibrium. Since $V_i$ is bounded, and since the intervals are disjoint across messages, there can be only countably many nonrevealing messages that are sent by types in $V_i^2$. Since each of these messages is sent by at most two types in $V_i^2$ (because all the types between them do not randomize), then $V_i^2$ is a set of zero measure.

To conclude, suppose that in some minimal intermediation mechanism there is a measure zero of types who do randomize. We can then modify the strategies of the randomizing types to send one of the messages in the support of their original strategy. Obviously, this change has no effect on the expected payoff of each type of each agent in equilibrium and on the expected social surplus. In the modified mechanism, all types employ pure strategies.

### *Proof of Proposition 2*

By Lemmata 2 and 3, for any intermediation mechanism there exists a payoff-equivalent minimal intermediation mechanism in which both agents employ pure strategies. Given such a mechanism, we partition each agent's set of types according to the messages they send in equilibrium and rename each message to be the set of types that send it. Thus, each agent's messages partition his type space.

We now show that every message (i.e., subset of types) of each agent is convex. To see this, suppose that two buyer types, $v_b^l$ and $v_b^h$ where $v_b^l < v_b^h$, send the message $m_b \in M_b$. Suppose further, by way of contradiction, that some type $v_b' \in (v_b^l, v_b^h)$ sends the message

$m'_b \neq m_b$, so that the set of types that send $m_b$ is not convex. Then it must be the case that $\bar{p}_b(m'_b) \cdot v'_b + \bar{t}_b(m'_b) \geq \bar{p}_b(m_b) \cdot v'_b + \bar{t}_b(m_b)$ where $\bar{t}_b(m_b) = \mathbb{E}_{m_s} t(m_s, m_b)$. Recall that the mechanism is minimal and, therefore, either $\bar{p}_b(m_b) < \bar{p}_b(m'_b)$ or $\bar{p}_b(m_b) > \bar{p}_b(m'_b)$. In the first case, $\bar{p}_b(m'_b) \cdot v^h_b + \bar{t}_b(m'_b) > \bar{p}_b(m_b) \cdot v^h_b + \bar{t}_b(m_b)$, contradicting the optimality of $m_b$ for type $v_b$. In the latter case, $\bar{p}_b(m'_b) \cdot v^l_b + \bar{t}_b(m'_b) > \bar{p}_b(m_b) \cdot v^l_b + \bar{t}_b(m_b)$, contradicting the optimality of $m_b$ for type $v_l$. The proof for the seller types is analogous.

## *Proof of Lemma 4*

Suppose the buyer has $k_b$ messages contained in the interval $\hat{V}$. Then he has $\hat{k}_b \leq k_b + 2$ (consecutive) messages intersecting the interval. Denote them by $[v^0, v^1], [v^1, v^2] \cdots [v^{\hat{k}_b - 1}, v^{\hat{k}_b}]$. Since $p$ is credible-minimal, then for any seller message $m_s \subset \hat{V}$ there is some cutoff $v^j \in \{v^0, \ldots, v^{\hat{k}_b}\}$ such that $p(m_b, m_s) = 0$ for all buyer messages below $v^j$ and $p(m_b, m_s) = 1$ above $v^j$. Now, since minimality of the mechanism implies that different seller messages must differ in their trade decision for at least one buyer message, then there are at most $\hat{k}_b + 1$ seller messages contained in $\hat{V}$. Thus, $k_s \leq k_b + 3$. A parallel argument shows that $k_b \leq k_s + 3$.

## *Proof of Lemma 5*

Suppose that $\Gamma = \langle M, p, t_s, t_b \rangle$ is a partition-direct intermediation mechanism that is ex ante budget-balanced. We construct a partition-direct intermediation mechanism $\Gamma' = \langle M, p, t'_b, t'_s \rangle$ (with the same message set and the same allocation rule as in $\Gamma$) that is ex post budget balanced in two steps: First, we define two transfer rules $t'_b(m_s, m_b)$ and $t'_s(m_s, m_b)$ such that the expected payment for each agent $i \in \{s, b\}$ under $t_i$ (the original payment rule) and under $t'_i$ (the new payment rule) are the same for every message $m_i$. We then adjust the transfer rules $t'_i$ to ensure that if some message $m'_i$ is an opt-out message under $\Gamma$ (i.e., $p(m_i, m_{-i}) = t_i(m_i, m_{-i}) = 0$ for all $m_{-i}$), it would also be an opt-out message under $\Gamma'$ ($p(m_i, m_{-i}) = t'_i(m_i, m_{-i}) = 0$ for all $m_{-i}$).

We begin by defining the transfer rules $t'_b(m_s, m_b)$ and $t'_s(m_s, m_b)$ as follows:

$$t'_s(m_s, m_b) = \frac{1}{2} t_s(m_s, m_b) - \frac{1}{2} t_b(m_s, m_b) + \frac{1}{2}\big[\mathbb{E}_{m'_b}\big[d(m_s, m'_b)\big] - \mathbb{E}_{m'_s}\big[d(m'_s, m_b)\big]\big],$$

$$t'_b(m_s, m_b) = -t'_s(m_s, m_b),$$

where $d(m_s, m_b) = t_s(m_s, m_b) + t_b(m_s, m_b)$. Since $t'_b(m_s, m_b) = -t'_s(m_s, m_b)$ for any $(m_b, m_s) \in M$, then $\Gamma'$ is ex post budget-balanced. Recall that since $\Gamma$ is ex ante budget-balanced, then $\mathbb{E}_{m_s}\mathbb{E}_{m_b}[d(m_s, m_b)] = 0$. It is then easy to verify that, for each agent $i$, this change does not affect the expected monetary transfers for any message $m_i$, nor the expected payoffs for each of his types or their incentive to report truthfully.

Next, suppose that $m'_i$ is an opt-out message for agent $i$ in $\Gamma$, i.e., $p(m'_i, m_{-i}) = t_i(m'_i, m_{-i}) = 0$ for all $m_{-i}$ and, therefore, $\mathbb{E}_{m_{-i}} t_i(m'_i, m_{-i}) = \mathbb{E}_{m_{-i}} t'_i(m'_i, m_{-i}) = 0$. Note that it could be the case that $t'_i(m'_i, m_{-i})$, as defined above, is not zero for some $m_{-i}$, and hence $m'_i$ is not an opt-out message in $\Gamma'$, which may violate credibility. To correct this, we now slightly modify the transfer rule $t_i'$ as follows. Denote by $F_i(m_i')$ the probability measure of the interval (or singleton) of types that send the message $m_i'$. For any $m_i'' \neq m_i'$, increase $t_i'(m_i'', m_{-i})$ by $(F_i(m_i')/(1-F_i(m_i'))) \cdot t'(m_i', m_{-i})$ and set $t(m_i', m_{-i})$ to be zero, for all $m_{-i}$. This makes $m_i'$ an opt-out message in $\Gamma'$ while not changing the expected payment for any type of agent $-i$. Moreover, since $\mathbb{E}_{m_{-i}} t_i'(m_i', m_{-i}) = 0$, the expected monetary transfer to each type of agent $i$ remains unchanged. Thus, expected payoffs and the incentive to report truthfully remain unchanged. Since the message set, the allocation rule, and the set of opt-out messages are the same in $\Gamma$ and $\Gamma'$, then $\Gamma'$ is credible. We then have that $\Gamma' = \langle M, p, t_b', t_s' \rangle$ is an ex post budget-balanced partition-direct intermediation mechanism, as desired.

### *Proof of Lemma 6*

Suppose that $M = M_s \times M_b$ is a message set such that $M_i$ consists of intervals and singletons that partition $V_i$ for each agent $i$, and $p$ is an allocation rule that is credible-minimal with respect to $M$. We will first show that when $\psi(p) = 0$ we can use the functions $\omega_s(\cdot)$ and $\omega_b(\cdot)$ to define two transfer rules which—along with $M$ and $p$—constitute an ex ante budget-balanced intermediation mechanism (as defined in Section 4.2). Then, by Lemma 5, there exists $t_b' : M \to R$ and $t_s' : M \to R$ such that $\langle M, p, t_b', t_s' \rangle$ is an (ex post budget-balanced) intermediation mechanism.

First, note that if $v_i$ and $v_i'$ are two types of agent $i$ that send the same message $m_i$, i.e., $v_i \in m_i$ and $v_i' \in m_i$, then by definition $\omega_{-i}(v_i) = \omega_{-i}(v_i')$. We can therefore define $t_s(m_s, m_b) = \{\omega_s(v_b) : v_b \in m_b\}$ and $t_b(m_s, m_b) = \{-\omega_b(v_s) : v_s \in m_s\}$ if $p(m_s, m_b) = 1$, and $t_s(m_s, m_b) = t_b(m_s, m_b) = 0$ otherwise. Thus, whenever there is trade the monetary transfer to agent $i$ equals the unique value that $\omega_{-i}(v_i)$ attains for all types $v_i \in m_i$. When there is no trade the monetary transfer is zero.

The mechanism $\Gamma = \langle M, p, t_s, t_b \rangle$ (along with its truth-telling equilibrium) is an ex ante budget-balanced intermediation mechanism. To see this, note that the credibility of $\Gamma$ is satisfied because $p$ is credible-minimal with respect to $M$. Individual rationality is satisfied because, for every message $m_s \in M_s$, a buyer of type $v_b$ pays $\omega_b(v_s)$ *if the object is traded*, which is by definition, lower than $v_b$ (recall that $\omega_b(v_s)$ is the *lowest* buyer type that trades with seller type $v_s$). A similar argument applies to the seller. Incentive-compatibility is satisfied because the monetary transfer to agent $i$, conditional on the object being traded, is determined solely agent $-i$'s message and, therefore, deviating from truth-telling is not beneficial: by misreporting his type agent $i$ can either avoid trade at a price that is profitable to him or induce trade at some nonprofitable price. Finally, the mechanism's expected budget deficit is given by

$$\int_{(v_b, v_s): p(\sigma_s(v_s), \sigma_b(v_b))=1} t_s\big(\sigma_s(v_s), \sigma_b(v_s)\big) + t_b\big(\sigma_s(v_s), \sigma_b(v_s)\big) \cdot dF(v_b) \cdot F(v_s)$$
$$= \int_{(v_b, v_s): p(\sigma_s(v_s), \sigma_b(v_b))=1} \big(\omega_s(v_b) - \omega_b(v_s)\big) \cdot dF(v_b) \cdot F(v_s),$$

which is exactly $\psi(p)$. Thus, when $\psi(p) = 0$ the intermediation mechanism $\Gamma$ is ex ante budget-balanced and, therefore, by Lemma 5, there exists a payoff-equivalent counterpart.

Now consider the case in which $\psi(p) > 0$, which implies that the mechanism $\Gamma$ defined above creates a budget deficit. A well-known result in mechanism design (see, e.g., Krishna (2009)) is that, up to an additive constant, the expected payoff of each type of each agent in any incentive-compatible mechanism depends only on the allocation rule, and the constant is the expected utility of the type most reluctant to trade of that agent (namely, $\bar{u}_s(\bar{v}_s)$ for the seller and $\bar{u}_b(\underline{v}_b)$ for the buyer in our case). The mechanism $\Gamma = \langle M, p, t_s, t_b \rangle$ is incentive-compatible and provides the lowest possible expected payoff to the reluctant types in order to sustain individual rationality, that is, $\bar{u}_s(\bar{v}_s) = \bar{u}_b(\underline{v}_b) = 0$. Therefore, there is no other incentive-compatible mechanism with message set $M$ and allocation rule $p$ that pays (in expectation) less to the seller or collects (in expectation) more from the buyer. In other words, any other mechanism creates a (weakly) higher budget deficit. Thus, when $\psi(p) > 0$ there exists no (budget-balanced) intermediation mechanism with message set $M$ and allocation rule $p$.

## *Proof of Proposition 3*

The proof is by way of contradiction and consists of two parts. Throughout the proof, whenever the mechanism's message sets, or its allocation rule, are modified we assume that the monetary transfers $t_b$ and $t_s$ adjust immediately in order to sustain incentive compatibility, i.e.,

$$\bar{t}_b(m_b) = \int_{\underline{v}_b}^{\underline{m}_b} \bar{p}_b\big(\sigma_b\big(v'_b\big)\big)\, dv'_b - \bar{p}_b(m_b) \cdot \underline{m}_b, \tag{3}$$

$$\bar{t}_s(m_s) = \int_{\bar{m}_s}^{\bar{v}_s} \bar{p}_s\big(\sigma_s\big(v'_s\big)\big)\, dv'_s + \bar{p}_s(m_s) \cdot \bar{m}_s, \tag{4}$$

where $\underline{m}_b = \inf\{v_b | v_b \in m_b\}$ and $\bar{m}_s = \sup\{v_s | v_s \in m_s\}$. Importantly, we use the fact that adding (removing) small amounts of trade requires paying higher (lower) information rents to the agents.

*Part I.* Suppose that $\Gamma$ is an optimal intermediation mechanism in which the allocation rule $p$ is credible-minimal but $\psi(p) < 0$. Thus, for any type-pair $(v_b, v_s)$ the probability of trade $p(\sigma_b(v_b), \sigma_s(v_s))$ is either 0 or 1. Assume that the set of type pairs that trade in $\Gamma$ is closed, i.e., that if $\lim_{v'_b \to v_b, v'_s \to v_s} p(\sigma_b(v'_b), \sigma_s(v'_s)) = 1$ then $p(\sigma_b(v_b), \sigma_s(v_s)) = 1$. This assumption somewhat simplifies the exposition since it guarantees that the type pair $(v_i, v_{-i})$ trades whenever $v_i = \omega_i(v_{-i})$.[33]

Recall that $\psi(p)$ is the minimal expected transfer to the agents that is required to sustain truth-telling and individual rationality in the mechanism. Since $\Gamma$ is (by definition) budget balanced, then $\psi(p) < 0$ implies that there is at least one agent $i$ for whom all types expect a strictly positive utility from participating in the mechanism. We therefore slightly decrease the monetary transfer $t_i(m_b, m_s)$ for all $(m_b, m_s)$ by a small constant,

[33]If the set of type-pairs that trade in $\Gamma$ is not closed, we can slightly modify the message set and allocation rule to make it closed (potentially losing credible minimality temporarily). Since such a change affects only a measure zero of types (and hence does not affect the agents' incentives to report truthfully, the generated surplus, or whether the mechanism is credible), the rest of the argument remains valid.

thereby creating a small (temporary) slack in the budget. This modification does not affect the agents' incentives to report truthfully or the mechanism's generated surplus.

We now add beneficial trade to the mechanism. Since $\psi(p) < 0$, there is at least one type-pair $(v_b, v_s)$ that trades and for which $\omega_s(v_b) - \omega_b(v_s) < 0$ (for instance, the one that corresponds to the black triangle in Figure 2(b)). Consider then the type-pair $(v_b', v_s')$ where $v_b' = \omega_b(v_s)$ and $v_s' = \omega_s(v_b')$ (namely, the one that corresponds to the black dot in Figure 2(b)). It therefore must be that $v_s' - v_b' < 0$ (i.e., the black dot in Figure 2(b) is below the equi-type line). To see this, note that since $v_s' = \omega_s(v_b')$ is the highest seller type that trades with $v_b'$, then $v_s \leq v_s'$, and since $\omega_b$ is weakly increasing then $\omega_b(v_s) \leq \omega_b(v_s')$. On the other hand, since $\omega_b(v_s')$ is the lowest buyer type that trades with $v_s'$, and since the type-pair $(v_s', v_b')$ trades, then $\omega_b(v_s') \leq v_b' = \omega_b(v_s)$. Thus, $\omega_b(v_s) = \omega_b(v_s')$. Now, since $v_b' \leq v_b$ and since $\omega_s$ is increasing then $\omega_s(v_b') \leq \omega_s(v_b)$. Therefore, $\omega_s(v_b') - \omega_b(v_s') < \omega_s(v_b) - \omega_b(v_s) < 0$, or equivalently, $v_s' - v_b' < 0$.

Since $v_s' < v_b'$ then neither $v_s'$ nor $v_b'$ belongs to an interval of types that are fully separated. This is because, by the corollary Lemma 4, if all of agent $i$'s types over some interval are fully separated, then so are all types of agent $-i$ on the same interval. But then, by credibility, the highest seller type that trades with buyer type $v_b'$ must be equal to $v_b'$. Thus, it must be the case that $v_b'$ is a lower bound of some buyer interval message $m_b$ and $v_s'$ is the upper bound of some seller interval message $m_s$, and $p(m_b, m_s) = 1$. Note that the message pair $(m_b, m_s)$ is on the "frontier" of the trade area—no buyer message below $m_b$ trades with $m_s$, and no seller message above $m_s$ trades with $m_b$. Moreover, since $v_s' < v_b'$ then the lowest buyer type in $m_b$ is strictly higher than the highest seller type in $m_s$.

Consider now the following modification (illustrated in Figure 2(b)). We slightly decrease the lower bound of $m_b$ and slightly increase the upper bound of $m_s$. Since, by credible minimality, trade occurs only when the expected surplus is strictly positive, we make the changes sufficiently small that the expected buyer type remains higher than the expected seller type for the three message pairs $(m_b^-, m_s^-)$, $(m_b, m_s)$, and $(m_b^+, m_s^+)$, where $m_i^-$ and $m_i^+$ are the messages that come immediately below and above $m_i$ for any $i \in \{b, s\}$. This, in turn, implies that trade remains beneficial for all the (potentially infinite number of) other message pairs for which trade occurs. Since the lower bound of $m_b$ is higher than the upper bound of $m_s$, this modification adds beneficial trade to the mechanism. Importantly, the modification does not change the allocation rule for any message pair in the mechanism (this assertion is explained after the proposition in the body of the text). Moreover, we choose the size of the modification so that the additional information rents exactly offset the budget slack (recall that when we generated the slack it could be made arbitrarily small). Therefore, the modified mechanism is an intermediation mechanism that achieves a higher surplus than $\Gamma$, contradicting the optimality of $\Gamma$.

*Part II.* Suppose that the allocation rule $p$ is not credible-minimal with respect to $M$. Thus, there is at least one message pair $(m_b', m_s')$ for which $\mathbb{E}[v_b | v_b \in m_b'] = \mathbb{E}[v_s | v_s \in m_s']$ and $p(m_b', m_s') > 0$. Denote by $m_i^0$ and $m_i^1$ the first and second messages of agent $i$. We proceed according to the following cases:

*Case 1*: $p(m_b^0, m_s^0) = 0$ and $\mathbb{E}[v_b|v_b \in m_b^0] > \mathbb{E}[v_s|v_s \in m_s^0]$. In this case, we slightly reduce the probability of trade $p(m_b', m_s')$ and slightly decrease the boundary between $m_b^0$ and $m_b^1$. The first modification reduces the amount of zero-surplus trade in the mechanism, thereby creating a small (temporary) slack in the budget. The second modification increases the amount of trade and the required information rents. Note that the additional trade is between buyer types at the high end of (the original) $m_b^0$ and seller types in $m_s^0$. The fact that, in the original mechanism, $\mathbb{E}[v_b|v_b \in m_b^0] > \mathbb{E}[v_s|v_s \in m_s^0]$ guarantees that this additional trade is (on average) beneficial. We choose the size of the modifications so that the expected slack in budget is exactly offset by the increase in information rents. Thus, the modified mechanism is an intermediation mechanism that generates higher surplus than $\Gamma$, contradicting its optimality.

*Case 2*: $p(m_b^0, m_s^0) = 1$ or $\mathbb{E}[v_b|v_b \in m_b^0] \leq \mathbb{E}[v_s|v_s \in m_s^0]$. We set $p(m_b, m_s) = 0$ for *all* message pairs $(m_b, m_s)$ satisfying $\mathbb{E}[v_b|v_b \in m_b] = \mathbb{E}[v_s|v_s \in m_s]$. This modification eliminates zero-surplus trade and creates (temporary) budget slack. We restore budget balance by giving a lump sum to the buyer. Obviously, this does not change the agents' incentives to report truthfully nor the surplus that is generated by the mechanism. After the modification, trade occurs for a message pair $(m_b, m_s)$ if and only if $\mathbb{E}[v_b|v_b \in m_b] > \mathbb{E}[v_s|v_s \in m_s]$, implying that the allocation rule is credible minimal.[34] Thus, the resulting intermediation mechanism is credible minimal, it generates the same expected surplus as $\Gamma$ and $\psi(p) < 0$ (the fact that the buyer gets a positive lump sum means that $p$ can also be supported in equilibrium that satisfies IR in which the net expected transfer to the agents is negative, i.e., $\psi(p) < 0$). By part I of the proof, this new mechanism is not optimal, contradicting the optimality of $\Gamma$.

## Appendix B: Proofs for Section 5

### *Preliminaries*

Denote the lower bound of an interval message $m_i$ by $\underline{m}_i = \inf\{v_i|v_i \in m_i\}$ and its upper bound by $\bar{m}_i = \sup\{v_i|v_i \in m_i\}$.

Given an intermediation mechanism $\Gamma$, *extreme vertices* are the type pairs marked by black dots in Figure 4, and formally defined as follows.

Definition 4. Type-pair $(c_b, c_s)$ is an extreme vertex if, for any $(v_b, v_s)$, (i) $v_b \geq c_b$ and $v_s \leq c_s$ implies that $(v_b, v_s)$ trades in $\Gamma$; (ii) $v_b \leq c_b$ and $v_s \geq c_s$, with at least one of the inequalities being strict, implies that $(v_b, v_s)$ does not trade in $\Gamma$.[35]

The proof of Proposition 4 will show that in an optimal intermediation mechanism the number of extreme vertices is finite. This implies that the number of messages for each agent is finite.

[34] Note that if $m_b^0$ is an opt-out message in the original mechanism then $p(m_b^0, m_s) = 0$ for all seller messages $m_s \in M_s$. While in the modified mechanism $m_b^0$ is not an opt out message (because the buyer types in $m_b^0$ get a positive monetary transfer), credibility is still satisfied because the fact that $\mathbb{E}[v_b|v_b \in m_b^0] \leq \mathbb{E}[v_s|v_s \in m_s^0]$ implies that the mean buyer type in $m_b^0$ is smaller than the mean seller type in $m_s$ for all $m_s \in M_s$.

[35] For the purpose of the definition, if the set of type pairs who trade in $\Gamma$ is not closed we take its closure.

The allocation rule of an intermediation mechanism is monotone (i.e., if buyer type $v_b$ trades with seller type $v_s$, then all type-pairs $(v_b', v_s')$ for which $v_b' > v_b$ and $v_s' < v_s$ also trade). This implies that if $c = (c_b, c_s)$ and $c' = (c_b', c_s')$ are extreme vertices and $c_b < c_b'$, then $c_s < c_s'$. Thus, extreme vertices can be ordered. If $c = (c_b, c_s)$ and $c' = (c_b', c_s')$ are two extreme vertices for which $c_b < c_b'$ (and therefore also $c_s < c_s'$), then we say that $c$ is *smaller than* $c'$, and denote $c < c'$. We say that $c$ and $c'$ are *consecutive* extreme vertices if there is no extreme vertex between them. Note that if $c$ and $c'$ are two consecutive extreme vertices, where $c < c'$, then there is a buyer (resp., seller) interval message for which the lower bound is $c_b$ (resp. $c_s$) and the upper bound is $c_b'$ (resp., $c_s'$).

The credibility of the intermediation mechanism $\Gamma$ implies the following properties:

(P1) If $c$ and $c'$ are consecutive extreme vertices and $c < c'$, then $\mathbb{E}_{F_b}[c_b, c_b'] \leq \mathbb{E}_{F_s}[c_s, c_s']$.

(P2) If $c = (c_b, c_s)$ and $c' = (c_b', c_s')$ are two extreme vertices and $c' > c$, then $c_s' > c_b$.

(P3) If $c = (c_b, c_s)$, $c' = (c_b', c_s')$ and $c'' = (c_b'', c_s'')$ are three extreme vertices and $c'' > c' > c$, then $c_b'' > c_s$.

To see property (P1), suppose that $m_b$ is the buyer message for which $c_b = \underline{m}_b$, and that $m_s$ is the seller message for which $c_s' = \bar{m}_s$ (e.g., in Figure 4(a)), the smallest extreme vertex satisfies $c_b = \underline{m}_b^1$ and $c_s = \bar{m}_s^1$). Since $c$ and $c'$ are extreme vertices, there is no trade for the message pair $(m_b, m_s)$. By credibility, it must be the case that the buyer's mean type in $m_b$ is lower than the seller's mean type $m_s$. Property (P2) immediately follows since if $c_s' \leq c_b$ then, for any two consecutive extreme vertices $c'' \geq c$ and $c''' \leq c'$ for which $c'' < c'''$, we have that $c_s''' \leq c_b''$, contradicting (P1). To see property (P3), suppose that $m_b'$ is the buyer message for which $c_b' = \underline{m}_b'$, and that $m_s'$ is the seller message for which $c_s' = \bar{m}_s'$. Since $c'$ in an extreme vertex, there is trade for the message pair $(m_b', m_s')$. By credibility, it must be that $\bar{m}_b' > \underline{m}_s'$. Since $\bar{m}_b' \leq c_b''$ and $\underline{m}_s' \geq c_s$, property (P3) follows.

Finally, we adopt several definitions that are related to the agents' type distributions. For each agent $i$, we denote by $f_i^{\max} = \max_{v_i}(f_i(v_i))$ and $f_i^{\min} = \min_{v_i}(f_i(v_i))$ the maximum and minimum of the probability density function $f_i$, respectively. We define $f(v_b, v_s) = f_b(v_b) \cdot f_s(v_s)$ to be the joint probability density over $V$ and $f^{\min} = \min_{v_b, v_s}(f(v_b, v_s))$ and $f^{\max} = \max_{v_b, v_s}(f(v_b, v_s))$ to be the minimal and maximal values of $f(v_b, v_s)$, respectively, over $V$.

## *Proof of Proposition 4*

Fix the type distributions $F = \{F_s, F_b\}$. Denote by $S^*$ the supremum surplus that is attainable by all (budget-balanced) intermediation mechanisms when the type distributions are given by $F$. Define $S^{pp} = \max_x \int_{\underline{v}_s}^{x} \int_{x}^{\bar{v}_b} (v_b - v_s)\, dF_b(v_b)\, dF_s(v_s)$ to be the surplus that is generated by the optimal posted price when the type distributions are given by $F$. If $S^* = S^{pp}$, the proof is complete since, as we show in Proposition 5, an intermediation mechanism that implements the optimal posted price exists. In this intermediation mechanism, the message set of each agent contains two messages.

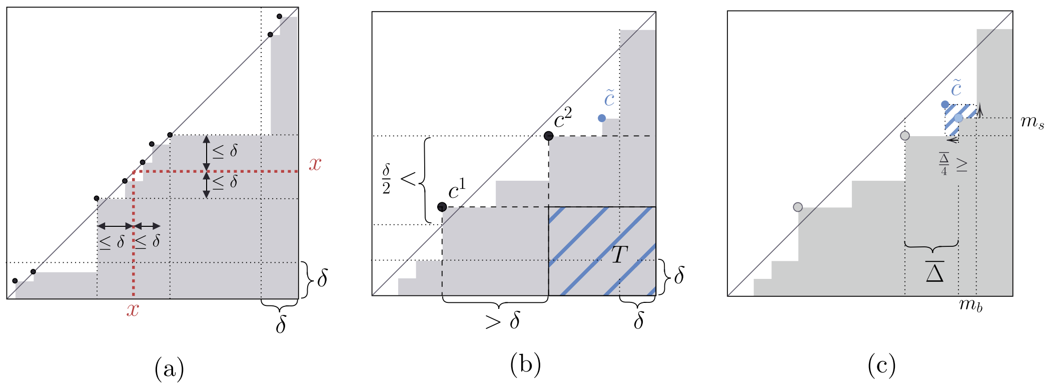


FIGURE 6. (a) A mechanism that is $\delta$-close to posted price $x$. (b), (c) Steps in the proof of Lemma A.2.

Suppose that $S^* > S^{pp}$. We begin by showing that if an intermediation mechanism generates a surplus higher than $\frac{1}{2}(S^* + S^{pp})$ then its extreme vertices are not all "too close" to some posted price. We say that an intermediation mechanism is $\delta$-close to some posted price $x$ if all extreme vertices for which the buyer type is not close to $\bar{v}_b$ and the seller type is not close to $\underline{v}_s$ are close to $(x, x)$.

DEFINITION 5. Given $\delta > 0$, an intermediation mechanism is $\delta$-close to the posted-price $x$ if any extreme vertex $(c_b, c_s)$ for which $c_b < \bar{v}_b - \delta$ and $c_s > \underline{v}_s + \delta$ satisfies $|c_b - x| < \delta$ and $|c_s - x| < \delta$.

LEMMA A.1. *If $S^* > S^{pp}$, then there exists $\hat{\delta} > 0$ such that any intermediation mechanism that generates a surplus higher than $\frac{1}{2}(S^* + S^{pp})$ is not $\hat{\delta}$-close to any posted price.*

Figure 6(a) illustrates a mechanism that is $\delta$-close to posted price $x$. Intuitively, when $\delta$ is sufficiently small, the surplus generated by the mechanism cannot be substantially higher than that of the posted price and, therefore, cannot exceed $\frac{1}{2}(S^* + S^{pp})$.

In what follows, we show that given any intermediation mechanism that is not $\hat{\delta}$-close to any posted price and has more than some fixed number ($\bar{K}$) of extreme vertices, we can construct another intermediation mechanism with less than that number of extreme vertices that generates a higher surplus. To simplify the exposition, we temporarily relax the budget balance restriction on intermediation mechanisms and allow intermediation mechanisms to end up with a budget surplus/deficit. We emphasize that the final intermediation mechanism we construct is exactly budget-balanced.

The next two lemmata comprise the main part of the proof.

LEMMA A.2. *Given $\delta < (f^{\min})^{-0.5}$, there exist $r > 0$ and $\bar{s}_1 > 0$ such that for any intermediation mechanism with $\psi(p) \leq 0$ that is not $\delta$-close to any posted-price, and any $s \leq \bar{s}_1$, one can construct an intermediation mechanism with weakly fewer extreme vertices, such that: (i) the generated surplus increases by at least $s$, and (ii) $\psi$ increases by no more than $s/r$.*

Lemma A.3. *Given $r > 0$ and $\bar{s}_2 > 0$, there exists $\bar{K}$ such that, for any intermediation mechanism with more than $\bar{K}$ extreme vertices, one can construct an intermediation mechanism with fewer than $\bar{K}$ extreme vertices, such that: (i) the generated surplus decreases by some $s < \bar{s}_2$, and (ii) $\psi$ decreases by at least $s/r$.*

Compute $\hat{\delta}$ from Lemma A.1 and compute $r$ and $\bar{s}_1$ from Lemma A.2 with $\delta = \min\{\hat{\delta}, (f_{\min})^{-0.5}\}$. Compute $\bar{K}$ from Lemma A.3 with $r$ and $\bar{s}_2 = \min(\bar{s}_1, \frac{1}{4}(S^* - S^{pp}))$. The explicit formulas for $r$, $\bar{s}_1$, and $\bar{K}$ appear in the proofs of the lemmata. Importantly, they all depend solely on $F$ and not on any particular intermediation mechanism.

Now suppose that $\Gamma$ is an intermediation mechanism with $\psi(p) = 0$, that generates a surplus which is higher than $\frac{3}{4}S^* + \frac{1}{4}S^{pp}$ and has more than $\bar{K}$ extreme vertices. Using $\Gamma$, construct an intermediation mechanism $\Gamma'$ according to Lemma A.3, with $r$ and $\bar{s}_2$ as computed above. Lemma A.3 guarantees that the lost surplus, $s$, is smaller than $\bar{s}_2 \le \frac{1}{4}(S^* - S^{pp})$. Thus, $\Gamma'$ generates a surplus which is higher than $\frac{1}{2}(S^* + S^{pp})$, and according to Lemma A.1, is not $\hat{\delta}$-close to any posted price. Lemma A.3 also guarantees that $\psi(p') < \psi(p) - s/r = -s/r$, where $p'$ is the allocation rule of $\Gamma'$.

Using $\Gamma'$, construct an intermediation mechanism $\Gamma''$ according to Lemma A.2 (with $\delta$ and $s$ as computed above). The surplus generated by $\Gamma''$ is at least as large as that generated by $\Gamma$. We also have that $\psi(p'') < \psi(p') + s/r < 0$, where $p''$ is the allocation rule of $\Gamma''$.

The mechanism $\Gamma''$ is not $\hat{\delta}$-close to any posted-price (because it generates surplus higher than $\frac{3}{4}S^* + \frac{1}{4}S^{pp}$). Now, apply Lemma A.2 iteratively on the mechanism at hand, with $s = \bar{s}_1$, until the resulting mechanism is such that one additional iteration would imply a positive value of $\psi$. Since each iteration increases the mechanism's surplus by $\bar{s}_1$, which is a constant that is independent of the mechanism, the process necessarily converges (this is because there is an upper-bound on the surplus that an intermediation mechanism with a negative $\psi$ can generate; for example, it cannot do better than a budget-balanced full-commitment trade mechanism). Then we apply Lemma A.2 once again, picking $s \in (0, \bar{s}_1)$ such that the resulting mechanism would have $\psi$ that is exactly equal to 0. The existence of such a value of $s$ is guaranteed because, as we explain in the proof of Lemma A.2, whenever the $\psi$ increases, this change is continuous in $s$. This resulting mechanism is exactly budget balanced.

We have thus shown that any budget-balanced intermediation mechanism with a surplus above $(3S^* + S^{pp})/4$ and more than $\bar{K}$ extreme vertices can be improved upon by a budget-balanced intermediation mechanism with fewer than $\bar{K}$ extreme vertices. Thus, $S^*$ is the supremum surplus of the class of budget-balanced intermediation mechanisms with fewer than $\bar{K}$ extreme vertices. The following compactness argument then completes the proof.

Lemma A.4. *For any integer $K > 0$, the supremum surplus over the set of budget-balanced intermediation mechanisms with no more than $K$ extreme vertices is attainable.*

Thus, $S^*$ is attainable by a budget-balanced intermediation mechanism with fewer than $\bar{K}$ extreme vertices.

Finally, note that if an intermediation mechanism has $\bar{K}$ extreme vertices (or less), then the buyer's and seller's message sets cannot contain more than $\bar{K}+1$ messages each. In the statement of the proposition, we refer to this number as $\bar{K}_F$ to emphasize that it depends solely on the type distributions $F$.

### *Proof of Lemma A.1*

Suppose that $S^* > S^{pp}$ and consider an intermediation mechanism $\Gamma$ with a set of extreme vertices $C$. The set of type-pairs that trade in $\Gamma$ is given by

$$T_\Gamma = \bigcup_{(c_b, c_s) \in C} \{(v_s, v_b) \in V : v_s \leq c_s, v_b \geq c_b\}.$$

Suppose that $\Gamma$ is $\delta$-close to some posted price $x$. Then, by Proposition 5 there exists an intermediation mechanism that implements $x$ and in which the set of type-pairs that trade is

$$T_x = \{(v_s, v_b) \in V : v_s \leq x, v_b \geq x\}.$$

The sets $T_\Gamma$ and $T_x$ differ only for the type-pairs $(v_s, v_b)$ for which $\bar{v}_b - v_b < \delta$, $v_s - \underline{v}_s < \delta$, $|x - v_b| < \delta$ or $|x - v_s| < \delta$ (see, e.g., the illustration in Figure 6(a)). Thus, the probability mass of the set $T_\Gamma \oplus T_x$ (i.e. the set of type-pairs that are in either $T_\Gamma$ or $T_x$, but not both) goes to zero as $\delta$ goes to zero. Since the social surplus that is generated by trade for each type-pair is bounded above by $\bar{v}_b - \underline{v}_s$, then the difference in the ex ante social surpluses that are generated by $\Gamma$ and by the posted price $x$ goes to zero as $\delta$ goes to zero. Therefore, there exists $\hat{\delta} > 0$ such that if $\Gamma$ is $\delta$-close to some posted price $x$ then its surplus cannot exceed $S^{pp} + \frac{1}{2}(S^* - S^{pp})$.

### *Proof of Lemma A.2*

The proof consists of two parts. First, given $\delta < (f^{\min})^{-0.5}$ we show that any intermediation mechanism that is not $\delta$-close to any posted price, and for which $\psi(p) \leq 0$, must have an extreme vertex $\tilde{c} = (\tilde{c}_b, \tilde{c}_s)$ for which $\tilde{c}_b - \tilde{c}_s \geq \bar{\Delta} \equiv \frac{\delta^3}{2} f^{\min}$, i.e., $\tilde{c}$ is "significantly" below the equi-type line. Then we show a modification of the mechanism that "adjusts the position" of $\tilde{c}$ and increases the mechanism's surplus by at least $s$, for any $s \leq \bar{s}_1 \equiv \bar{\Delta}/2 \cdot (\bar{\Delta}/4)^2 \cdot f^{\min}$. This modification increases the mechanism's budget deficit by no more than $s/r$, where $r = \bar{\Delta}/(2(1/f_b^{\min} + 1/f_s^{\min}) - \bar{\Delta})$.

*Part I.* Suppose that $\Gamma = \langle M, p, t_b, t_s \rangle$ is an intermediation mechanism for which $\psi(p) \leq 0$. If $\Gamma$ has an extreme vertex $c$ for which $(c_b - c_s) \geq \delta/2$, then it is the vertex $\tilde{c}$ we are looking for, since $\delta/2 > \bar{\Delta}$. Otherwise, all extreme vertices $(c_b, c_s)$ satisfy $c_b - c_s < \delta/2$. Let $c^1$ be an extreme vertex that satisfies $c_s^1 > \underline{v}_s + \delta$ and for which there is no other extreme vertex $c'$ with $c'_b < c_b^1 - \delta/2$ and $c'_s > \underline{v}_s + \delta$. For example, $c^1$ can be the minimal vertex satisfying $c_s^1 > \underline{v}_s + \delta$ if one exists, or an extreme vertex that is sufficiently close to the lowest accumulation point of extreme vertices otherwise. Since the mechanism is not $\delta$-close to any posted price, in particular it is not $\delta$-close to the posted price $x = c_b^1$, then there exists an extreme vertex $c^2 > c^1$ for which either $c_s^2 > x + \delta$ or $c_b^2 > x + \delta$ (it cannot be that $c_b^2 < x - \delta$ or $c_s^2 < x - \delta$ because $c_b^2 \geq c_b^1 - \delta/2$ and $c_b^2 - c_s^2 \leq \delta/2$). Since $c_b^2 - c_s^2 < \delta/2$, we are guaranteed that $c_s^2 > x + \delta/2$ in both cases. Therefore, $c_s^2 - c_b^1 > \delta/2$. See Figure 6(b) for an illustration.

Consider the set of type-pairs $T = \{(v_s, v_b) \in V : v_b \geq c_b^2 \text{ and } v_s \leq c_s^1\}$ (see Figure 6(b)). All type-pairs in $T$ trade in the mechanism $\Gamma$. Let $\bar{T}$ be the set of type pairs that trade in $\Gamma$ and are not in the set $T$. Thus, we can decompose $\psi(p)$ as follows:

$$\psi(p) = \int\int_{(v_b, v_s) \in T} \big(\omega_s(v_b) - \omega_b(v_s)\big)\, dF(v_s)\, dF(v_b) + \int\int_{(v_b, v_s) \in \bar{T}} \big(\omega_s(v_b) - \omega_b(v_s)\big)\, dF(v_s)\, dF(v_b) \tag{5}$$

Since $c_b^2 < \bar{v}_b - \delta$ and $c_s^1 > \underline{v}_s + \delta$, the probability mass of all type-pairs in $T$ is at least $\delta^2 \cdot f^{\min}$. Also, since $\omega_s$ and $\omega_b$ are both weakly increasing functions, then for every type-pair $(v_b, v_s) \in T$ we have that $\omega_s(v_b) \geq c_s^2$ and $\omega_b(v_s) \leq c_b^1$ and, therefore, $\omega_s(v_b) - \omega_b(v_s) > \delta/2$. Thus, the first term in the right-hand side of equation (5) is at least $\bar{\Delta} = (\delta/2) \cdot \delta^2 f^{\min}$. Since $\psi(p) \leq 0$, it must be that the second term in the right-hand side of equation (5) is smaller than $-\bar{\Delta}$. Since the probability mass of all type-pairs in $\bar{T}$ is obviously bounded above by 1, there must be at least one type-pair $(v_s, v_b)$ for which $(\omega_s(v_b) - \omega_b(v_s)) < -\bar{\Delta}$. Therefore, $\tilde{c} = (\omega_b(v_s), \omega_s(v_b))$ is an extreme vertex satisfying $\tilde{c}_b - \tilde{c}_s \geq \bar{\Delta}$.[36]

*Part II.* Since $\tilde{c}_s < \tilde{c}_b$ then neither the buyer type $\tilde{c}_b$ nor the seller type $\tilde{c}_s$ belong to an interval of types that are fully separated. This is because, by the corollary of Lemma 4, if all of agent $i$'s types over some interval are fully separated, then so do all types of agent $-i$ on the same interval. But then, by credibility, the highest seller type that trades with buyer type $c_b$ must be equal to $c_b$. Let $m_b$ be the buyer's message (interval) satisfying $\underline{m}_b = \tilde{c}_b$, and let $m_s$ be the seller's message (interval) satisfying $\bar{m}_s = \tilde{c}_s$ (see Figure 6(c))

Consider the following modification of the mechanism, that is parametrized by some $\Delta \in (0, \bar{\Delta}/4]$ and illustrated in Figure 6(c): decrease $\underline{m}_b$ by $\Delta$, increase $\bar{m}_s$ by $\Delta$ and update the allocation rule $p$ to be credible-minimal with respect to the modified message set. This modification, which resembles the one performed in the proof of Proposition 3, adds trade (only) for type-pairs that correspond to the hatched area in Figure 6(c) (for an explanation why, see the discussion that follows Proposition 3 in the text). In some cases, the modification may also result in the elimination of some trade. This happens if, for some message pair, the seller's mean type switched from being smaller than the buyer's mean type to being (weakly) larger. The allocation rule for this messages pair then changes from "trade" to "no trade." Note, however, that if such a change happens, it only increases the surplus generated by the mechanism, since it eliminates nonbeneficial trade. Furthermore, since it eliminates trade it also reduces the amount of information rents that have to be paid to the agents.

[36]To see that $(\omega_b(v_s), \omega_s(v_b))$ is an extreme vertex, note that $\omega_b(v_s)$ is the buyer-type coordinate of some extreme vertex $c'$, and $\omega_s(v_b)$ is the seller-type coordinate of some extreme vertex $c''$, and $c' \leq c''$; thus, if $c' \neq c''$ then $c_s'' < c_b'$, contradicting property (P2) defined above.

Denote the probability mass of all the type pairs for which trade is added by $P_\Delta$ (these are the type pairs in the hatched area in Figure 6(c)). Since every such type-pair $(v_b, v_s)$ satisfies $v_b - v_s > \bar{\Delta}/2$, then the increase in the generated surplus, $s$, is at least $P_\Delta \cdot \bar{\Delta}/2$. In order to see how large the increase in surplus $s$ can be and to determine the value of $\bar{s}_1$, note that when $\Delta = \bar{\Delta}/4$ the area of added trade is greater than $(\bar{\Delta}/4)^2$ and, therefore, we are guaranteed that $P_\Delta \geq (\bar{\Delta}/4)^2 f^{\min}$. Thus, $s$ can take any value between 0 and $\bar{s}_1 = (\bar{\Delta}/2) \cdot (\bar{\Delta}/4)^2 \cdot f^{\min}$.

Adding trade to the mechanism requires paying more information rents to the agents. The amount of additional rents, denoted $I_\Delta$, satisfies $I_\Delta < P_\Delta \cdot (1/f^b_{\min} + 1/f^s_{\min})$. To see this, recall that in any incentive-compatible mechanism, the expected information rent paid to buyer type $v_b$ (above the payoff of the lowest type $\underline{v}_b$) is $I(v_b) \equiv \int_{\underline{v}}^{v_b} \bar{p}_b(x)\, dx$, where $\bar{p}_b(x)$ is the expected probability of trade for type $x \in V_b$. Thus, the maximal increase in information rent to any buyer type due to the modification is $P_\Delta / f_b^{\min}$.[37] Since the mass of *all* buyer types is 1, an upper bound on the added information rent to all buyer types is $1 \cdot P_\Delta / f_b^{\min}$. The same argument for the seller gives an upper bound of $1 \cdot P_\Delta / f_s^{\min}$. The total effect is then bounded above by $P_\Delta \cdot (1/f_b^{\min} + 1/f_s^{\min})$.

Denote the change in $\psi$ due to the modification by $\psi_\Delta$. A well-known result in standard mechanism design (see, e.g., Börgers (2015)), that holds also for intermediation mechanisms, is that the sum of the expected transfers to the agents is equal to the sum of the expected information rents minus the expected social surplus (plus the difference between the payoff of the buyer's lowest type $\underline{v}_b$ and the seller's highest type $\bar{v}_s$). Thus, $\psi_\Delta = I_\Delta - s$. Using the bounds $I_\Delta < P_\Delta \cdot (1/f^b_{\min} + 1/f^s_{\min})$ and $s > P_\Delta \cdot \bar{\Delta}/2$, we established before, we then obtain

$$\psi_\Delta = s\left(\frac{I_\Delta}{s} - 1\right) < s\left(\frac{P_\Delta \cdot \left(\dfrac{1}{f^b_{\min}} + \dfrac{1}{f^s_{\min}}\right)}{P_\Delta \cdot \bar{\Delta}/2} - 1\right) = \frac{s}{r},$$

where $r \equiv \bar{\Delta}/(2(1/f^b_{\min} + 1/f^s_{\min}) - \bar{\Delta})$.

Finally, note that the measure of the set of types for which trade is added (the hatched area in Figure 6(c)) is continuous in and so is $P_\Delta$. Therefore, if the change in information rents is positive, then it is continuous in $s$ (i.e., $I_\Delta$ does not have positive discontinuous jumps as $s$ increases).[38] Consequently, whenever $\psi$ increases, $\psi_\Delta$ is continuous in $s$.

### *Proof of Lemma A.3*

We start by dividing the intersection $V_s \cap V_b$ into $l$ equal-sized segments. To simplify notation, we assume throughout the proof that $|V_s \cap V_b| = 1$. Then the type-space $V$ is

[37] To see this, let the function $\xi(x, y)$ equal 1 if, in the modification described above, the type-pair $(x, y)$ has been added to trade and 0 otherwise. We ignore type-pairs for which trade was eliminated (if there are any), since this only reduced the amount of required information rents. Then the *difference* in the information rent to type $v_b$ due to the modification is at least $\int_{\underline{v}_b}^{v_b} \int_{\underline{v}_s}^{\bar{v}_s} \xi(x, y) f_s(y)\, dy dx < \frac{1}{f^b_{\min}} \int_{\underline{v}_b}^{\bar{v}_b} \int_{\underline{v}_s}^{\bar{v}_s} \xi(x, y) f_s(y) f_b(x)\, dy dx = \frac{1}{f^b_{\min}} \cdot P_\Delta$.

[38] $I_\Delta$ may have negative discontinuous jumps if the modification eliminates nonbeneficial trade.

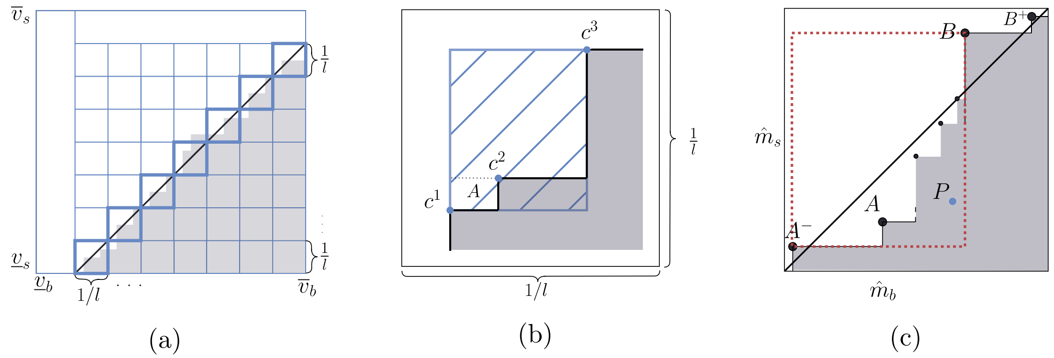


FIGURE 7. Steps in the proof of Lemma A.3.

split into (at most) $l^2 + 2$ regions, i.e., $l^2$ *boxes* of size $(1/l) \times (1/l)$ and 2 rectangles, as depicted in Figure 7(a). We say that a box $\beta = [v_b, v_b + 1/l] \times [v_s, v_s + 1/l]$ is *on the equi-type line* if $v_b = v_s$. Note that only boxes on the equi-type line can contain more than two extreme vertices. This is because, in the case of all other regions, having three extreme vertices (or more), e.g., $c^1$, $c^2$, and $c^3$ where $c^1 < c^2 < c^3$, violates (P2) if the region is below the equi-type line (since in that case $c_b^1 < c_s^2$) and violates (P3) if the region is above that line (since $c_b^3 > c_s^1$).

The proof consists of four steps. In step I, we find two lower bounds on the number of boxes $l$, denoted $\hat{l}_1$ and $\hat{l}_2$. Assuming that $l$ is larger than $\max\{\hat{l}_1, \hat{l}_2\}$, in step II we show a modification that is applicable to any intermediation mechanism, after which no box or rectangle in the division we defined above contains more than 5 extreme vertices. Thus, after the modification the mechanism has no more than $5l + 2(l^2 - l + 2)$ extreme vertices.[39] Crucially, the modification maintains the credibility of the mechanism and changes the allocation rule only for type-pairs within boxes on the equi-type line. In step III, we find another lower bound on the number of boxes, denoted $\hat{l}_3$, which guarantees that the surplus loss due to the modification $(s)$ is smaller than the given constant $\bar{s}_2$. In step IV, we find one more lower bound, denoted $\hat{l}_4$, which guarantees that the reduction in $\psi$ due to the modification is *at least* $s/r$, where $r$ is the constant given in the statement of the lemma. We conclude by explaining how the constants we find are combined to construct the desired intermediation mechanism.

In what follows, we use the following definitions: $R \equiv \max\{f_b^{\max}/f_b^{\min}, f_s^{\max}/f_s^{\min}\} \geq 1$ and $\hat{\kappa} \equiv 1/(224R + 8)$.

*Step I: Finding two lower bounds on the box size $l$.* Consider a box $\beta \subset V_b \times V_s$. Let $\rho_\beta = (\max_{(v_b, v_s) \in \beta} f(v_b, v_s))/(\min_{(v_b, v_s) \in \beta} f(v_b, v_s))$ be the upper bound on the ratio of values of $f$ for any two type-pairs within the box $\beta$. Define $\rho \equiv \max_\beta \rho_\beta$ to be the uniform upper bound over all the boxes. Because $f(v_b, v_s)$ is uniformly continuous and bounded away from zero over the entire type space, then $\rho$ approaches 1 as the size of the boxes approaches 0. In other words, the conditional type distribution within each

[39]That is, at most 5 vertices in each box on the equi-type line, and at most 2 extreme vertices in every other region.

box becomes close to uniform as the number of boxes increases. Our first lower bound on the number of boxes $l$ guarantees that $\rho$ is small:

*Lower bound 1.* Denote by $\hat{l}_1$ the number of boxes for which $l > \hat{l}_1$ implies $\rho \leq 2$.

Thus, when the number of boxes is greater than $\hat{l}_1$, we are guaranteed that for any box $\beta$ on the equi-type line, and any two disjoint subsets of type-pairs $T_1, T_2 \subset \beta$, if the ratio between the Lebesgue measures of $T_1$ and $T_2$ is $\alpha$, then the ratio between the probability mass of type-pairs in $T_1$ and the probability mass of type-pairs in $T_2$ is at least $\alpha/2$.[40]

For the second lower bound on the number of boxes $l$, we start with the following lemma.

Lemma A.5. *For any $\kappa > 0$, there exists $\tilde{l}$ such that for every number of boxes $l > \tilde{l}$ and every three extreme vertices $c^1 \leq c^2 \leq c^3$ that are all within the same box:*

*(i) If $(c_s^3 - c_s^2)/(c_b^3 - c_b^2) > 1 + \kappa$, then $\mathbb{E}[c_s^1, c_s^2] \geq \mathbb{E}[c_b^1, c_b^2]$ implies $\mathbb{E}[c_s^1, c_s^3] \geq \mathbb{E}[c_b^1, c_b^3]$, and*

*(ii) If $(c_b^2 - c_b^1)/c_s^2 - c_s^1 > 1 + \kappa$, then $\mathbb{E}[c_s^2, c_s^3] \geq \mathbb{E}[c_b^2, c_b^3]$ implies $\mathbb{E}[c_s^1, c_s^3] \geq \mathbb{E}[c_b^1, c_b^3]$.*

*These properties hold even if $c^1$, $c^2$ or $c^3$ is not an extreme vertex but an accumulation point of extreme vertices.*

In order to highlight the intuition of the first part of the lemma, consider the two consecutive extreme vertices $c^1$ and $c^2$. By Property (P1), there is no trade when the agents report messages $m_b = [c_b^1, c_b^2]$ and $m_s = [c_s^1, c_s^2])$. This is represented by area A in Figure 7(b). Suppose we merge $m_b$ (resp., $m_s$) with all buyer (resp., seller) messages up to $c_b^3$ (resp., $c_s^3$). What would guarantee that there is no trade for the unified message pair $([c_b^1, c_b^3], [c_s^1, c_s^3])$? If the distribution were uniform, then knowing that we added more seller types (to $m_s$) than buyer types (to $m_b$), i.e., $(c_s^3 - c_s^2)/(c_b^3 - c_b^2) > 1$, would suffice. But since the conditional distribution is only close to uniform within the small box, then we need to correct by adding a small proportion $\kappa$ of additional high seller types. The intuition for the second part is analogous. Our second lower bound is then:

*Lower bound 2.* $\hat{l}_2$ is the bound $\tilde{l}$ determined by Lemma A.5 when $\kappa = \hat{\kappa}$ (where $\hat{\kappa}$ is the constant defined above).

*Step II: Reducing the number of extreme vertices.* Assume that $l > \max\{\hat{l}_1, \hat{l}_2\}$. In this step, we show a modification of the mechanism that reduces the number of extreme vertices in each box on the equi-type line to be at most 5. The modification affects trade only for type-pairs within boxes on the equi-type line.

Consider a box $\beta$ on the equi-type line with more than 5 extreme vertices, like the one illustrated in Figure 7(c). Denote the smallest extreme vertex in the box by $A^-$. Denote the second-smallest vertex by $A$. If a smallest vertex does not exist, i.e., there is a converging sequence of vertices, then denote the smallest accumulation point of vertices by $A$, and let $A^- = A$. Similarly, denote the largest vertex in the box by $B^+$ and

[40]The probability of type-pairs in a set $T \subset \beta$ is given by $\int_{(v_b, v_s) \in T} f(v_b, v_s)\, dv_b\, dv_s$.

the second-largest by $B$. If a largest vertex does not exist, then denote the greatest accumulation point of vertices by $B$, and let $B^+ = B$. Lastly, denote by $P$ the type-pair $(P_b, P_s) = (\frac{7}{8}B_b + \frac{1}{8}A_b\,,\ \frac{1}{8}B_s + \frac{7}{8}A_s)$, where $A_s$ and $A_b$ are the seller- and buyer-type coordinates of the extreme vertex $A$ (resp.), and $B_s$ and $B_b$ are the seller- and buyer-type coordinates of $B$.

In order to reduce the number of extreme vertices in the box $\beta$, we merge the small messages (intervals) of each player into coarser ones. In what follows, when we say that we merge messages of player $i$ to create a message $m_i = [v_i', v_i'']$ we mean that: (i) we replace all agent $i$'s messages between $v_i'$ and $v_i''$ by a single message $m_i = [v_i', v_i'']$, (ii) we update the mechanism's allocation rule to be credible-minimal with respect to the updated message sets, and (iii) we adjust the transfer rules so that incentive compatibility is maintained.

Note that if $c^1$ and $c^2$ are two extreme vertices such that $c^1 < c^2$, and we merge the buyer's messages to create $\hat{m}_b = [c_b^1, c_b^2]$ and merge the seller's messages to create $\hat{m}_s = [c_s^1, c_s^2]$, then the modification does *not* affect the allocation rule for type-pairs outside of $\hat{m}_b \times \hat{m}_s$. To see this, consider, e.g., a seller type $v_s < c_s^1$ who sends a message $m_s \in M_s$ in the original mechanism. Since the modification does not affect message $m_s$, type $v_s$ keeps on sending $m_s$ even after the modification. Moreover, in the original mechanism $v_s$ trades with all buyer types in $[c_b^1, c_b^2]$, and we therefore deduce that the seller mean type in the message (interval) $m_s$ is smaller than the buyer mean type in all the messages (i.e., subintervals) that buyer types in $[c_b^1, c_b^2]$ send in the original mechanism. Thus, the seller mean type in $m_s$ is smaller than the buyer mean type also in the merged message $\hat{m}_b$ and, therefore, $p(m_s, \hat{m}_b) = 1$ by credibility. An analogous argument shows that seller type $v_s > c_s^2$, who did not not trade with any buyer type in $[c_b^1, c_b^2]$ before the modification, does not trade with any buyer type in $[c_b^1, c_b^2]$ after the modification. The arguments are analogous for the buyer side.

For each box $\beta$, we modify the mechanism according to the relevant case.

*Case 1*: $(B_s - A_s)/(B_b - A_b) \geq 2$ *or* $(B_s - A_s)/(B_b - A_b) \leq \frac{1}{2}$. Suppose that $(B_s - A_s)/(B_b - A_b) \geq 2$, as depicted in Figure 7(c). We then merge the buyer messages to create $\hat{m}_b = [A_b^-, B_b]$ and merge seller messages to create $\hat{m}_s = [A_s^-, B_s]$. After this modification, no type-pair in $\hat{m}_b \times \hat{m}_s$ trades. To see this, note that since $\hat{\kappa} < 1$ then $(B_s - A_s)/(B_b - A_b) > (1 + \hat{\kappa})$ and recall that $l > \hat{l}_2$ and that $\mathbb{E}[A_s^-, A_s] \geq \mathbb{E}[A_b^-, A_b]$ by property (P1) defined above. Therefore, by the first part of Lemma A.5 we have $\mathbb{E}[A_s^-, B_s] \geq \mathbb{E}[A_b^-, B_b]$.[41] Therefore, credibility implies that $p(\hat{m}_b, \hat{m}_s) = 0$.

If $(B_s - A_s)/(B_b - A_b) \leq \frac{1}{2}$, we merge buyer and seller messages to create $\hat{m}_b = [A_b, B_b^+]$ and $\hat{m}_s = [A_s, B_s^+]$, respectively. An analogous argument (using the second part of Lemma A.5) guarantees that after the modification no type pair in $\hat{m}_b \times \hat{m}_s$ trades.

In either case, in the modified mechanism there are at most 3 extreme vertices in the box $\beta$.

*Case 2*: $\frac{1}{2} < (B_s - A_s)/(B_b - A_b) < 2$ *and* $P_s > P_b$. In this case, which is illustrated in Figure 8(a), we merge buyer and seller messages to create $\hat{m}_b = [A_b, B_b]$ and $\hat{m}_s = [A_s, B_s]$, respectively. After this modification, no type-pair in $\hat{m}_b \times \hat{m}_s$ trades.

[41] In terms of Lemma A.5, $A^-$ is $c^1$, $A$ is $c^2$, and $B$ is $c^3$.

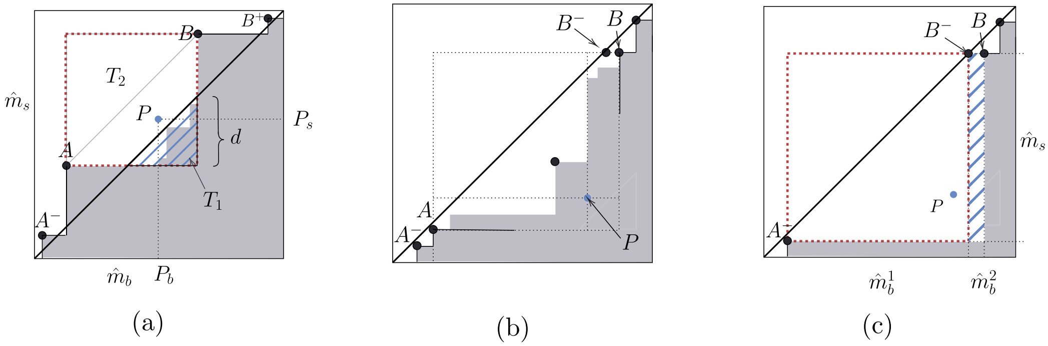


FIGURE 8. Steps in the proof of Lemma A.3.

To see this, note that the subset of type-pairs in $\hat{m}_b \times \hat{m}_s$ for which the surplus from trade is positive is $\{(v_b, v_s) \in \hat{m}_b \times \hat{m}_s | v_b \geq v_s\}$. This set corresponds to the triangle $T_1 = \langle (A_s, A_s), (B_b, A_s), (B_b, B_b) \rangle$ in the figure. The subset of type pairs in $\hat{m}_b \times \hat{m}_s$ for which the surplus from trade is negative is $\{(v_b, v_s) \in \hat{m}_b \times \hat{m}_s | v_b < v_s\}$. Consider its proper subset $\{(v_b, v_s) \in \hat{m}_b \times \hat{m}_s | A_s + (v_b - A_b)(B_s - A_s)/(B_b - A_b) < v_s\}$, which corresponds to the triangle $T_2 = \langle (A_b, A_s), (A_b, B_s), (B_b, B_s) \rangle$ in the figure.

The ratio between the areas of $T_2$ and $T_1$ is *at least* $\frac{128}{9}$.[42] Since $l > \hat{l}_1$, the probability mass of type-pairs in $T_2$ is larger than the probability mass of type pairs in $T_1$. Furthermore, for each type-pair in $T_1$ the surplus from trade is at most $d \equiv B_b - A_s$, while for each type-pair in $T_2$ the surplus destroyed by trade is at least $d$.[43] Thus, when the buyer reports $\hat{m}_b$ and the seller reports $\hat{m}_s$ then, on average, trade is not beneficial and, therefore, credibility implies $p(\hat{m}_s, \hat{m}_b) = 0$.

After the modification, there are no more than 4 extreme vertices in the box $\beta$.

*Case 3*: $\frac{1}{2} < (B_s - A_s)/(B_b - A_b) < 2$ *and* $P_s < P_b$ *and there is an extreme vertex in* $[A_b, P_b] \times [P_s, B_s]$. This case is depicted in Figure 8(b). Since there is an extreme vertex in $[A_b, P_b] \times [P_s, B_s]$, then all type-pairs in $[P_b, B_b] \times [A_s, P_s]$ trade. Suppose that $B_s - A_s \geq B_b - A_b$, and denote by $B^-$ the type pair $(B_b^-, B_s^-) = (\hat{\kappa} A_b + (1 - \hat{\kappa}) B_b, B_s)$ (the constant $\hat{\kappa}$ is defined before Step I). Now merge seller messages to create $\hat{m}_s = (A_s^-, B_s)$ and buyer messages to create $\hat{m}_b^1 = (A_b^-, B_b^-)$ and $\hat{m}_b^2 = (B_b^-, B_b)$, as depicted in Figure 8(c). After the modification, no type-pair in $\hat{m}_b^1 \times \hat{m}_s$ trades. This is because $(B_s - A_s)/(B_b^- - A_b) = (1/(1 - \hat{\kappa})) \cdot (B_s - A_s)(B_b - A_b) > 1 + \hat{\kappa}$ and $\mathbb{E}[A_s^-, A_s] \geq \mathbb{E}[A_b^-, A_b]$ (which is implied by property (P1), since $A^-$ and $A$ are consecutive extreme vertices). Therefore, by the first part of Lemma A.5, and since $l > \hat{l}_1$, we have that $\mathbb{E}[\hat{m}_s] \geq \mathbb{E}[\hat{m}_b^1]$. By credibility, we then have that $p(\hat{m}_s, \hat{m}_b^1) = 0$. Whether or not the type pairs in $\hat{m}_b^2 \times \hat{m}_s$ trade depends on the type distributions.

42 Denote $x = \max(B_s - A_s, B_b - A_b)$ and $y = \min(B_s - A_s, B_b - A_b)$. The area of triangle $T_2$ is $0.5(x \cdot y)$. The area of triangle $T_1$ is at most $0.5 \cdot (\frac{1}{8}(x + y))^2$. Since $x \leq 2y$, the ratio $0.5(x \cdot y)/(0.5 \cdot ((1/8) \cdot (x + y))^2)$ is *at least* $\frac{128}{9}$ (when $x = 2y$).

43 This is because, using the notation of footnote 42, trade between each type-pair in $T_2$ creates a negative surplus that is worse than $-(y - d)$. Note that $d \leq \frac{1}{8}(x + y) \leq \frac{3y}{8}$, where the first inequality is because $P_s > P_b$ and the second is because $\frac{x}{y} \leq 2$. Thus, $|-(y - d)| > d$.

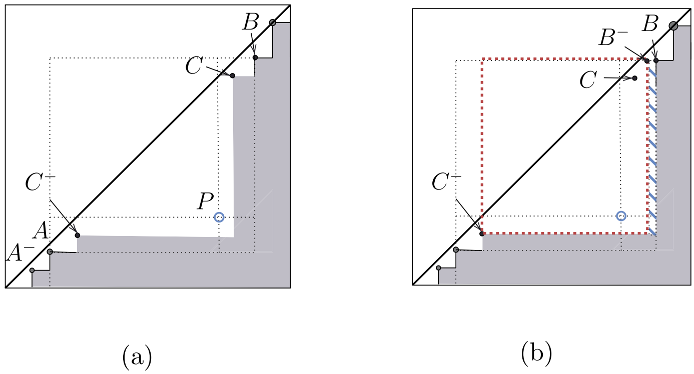


FIGURE 9. Steps in the proof of Lemma A.3.

Relative to the original mechanism, the modification eliminates trade for at least all type-pairs in $[P_b, B_b^-] \times [A_s, P_s]$. If the modification adds trade, it does so only for a subset of type pairs in $[B_b^-, B_b] \times [P_s.B_s]$. The ratio between the area in which trade is eliminated and the area in which trade is added is at least $((B_b^- - P_b) \cdot (P_s - A_s))/((B_b - B_b^-) \cdot (B_s - P_s)) = (1 - 8\hat{\kappa})/56\hat{\kappa} = 4R$. Since $l > \hat{l}_1$, we are guaranteed that the ratio between the probability mass of the type-pairs for which trade is eliminated and the probability mass of type-pairs for which trade is added is at least $2R$.

The analysis of the case in which $B_s - A_s < B_b - A_b$ is similar. Denote the point $A^+ \equiv \langle A_b, (1 - \hat{\kappa})A_b + \hat{\kappa}B_b\rangle$, merge buyer messages to create $\hat{m}_b = (A_b, B_b^+)$ and merge seller messages to create $\hat{m}_s^1 = (A_s, A_s^+)$ and $\hat{m}_s^2 = (A_s^+, B_s^+)$. Then use the second part of Lemma A.5 to show that the ratio between the probability mass of the type-pairs for which trade is eliminated and the probability mass of type-pairs for which type is added is at least $2R$.

In either case, in the modified mechanism there are at most 3 extreme vertices in the box $\beta$.

*Case 4*: $\frac{1}{2} < (B_s - A_s)/(B_b - A_b) < 2$ and $P_s < P_b$ *and there is no extreme vertex in* $[A_b, P_b] \times [P_s, B_s]$. This case is depicted in Figure 9(a). Denote the smallest extreme vertex in $[P_b, B_b] \times [P_s, B_s]$ by $C$ (if there is no such smallest vertex, then let $C$ be the smallest accumulation point of vertices). Denote the first extreme vertex (or accumulation point) before it by $C^-$. If $(B_s - C_s)/(B_b - C_b) \geq 2$, merge buyer and seller messages to create $\hat{m}_b = (C_b^-, B_b)$ and $\hat{m}_s = (C_s^-, B_s)$, respectively. The argument in case 1 above shows that after this modification no type pair in $\hat{m}_b \times \hat{m}_s$ trades. Similarly, if $(B_s - C_s)/(B_b - C_b) \leq \frac{1}{2}$, then merge buyer and seller messages to create $\hat{m}_b = (C_b, B_b^+)$ and $\hat{m}_s = (C_s, B_s^+)$, respectively, in order to guarantee that no type-pair in $\hat{m}_b \times \hat{m}_s$ trades.

Otherwise, $\frac{1}{2} \leq (B_s - C_s)/(B_b - C_b) \leq 2$. Denote the type pair $B^- \equiv \langle \hat{\kappa}C_b + (1 - \hat{\kappa})B_b, B_s\rangle$ as illustrated in Figure 9(b). Then merge seller messages to create $\hat{m}_s = (C_s^-, B_s)$ and buyer messages to create $\hat{m}_b^1 = (C_b^-, B_b^-)$ and $\hat{m}_b^2 = (B^-, B_b)$. The argument in case 3 above shows that in the modified mechanism no type pair in $\hat{m}_b^1 \times \hat{m}_s$ trades. Whether or not type pairs in $\hat{m}_b^2 \times \hat{m}_s$ trade depends on the type distributions.

Relative to the original mechanism, the modification eliminated trade for at least all type pairs in $T_1 = [C_b, B_b^-] \times [P_s, C_s]$. If it adds trade, then it does so only for type pairs in

$T_2 = [B_b^-, B_b] \times [C_s, B_s]$. However, note that $C_s - P_s > \frac{3}{8}(B_s - A_s)$ and $B_s - C_s < \frac{4}{8}(B_s - A_s)$.[44] Therefore, the ratio between the areas of $T_1$ and $T_2$ is at least $((B_b^- - C_b) \cdot \frac{3}{8}(B_s - A_s))/((B_b - B_b^-) \cdot \frac{4}{8}(B_s - A_s)) = (3 - 3\hat{\kappa})/4\hat{\kappa} > 4R$. Since $l > \hat{l}_1$, we are guaranteed that the ratio between the probability mass of the type-pairs for which trade is eliminated and the probability mass of type-pairs for which trade is added is at least $2R$.

Now denote the rightmost extreme vertex in $[A_b, P_b] \times [A_s, P_s]$ by $D$ and the extreme vertex after it by $D^+$ and apply a symmetric argument to remove trade for type pairs in $[A_b, P_b] \times [A_s, P_s]$. Note that, by property (P2), there cannot be more than one extreme vertex in $[P_b, B_b] \times [A_s, P_s]$. Thus, after the modification there are no more than 5 extreme vertices in the box $\beta$.

*Step III—An upper bound on the lost surplus.* The modification described in step II changes the allocation rule only for type-pairs in the $l$ boxes on the equi-type line. An upper bound on the (expected) loss of surplus due to the modification is $\frac{f^{\max}}{l^2}$. This is because for each type-pair $(v_b, v_s)$ in a box on the equi-type line we have $|v_b - v_s| \leq 1/l$, i.e., trade between $v_b$ and $v_s$ cannot create or destroy surplus of more than $1/l$. In addition, the probability mass of all type-pairs in a box is no more than $(1/l^2) \cdot f^{\max}$. Thus, the expected lost surplus across all $l$ boxes on the equi-type line cannot exceed $l \cdot (1/l) \cdot (f^{\max}/l^2)$. The third lower bound on the number of boxes $l$ guarantees that this loss of surplus is smaller than $\bar{s}_2$:

*Lower bound 3.* $\hat{l}_3 = \sqrt{f^{\max}/\hat{s}}$.

*Step IV—Ratio between the lost surplus and the saved budget.* The modification described in step II reduces the amount of trade in the mechanism. In this step, we show that the ratio between the reduction in the required transfers to the agents $(\psi)$ and the lost surplus $(s)$ is increasing in the number of boxes $l$. Thus, for large enough $l$ the reduction in $\psi$ is larger than $s/r$, where $r$ is the given parameter.

Fix a box $\beta = [v, v + 1/l] \times [v, v + 1/l]$ on the equi-type line. Denote by $P_-$ the probability mass of all the type-pairs that trade in the original mechanism but not trade in the modified one. Denote by $P_+$ the probability mass of all type pairs that do not trade in the original mechanism but do trade in the modified one. $P_+$ is zero in cases 1 and 2 of step II. In cases 3 and 4, it may be positive but, as we showed, nonetheless satisfies

$$2R \cdot P_+ < P_-. \tag{6}$$

Denote the net reduction in trade probability by $P_\Delta \equiv P_- - P_+$. Since $R \geq 1$, then $P_+ < P_\Delta \leq P_-$ and $P_- + P_+ < 3P_\Delta$.

We begin by showing that the expected *reduction* in information rent to the buyer due to the modification in the box is at least

$$\left(\frac{1}{2}P_\Delta \cdot \frac{f_b^{\min}}{f_b^{\max}} \cdot \left(\bar{v}_b - \left(v + \frac{1}{l}\right)\right)\right) - \left(P_+ \cdot \frac{f_b^{\max}}{f_b^{\min}} \cdot \frac{1}{l}\right) \tag{7}$$

[44]To see that $B_s - C_s < \frac{4}{8}(B_s - A_s)$, note first that $B_s - C_s \leq 2(B_b - C_b)$. Since $C_b > P_b$, then $B_s - C_s < 2(B_b - P_b) = \frac{2}{8}(B_b - A_b)$, and since in case IV we assume that $(B_b - A_b) < 2(B_s - A_s)$ then $B_s - C_s < \frac{4}{8}(B_s - A_s)$. To see $C_s - P_s > \frac{3}{8}(B_s - A_s)$, note that $C_s - P_s = (B_s - A_s) - (B_s - C_s) - (P_s - A_s)$. Since $B_s - C_s < \frac{4}{8}(B_s - A_s)$ and $(P_s - A_s) = \frac{1}{8}(B_s - A_s)$, then $C_s - P_s > \frac{3}{8}(B_s - A_s)$.

To see this, recall again that the expected information rent paid to buyer type $v_b$ (above the payoff of the lowest type $\underline{v}_b$) in any incentive-compatible mechanism is given by $I(v_b) \equiv \int_{\underline{v}}^{v_b} \bar{p}_b(x)\,dx$. We divide the buyer types into three groups. First, all buyer types $v_b < v$ (recall that $v$ is the lowest buyer type in the box $\beta$) are unaffected by the modification, and their expected information rents are unchanged. Second, the expected information rents that are paid to buyer types $v_b \in [v, v + 1/l]$ *increase* by *at most* $P^+ \cdot (f_b^{\max}/f_b^{\min}) \cdot (1/l)$. This is because the probability mass of all buyer types in the interval $[v, v+1/l]$ is at most $(1/l) \cdot f_b^{\max}$ and the maximal increase in information rent paid to any buyer type in this interval cannot exceed $P_+/f_b^{\min}$.[45] Finally, the expected information rents that have to be paid to buyer types $v_b \in [v + 1/l, \bar{v}_b]$ *decrease* by *at least* $\frac{1}{2}(P_\Delta/f_b^{\max}) \cdot f_b^{\min}(\bar{v}_b - (v + 1/l))$. This is because the mass of all buyer types in the interval $[v + 1/l, \bar{v}_b]$ is at least $(\bar{v}_b - (v + 1/l)) \cdot f_b^{\min}$, and because for each buyer type in this interval the expected information rent decreases by at least $P_-/f_b^{\max} - P_+/f_b^{\min}$. By equation (6), we have that $P_-/f_b^{\max} - P_+/f_b^{\min} \geq (1 - (1/2R)(f_b^{\max}/f_b^{\min})) \cdot (P_-/f_b^{\max})$, and since $R \geq f_b^{\max}/f_b^{\min}$ and $P_\Delta \leq P_-$, then $(1 - (1/2R)(f_b^{\max}/f_b^{\min})) \cdot (P_-/f_b^{\max}) \geq \frac{1}{2}P_\Delta/f_b^{\max}$.

An analogous argument shows that the expected reduction in information rents paid to the seller is at least

$$\left(\frac{1}{2}P_\Delta \frac{f_s^{\min}}{f_s^{\max}}(v - \underline{v}_s)\right) - \left(P_+ \cdot \frac{f_s^{\max}}{f_s^{\min}} \cdot \frac{1}{l}\right). \tag{8}$$

Thus, the *total ex ante reduction* in information rents, denoted by $I_\Delta^\beta$, is greater than the sum of expressions (7) and (8) and, therefore,

$$I_\Delta^\beta > \frac{1}{2}P_\Delta \cdot \min\left\{\frac{f_b^{\min}}{f_b^{\max}}, \frac{f_s^{\min}}{f_s^{\max}}\right\} \cdot \left(\bar{v}_b - \underline{v}_s - \frac{1}{l}\right) - P_+ \cdot \max\left\{\frac{f_b^{\max}}{f_b^{\min}}, \frac{f_s^{\max}}{f_s^{\min}}\right\} \cdot \frac{2}{l}.$$

Since $P_\Delta > P^+$ and since for sufficiently large $l$, we have that $(\bar{v}_b - \underline{v}_s - \frac{1}{l}) > (\bar{v}_b - \underline{v}_s)/2$, then

$$I_\Delta^\beta > \frac{1}{2}P_\Delta \cdot \left[\min\left\{\frac{f_b^{\min}}{f_b^{\max}}, \frac{f_s^{\min}}{f_s^{\max}}\right\} \cdot \frac{\bar{v}_b - \underline{v}_s}{2} - \max\left\{\frac{f_b^{\max}}{f_b^{\min}}, \frac{f_s^{\max}}{f_s^{\min}}\right\} \cdot \frac{4}{l}\right].$$

The lost surplus due to the modification in the box ($s^\beta$) cannot exceed $\frac{1}{l}(P_+ + P_-)$: at the worst case, the reduced trade $P_-$ occurs for type-pairs with a positive surplus, which cannot exceed $1/l$, and the addition to trade $P_+$ occurs for type-pairs with a negative surplus, which cannot be less than $-1/l$. Since $(P_+ + P_-) < 3P_\Delta$, we have that $s^\beta < 3P_\Delta/l$ and, therefore,

$$I_\Delta^\beta > \frac{s^\beta \cdot l}{6} \cdot \left[\min\left(\frac{f_b^{\min}}{f_b^{\max}}, \frac{f_s^{\min}}{f_s^{\max}}\right) \cdot \frac{\bar{v}_b - \underline{v}_s}{2} - \max\left\{\frac{f_b^{\max}}{f_b^{\min}}, \frac{f_s^{\max}}{f_s^{\min}}\right\} \cdot \frac{4}{l}\right].$$

[45] To see this, let the function $\xi(x, y)$ equal 1 if, in the modification described in step II, the type-pair $(x, y)$ started trading and 0 otherwise. Ignore all type pairs for which the modification reduced trade. Then the *change* in the information rent paid to type $v_b$ due to the modification is no more than $\int_{\underline{v}_b}^{v_b} \int_{\underline{v}_s}^{\bar{v}_s} \xi(x, y) f_s(y)\,dy\,dx < (1/f_{\min}^b) \int_{\underline{v}_b}^{\bar{v}_b} \int_{\underline{v}_s}^{\bar{v}_s} \xi(x, y) f_s(y) f_b(x)\,dy\,dx = (1/f_{\min}^b) \cdot P_+$.

The *reduction* in $\psi$ due to the modification in the box, denoted by $\psi_\Delta^\beta$, equals to the reduction in the required information rents ($I_\Delta^\beta$) minus the reduction in the generated surplus ($s^\beta$), i.e., $\psi_\Delta^\beta = I_\Delta^\beta - s^\beta$. Therefore,

$$\frac{\psi_\Delta^\beta}{s^\beta} = \frac{I_\Delta^\beta - s^\beta}{s^\beta} > \frac{l}{6} \cdot \left[\min\left(\frac{f_b^{\min}}{f_b^{\max}}, \frac{f_s^{\min}}{f_s^{\max}}\right) \cdot \frac{\bar{v}_b - \underline{v}_s}{2} - \max\left\{\frac{f_b^{\max}}{f_b^{\min}}, \frac{f_s^{\max}}{f_s^{\min}}\right\} \cdot \frac{4}{l}\right] - 1. \tag{9}$$

The right-hand side of equation (9) is unboundedly increasing in $l$. Our fourth and last lower bound on the number of boxes $l$ guarantees that the ratio between the reduction of $\psi$ and the reduction of surplus is larger than $r$:

*Lower bound 4.* $\hat{l}_4$ is the lowest integer for which the right-hand side of equation (9) is greater than $r$.

*In sum*, given $r$ and $\bar{s}_2$, let $\hat{l}$ be the lowest integer that is greater than $\max\{\hat{l}_1, \hat{l}_2, \hat{l}_3, \hat{l}_4\}$. Then define

$$\bar{K} = 5\hat{l} + 2(\hat{l}^2 - \hat{l} + 2).$$

Given any intermediation mechanism with more that $\bar{K}$ extreme vertices, we apply the modification described in step II above. The resulting intermediation mechanism has no more than $\bar{K}$ extreme vertices (the computation of $\bar{K}$ is explained at the beginning of this lemma's proof). Compared to the original mechanism, the surplus decreases by at most $s < \bar{s}_2$ and the budget deficit decreases by at least $s/r$.

*Proof of Lemma A.4*

A budget-balanced intermediation mechanism with (weakly) fewer than $K$ extreme vertices is essentially a vector of $2K$ elements $(c_b^1, c_s^1, \dots, c_b^K, c_s^K)$ that characterize the boundaries of trade and satisfy the following three conditions:

(A) $c_b^1 \le c_b^2 \le \dots \le c_b^K$ and $c_s^1 \le c_s^2 \le \dots \le c_s^K$,

(B) Credibility: for every $l = \{0, \dots, K\}$ such that $c_s^l \neq c_s^{l+1}$:

  (i) if $c_b^l \neq c_b^{l+1}$ then $\mathbb{E}_{F_s}[c_s^l, c_s^{l+1}] \ge \mathbb{E}_{F_b}[c_b^l, c_b^{l+1}]$, and

  (ii) if $c_b^{l+1} \neq c_b^{l+2}$ then $\mathbb{E}_{F_s}[c_s^l, c_s^{l+1}] \le \mathbb{E}_{F_b}[c_b^{l+1}, c_b^{l+2}]$,

  where $(c_b^0, c_s^0) \equiv (\underline{v}_b, \underline{v}_s)$ and $(c_b^{K+1}, c_s^{K+1}) \equiv (\bar{v}_b, \bar{v}_s)$.

(C) Budget balance: $\psi = \sum_{k=1,\dots,K} \sum_{l=1,\dots,k} (c_s^k - c_b^l) \cdot (F_b(c_b^{k+1}) - F_b(c_b^k)) \cdot (F_s(c_s^l) - F_s(c_s^{l-1})) = 0$.

The surplus of a mechanism $(c_b^1, c_s^1, \dots, c_b^K, c_s^K)$ is computed as follows:

$$\sum_{k=1,\dots,K} \sum_{l=1,\dots,k} \left(\mathbb{E}_{F_b}[c_b^k, c_b^{k+1}] - \mathbb{E}_{F_s}[c_s^{l-1}, c_s^l]\right) \cdot \left(F_b(c_b^{k+1}) - F_b(c_b^k)\right) \cdot \left(F_s(c_s^l) - F_s(c_s^{l-1})\right).$$

Define the set of all vectors with $2K$ elements that satisfy the above conditions as "the feasible set," and denote it by $\mathcal{M}_K$. Endow $\mathcal{M}_K$ with the Euclidean metric. Clearly, the surplus function is continuous and $\mathcal{M}_K$ is compact (since all the constraints are given by equalities and weak inequalities on continuous functions). Therefore, there exists a budget-balanced intermediation mechanism with fewer than $K$ extreme vertices that attains the supremum surplus over $\mathcal{M}_K$.

### *Proof of Lemma A.5*

We prove only the first part of the lemma, since the second part is analogous. Given a distribution $F_i$, and three points $x, y, z$ that satisfy $\underline{v}_i < x < y < z < \bar{v}_i$, we begin by showing that

$$\mathbb{E}[x, z] = \mathbb{E}[x, y] + \frac{1}{2}(z - y) + O\big((z - y)(y - x)\big) + O(z - y)^2 \tag{10}$$

when $z \to x$. To see this, denote $h(x, y) \equiv d\mathbb{E}[x, y]/dy = f(y)[y(F(y) - F(x)) - \int_x^y tf(t)\,dt] /(F(y) - F(x))^2$. Sequentially applying L'Hospital's rule yields $\lim_{y\to x} h(x, y) = \frac{1}{2}$ and $\lim_{y\to x}(h(x, y) - \frac{1}{2})/(y - x) = -f'(x)/(6f(x))$. Therefore, $h(x, y) = \frac{1}{2} + O(y - x)$ as $y \to x$. Now write the Taylor expansion of $\mathbb{E}[x, z]$ at $(x, y)$ to obtain

$$\mathbb{E}[x, z] = \mathbb{E}[x, y] + (z - y) \cdot h(x, y) + O(z - y)^2.$$

Plugging in $h(x, y) = \frac{1}{2} + O(y - x)$ yields equation (10).

Using equation (10), we can write the buyer mean on the interval $[c_b^1, c_b^3]$ as follows:

$$\mathbb{E}_{F_b}\big[c_b^1, c_b^3\big] = \mathbb{E}_{F_b}\big[c_b^1, c_b^2\big] + \frac{1}{2}\big(c_b^3 - c_b^2\big) + O\big(\big(c_b^3 - c_b^2\big)\big(c_b^2 - c_b^1\big)\big) + O\big(c_b^3 - c_b^2\big)^2,$$

where the subscript $F_b$ in the expectations operator indicates the distribution according to which the mean is evaluated.

Since $(c_s^3 - c_s^2)/(c_b^3 - c_b^2) > 1 + \kappa$, then $c_s^3 > c_s^2 + (1 + \kappa)(c_b^3 - c_b^2) > c_s^1$ and, therefore, $\mathbb{E}_{F_s}[c_s^1, c_s^3] > \mathbb{E}_{F_s}[c_s^1, c_s^2 + (1 + \kappa)(c_b^3 - c_b^2)]$. Using equation (10) again, we can write

$$\begin{aligned}\mathbb{E}_{F_s}\big[c_s^1, c_s^3\big] &> \mathbb{E}_{F_s}\big[c_s^1, c_s^2 + (1 + \kappa)\big(c_b^3 - c_b^2\big)\big] \\ &= \mathbb{E}_{F_s}\big[c_s^1, c_s^2\big] + \frac{1}{2}(1 + \kappa)\big(c_b^3 - c_b^2\big) + O\big(\big(c_s^2 - c_s^1\big)\big(c_b^3 - c_b^2\big)\big) + O\big(c_b^3 - c_b^2\big)^2.\end{aligned}$$

Therefore,

$$\begin{aligned}&\mathbb{E}_{F_s}\big[c_s^1, c_s^3\big] - \mathbb{E}_{F_b}\big[c_b^1, c_b^3\big] \\ &\quad > \big(\mathbb{E}_{F_s}\big[c_s^1, c_s^2\big] - \mathbb{E}_{F_b}\big[c_b^1, c_b^2\big]\big) + \frac{1}{2}\kappa\big(c_b^3 - c_b^2\big) \\ &\qquad + O\big(\big(c_s^2 - c_s^1\big)\big(c_b^3 - c_b^2\big)\big) + O\big(\big(c_b^3 - c_b^2\big)\big(c_b^2 - c_b^1\big)\big) + O\big(c_b^3 - c_b^2\big)^2\end{aligned}$$

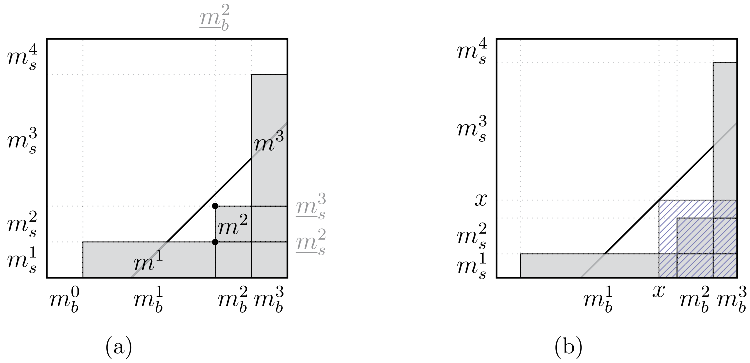


FIGURE 10. The left-hand panel shows an allocation rule that satisfies the conditions of Lemma A.8. The right-hand panel shows a posted price that generates a higher social surplus.

and since $\mathbb{E}_{F_s}[c_s^1, c_s^2] \geq \mathbb{E}_{F_b}[c_b^1, c_b^2]$ we have that

$$\begin{aligned}
&\mathbb{E}_{F_s}\big[c_s^1, c_s^3\big] - \mathbb{E}_{F_b}\big[c_b^1, c_b^3\big] \\
&\quad > \big(c_b^3 - c_b^2\big) \cdot \left[\frac{\kappa}{2} + O\big(c_s^2 - c_s^1\big) + O\big(c_b^2 - c_b^1\big) + O\big(c_b^3 - c_b^2\big)\right].
\end{aligned} \tag{11}$$

Recall that the three extreme vertices $c^1$, $c^2$, and $c^3$ are all within the same box, i.e., $(c_b^3 - c_b^2)$, $(c_b^2 - c_b^1)$, and $(c_s^2 - c_s^1)$ all go to zero as $l$ grows. Thus, given $\kappa$, there exists $\tilde{l}$ such that for any $l > \tilde{l}$ the left-hand side of equation (11) is positive.

*Proof of Proposition 6*

The proof consists of two parts. In part A, we show that when the type distributions are uniform the intermediary implements the optimal posted price $x^*$ as described in the statement of the proposition. In part B, we show that when the distributions are close to uniform, the intermediary implements an outcome that is close to that of the optimal posted price $x^*$.

*Part A.* Suppose that the type distributions are uniform. By Proposition 4, the buyer's and seller's message sets in the optimal intermediation mechanism are finite. Given a partition-direct intermediation mechanism $\Gamma = (M, p, t_b, t_s)$ with finite message sets, we order the seller's messages from lowest to highest and denote the $k$th message by $m_s^k$. We denote the lowest buyer message that trades with $m_s^1$ by $m_b^1$, and denote by $m_b^k$ the $k$th buyer message above it. The index of the highest buyer message is denoted by $K$ (for an illustration see Figure 10(a), in which $K = 3$). Thus, for any $k \in \{1, \ldots, K\}$, the message pair $(m_b^k, m_s^k)$ is in the frontier of trade, i.e., $m_s^k$ is the highest seller message that trades with $m_b^k$ and $m_b^k$ is the lowest buyer message that trades with $m_s^k$.

Using the fact that the message sets are finite, we can slightly simplify the expression of $\psi(p)$. Note that for any pair of messages, $m_s^l$ and $m_b^k$, we have $p(m_s^l, m_b^k) = 1$ if and only if $k \geq l$. Moreover, if $k \geq l$ then $\omega_s(v_b) - \omega_b(v_s) = \bar{m}_s^k - \underline{m}_b^l$ for any type-pair $(v_b, v_s) \in m_b^k \times m_s^l$ (where that $\underline{m}_i$ and $\bar{m}_i$ are the lower and upper bounds of the interval $m_i$, resp.).

We then have $\psi(p)=\sum_{k=1}^{K}\sum_{l=1}^{k}\phi(m_s^l,m_b^k)$, where $\phi(m_s^l,m_b^k)=(\underline{m}_s^{k+1}-\underline{m}_b^l)\cdot((\underline{m}_b^{k+1}-\underline{m}_b^k)/(\bar{v}_b-\underline{v}_b))\cdot((\underline{m}_s^{l+1}-\underline{m}_s^l)/(\bar{v}_s-\underline{v}_s))$ (recall that $\bar{m}_i^l=\underline{m}_i^{l+1}$ for all $l$ and for every agent $i$).

To prove part A, we proceed in two steps. The first presents our core argument and shows that if $K>1$, then $\Gamma$ can satisfy budget balance *only if* it attains a very specific structure that is qualitatively illustrated in Figure 10(a) (and formally characterized by Lemma A.8). The second step shows that if $\Gamma$ attains this particular structure, then there exists a posted-price intermediation mechanism (i.e., an intermediation mechanism with $K=1$) that generates a higher social surplus.

We therefore deduce that in the optimal intermediation mechanism it must be the case that $K=1$, and since the optimal posted price generates the highest surplus among all intermediation mechanisms with $K=1$, the desired result is attained.

In what follows, we denote the length of an interval message $m_i$ by $|m_i|=\bar{m}_i-\underline{m}_i$.

*Step I: The core argument*

For any $k\in\{1,\dots,K\}$, we denote $m^k=(m_s^k,m_b^k)$ and refer to each $m^k$ as an *extreme rectangle* of trade. The rectangles $m^k$ and $m^{k+1}$ are referred to as *consecutive* extreme rectangles. We refer to $(\underline{m}_s^k,\underline{m}_b^k)$ as the *bottom-left corner* of $m^k$ and to $(\underline{m}_s^{k+1},\underline{m}_b^k)$ as the *top-left corner* of $m^k$ (note that this is an extreme vertex in the mechanism). For convenience, we denote $\underline{m}_b^{K+1}=\bar{v}_b$ and $\underline{m}_s^{\hat{K}+1}=\bar{v}_s$, where $\hat{K}\in\{K,K+1\}$ is the index of the seller's highest interval.[46] Finally, we define TLB to be the set of extreme rectangles with top-left corner below the equi-type line, and BLB to be the set of extreme rectangles with bottom-left corner below the equi-type line:

$$\mathrm{TLB}=\{(m_s^k,m_b^k)\in M:\underline{m}_s^{k+1}<\underline{m}_b^k\},$$
$$\mathrm{BLB}=\{(m_s^k,m_b^k)\in M:\underline{m}_s^{k}<\underline{m}_b^k\}.$$

In the example illustrated in Figure 10(a), the top-left and bottom-left corners of the extreme rectangle $m^2$ are marked with small black dots. Note that in this example the extreme rectangle $m^2$ belongs to the set TLB, and the extreme rectangle $m^1$ does *not* belong to the set BLB.

The first lemma establishes that, due to credibility, there are no two consecutive extreme rectangles with their top-left corner below the equi-type line and no two consecutive extreme rectangles with their bottom-left corner above the equi-type line.

Lemma A.6. *For any $k\in\{1,\dots,K-1\}$:*

*(i) If $m^k\in\mathrm{TLB}$, then $m^{k+1}\notin\mathrm{TLB}$;*

*(ii) If $m^k\notin\mathrm{BLB}$, then $m^{k+1}\in\mathrm{BLB}$.*

If $K>1$ and $\psi(p)=0$, then there is at least one extreme rectangle $m^k$ that belongs to the set TLB. To see why, note that if $K>1$ then there is more than one rectangle

[46] If the index of the buyer's highest interval is $K$, then the index of the seller's highest interval is either $K$ or $K+1$.

with trade. Since only rectangles in the set TLB contribute negative summands to the computation of $\psi$, and since $\psi(p)=0$, then there must be at least one extreme vertex that belongs to TLB. The next lemma, however, asserts that if $m^k \in$ TLB but there is also a rectangle $m^{k-j}$ or a rectangle $m^{k+j+1}$, for some $j \geq 1$, with their bottom-left corner below the equi-type line, then there are positive terms in $\psi$ that outweigh the negative value that is contributed by $\phi(m^k)$. The geometric interpretation of this result is outlined in the body of the text.

Lemma A.7. *For any* $k \in \{1, \dots, K\}$ *and any* $m^k \in$ TLB*:*

*(i) If* $m^{k-j} \in$ BLB *for some* $1 \leq j < k-1$*, then* $\phi(m_s^k, m_b^k) + \sum_{z=1}^{j} \phi(m_s^{k-z}, m_b^k) > 0$.

*(ii) If* $m^{k+j+1} \in$ BLB *for some* $1 \leq j \leq K-k$*, then* $\phi(m_s^k, m_b^k) + \sum_{z=1}^{j} \phi(m_s^k, m_b^{k+z}) > 0$.

The third lemma characterizes the structure of any intermediation mechanism for $K > 1$ and $\psi(p) \leq 0$. The intuition is as follows: Unless the mechanism satisfies the specific structure (which is qualitatively illustrated in Figure 10(a)), then *every* rectangle $m^k \in$ TLB can be "associated" with a *distinct* set of rectangles in which trade takes place, such that the sum of $\phi$ over $m^k = (m_s^k, m_b^k)$ and the other elements of the set is *positive*. This would be a contradiction of $\psi(p) \leq 0$.

Lemma A.8. *If* $\Gamma = (M, p, t)$ *is an intermediation mechanism, then* $K \leq 3$*. Furthermore, if* $K > 1$ *it must be that either:*

*(i)* $K = 2$ *and* $m^1 \in$ TLB *and* $m^3 \notin$ BLB,[47] *or*

*(ii)* $K = 2$ *and* $m^2 \in$ TLB *and* $m^1 \notin$ BLB*, or*

*(iii)* $K = 3$ *and* $m^2 \in$ TLB *and* $m^1 \notin$ BLB *and* $m^4 \notin$ BLB.

*Step II: A posted price is better than an intermediation mechanism with* $K > 1$.

We will show that for any intermediation mechanism with $K > 1$, there exists another intermediation mechanism which implements some posted price that generates a higher expected social surplus.

Suppose, by way of contradiction, that $\Gamma$ is an *optimal* intermediation mechanism with $K > 1$. Since $\Gamma$ is an intermediation mechanism, it must satisfy the conditions of Lemma A.8. Consider the case in which $K = 3$ and $m^2 \in$ TLB and $m^1 \notin$ BLB and $m^4 \notin$ BLB which is illustrated in Figure 10(a).[48] The proof for the other two cases with $K = 2$ is similar.

Since $\Gamma$ is optimal it must be that

$$\underline{m}_b^1 + \underline{m}_b^2 = \underline{m}_s^2 + \underline{m}_s^3 \tag{12}$$

[47] Recall that when $K = 2$, then by definition, $\underline{m}_b^3 = \bar{v}_b$. Therefore, if the seller has 3 intervals then $m^3 \notin$ BLB whenever $\underline{m}_s^3 > \bar{v}_b$, and if the seller has only 2 intervals, then $m^3 \notin$ BLB whenever $\bar{v}_s > \bar{v}_b$.

[48] Since according to our notation $\underline{m}_b^4 = \bar{v}_b$, then $m^4 \notin$ BLB is equivalent to $\underline{m}_s^4 > \bar{v}_b$.

and

$$\underline{m}_b^2 + \underline{m}_b^3 = \underline{m}_s^3 + \underline{m}_s^4. \tag{13}$$

This is because if $\underline{m}_b^1 + \underline{m}_b^2 < \underline{m}_s^2 + \underline{m}_s^3$, then slightly increasing $\underline{m}_1^b$ does not violate credibility and increases the expected surplus since it eliminates nonbeneficial trade.[49] Similarly, if $\underline{m}_b^2 + \underline{m}_b^3 < \underline{m}_s^3 + \underline{m}_s^4$, then slightly decreasing $\underline{m}_s^4$ increases the expected surplus without violating credibility.

Suppose for now that $\underline{m}_s^3 + |m_b^3| < \underline{m}_b^2 - |m_s^1|$ (we will prove below that this must be true). The expected surplus of any posted price $x \in [\underline{m}_s^3 + |m_b^3|, \underline{m}_b^2 - |m_s^1|]$, which is illustrated in Figure 10(b), is then given by $S(x) \equiv ((x - \underline{v}_s)/(\bar{v}_s - \underline{v}_s)) \cdot ((\bar{v}_b - x)/(\bar{v}_b - \underline{v}_b)) \cdot ((\bar{v}_b + x)/2 - (x + \underline{v}_s)/2)$, which can be alternatively written as a sum of four terms:

$$\begin{aligned} S(x) = {} & \frac{\underline{m}_s^3 - \underline{v}_s}{\bar{v}_s - \underline{v}_s} \cdot \frac{\bar{v}_b - \underline{m}_b^2}{\bar{v}_b - \underline{v}_b} \cdot \left( \frac{\bar{v}_b + \underline{m}_b^2}{2} - \frac{\underline{m}_s^3 + \underline{v}_s}{2} \right) \\ & + \frac{x - \underline{m}_s^3}{\bar{v}_s - \underline{v}_s} \cdot \frac{\bar{v}_b - \underline{m}_b^2}{\bar{v}_b - \underline{v}_b} \cdot \left( \frac{\bar{v}_b + \underline{m}_b^2}{2} - \frac{x + \underline{m}_s^3}{2} \right) \\ & + \frac{\underline{m}_s^3 - \underline{v}_s}{\bar{v}_s - \underline{v}_s} \cdot \frac{\underline{m}_b^2 - x}{\bar{v}_b - \underline{v}_b} \cdot \left( \frac{\underline{m}_b^2 + x}{2} - \frac{\underline{m}_s^3 + \underline{v}_s}{2} \right) \\ & + \frac{x - \underline{m}_s^3}{\bar{v}_s - \underline{v}_s} \cdot \frac{\underline{m}_b^2 - x}{\bar{v}_b - \underline{v}_b} \cdot \left( \frac{\underline{m}_b^2 + x}{2} - \frac{x + \underline{m}_s^3}{2} \right) \end{aligned}$$

There are four positive summands in the right-hand side of the equation. The first equals the expected social surplus generated when the buyer types in $m_b^2 \cup m_b^3 = [\underline{m}_b^2, \bar{v}_b]$ trade with the seller types in $m_s^1 \cup m_s^2 = [\underline{v}_s, \underline{m}_s^3]$ in the mechanism $\Gamma$. The second summand is *weakly greater* than the expected social surplus generated when buyer types in $m_b^3$ trade with seller types in $m_s^3$ in the mechanism $\Gamma$.[50] Similarly, the third summand is *weakly greater* than the expected social surplus generated when buyer types in $m_b^1$ trade with seller types in $m_s^1$ in the mechanism $\Gamma$.

Thus, the sum of the first three arguments of $S(x)$ is (weakly) greater than the total surplus generated by the intermediation mechanism $\Gamma$. Since the fourth argument is also positive, then the total expected surplus generated by any posted price $x \in [\underline{m}_s^3 + |m_b^3|, \underline{m}_b^2 - |m_s^1|]$ is strictly greater than that of the intermediation mechanism $\Gamma$.

[49] Clearly, it cannot be that $\underline{m}_b^1 + \underline{m}_b^2 > \underline{m}_s^2 + \underline{m}_s^3$, because then $p(m_s^2, m_b^1) = 0$, thus contradicting credibility. Similarly, it cannot be that $\underline{m}_b^2 + \underline{m}_b^3 > \underline{m}_s^3 + \underline{m}_s^4$ because then $p(m_s^3, m_b^2) = 0$ contradicts credibility.

[50] To see this, note that the second summand is increasing in $x$ in the range $x \in [\underline{m}_s^3 + |m_b^3|, \underline{m}_b^2 - |m_s^1|]$, and is thus (weakly) greater than $\frac{(\underline{m}_s^3 + |m_b^3|) - \underline{m}_s^3}{\bar{v}_s - \underline{v}_s} \cdot \frac{\bar{v}_b - \underline{m}_b^2}{\bar{v}_b - \underline{v}_b} \cdot (\frac{\bar{v}_b + \underline{m}_b^2}{2} - \frac{(\underline{m}_s^3 + |m_b^3|) + \underline{m}_s^3}{2})$, which is equal to $\frac{|m_s^3|}{\bar{v}_s - \underline{v}_s} \cdot \frac{|m_b^3|}{\bar{v}_b - \underline{v}_b} \cdot (\frac{|m_b^2| + |m_b^3|}{2})$ by equation (13) and since $\bar{v}_b = \bar{m}_b^3$. The expected surplus generated by the trade between buyers in $m_b^3$ and sellers in $m_s^3$ is $\frac{|m_b^3|}{\bar{v}_b - \underline{v}_b} \cdot \frac{|m_s^3|}{\bar{v}_s - \underline{v}_s} \cdot (\frac{\underline{m}_b^3 + \bar{v}_b}{2} - \frac{\underline{m}_s^3 + \underline{m}_s^4}{2})$, which is equal to $\frac{|m_b^3|}{\bar{v}_b - \underline{v}_b} \cdot \frac{|m_s^3|}{\bar{v}_s - \underline{v}_s} \cdot \frac{|m_b^2| + |m_b^3|}{2}$ by equation (13).

It remains to show that $\underline{m}_s^3 + |m_b^3| < \underline{m}_b^2 - |m_s^1|$. To do so, suppose by way of contradiction that $\underline{m}_b^2 - \underline{m}_s^3 \leq |m_s^1| + |m_b^3|$. Then

$$-\phi(m_s^2, m_b^2) = (\underline{m}_b^2 - \underline{m}_s^3) \cdot \frac{|m_b^2|}{\bar{v}_b - \underline{v}_b} \cdot \frac{|m_s^2|}{\bar{v}_s - \underline{v}_s} \leq (|m_s^1| + |m_b^3|) \cdot \frac{|m_b^2|}{\bar{v}_b - \underline{v}_b} \cdot \frac{|m_s^2|}{\bar{v}_s - \underline{v}_s}$$
$$< \frac{|m_s^1|}{\bar{v}_s - \underline{v}_s} \cdot \frac{|m_b^2|}{\bar{v}_b - \underline{v}_b} \cdot (\underline{m}_b^2 - \underline{m}_s^2) + \frac{|m_b^3|}{\bar{v}_b - \underline{v}_b} \cdot (\underline{m}_b^3 - \underline{m}_s^3) \cdot \frac{|m_s^2|}{\bar{v}_s - \underline{v}_s},$$

where the second inequality follows from $m^2 \in \text{TLB}$ (and therefore $\underline{m}_s^3 < \underline{m}_b^2$). Plugging in (12) and (13), we obtain:[51]

$$-\phi(m_s^2, m_b^2) < \frac{|m_s^1|}{\bar{v}_s - \underline{v}_s} \cdot \frac{|m_b^2|}{\bar{v}_b - \underline{v}_b} \cdot (\underline{m}_s^3 - \underline{m}_b^1) + \frac{|m_b^3|}{\bar{v}_b - \underline{v}_b} \cdot (\underline{m}_s^4 - \underline{m}_b^2) \cdot \frac{|m_s^2|}{\bar{v}_s - \underline{v}_s}$$
$$= \phi(m_s^1, m_b^2) + \phi(m_s^2, m_b^3)$$

Thus, $\phi(m_s^1, m_b^2) + \phi(m_s^2, m_b^3) + \phi(m_s^2, m_b^2) > 0$, which implies $\psi(p) > 0$, which contradicts $\Gamma = (M, p, t_b, t_s)$ being an intermediation mechanism.

*In sum*, in the first step, we showed that if $F_b$ and $F_s$ are uniform, then any intermediation mechanism with $K > 1$ must attain a very specific structure in order not to violate the budget-balance requirement. In the second step, we showed that for any intermediation mechanism with $K > 1$ that attains this structure there exists a posted price that yields a higher social surplus. By Proposition 5, this posted price can be implemented by an intermediation mechanism (with $K = 1$). Furthermore, since any intermediation mechanism with $K = 1$ and $\psi(p) = 0$ necessarily implements some posted price, it follows that the optimal intermediation mechanism implements the optimal posted price (i.e., the price that maximizes the social surplus). This price, $x^*$, is the maximizer of $\int_x^{\bar{v}_b} \int_{\underline{v}_s}^x (v_b - v_s)\, dv_s\, dv_b$ among all $x \in V_b \cap V_s$, and is characterized in the statement of the proposition.

*Part B.* We now turn to show that when the distributions are close to uniform, then the intermediary implements an outcome that is close to that of the optimal posted price $x^*$.

Define the distance between two type distributions $F = \{F_b, F_s\}$ and $F' = \{F'_b, F'_s\}$ as $\max_{i \in \{b,s\}} \max_{v_i \in V_i} |f_i(v_i) - f'(v_i)|$. Consider a sequence of distributions $\{F_n\}$ that converges to the uniform distribution. Since in the uniform distribution $f_i^{\max} = f_i^{\min} = 1$ for each agent $i$ (where $f_i^{\max}$ and $f_i^{\min}$ are the maximum and minimum of the probability density function $f_i$), then we assume without loss that for each $F_n$ we have $f_i^{\max} < 2$ and $f_i^{\min} > 0.5$. Thus, by the proof of Proposition 4, there exists a finite $K$ such that, for all $n$, the optimal intermediation mechanism $\Gamma_n^*$ under $F_n$ has no more than $K$ extreme vertices. We thus identify each mechanism $\Gamma_n^*$ with a $K$-pairs-vector of the coordinates of its $K$ extreme vertices (where if the mechanism has fewer than $K$ extreme vertices we pad the vector by replicating the last pair until it has $K$-pairs, and $(\underline{v}_b, \underline{v}_s)$ and $(\bar{v}_b, \bar{v}_s)$

[51] Note that $\underline{m}_b^2 - \bar{m}_s^1 = \bar{m}_s^2 - \underline{m}_b^1$ and $\bar{m}_b^2 - \bar{m}_s^2 = \bar{m}_s^3 - \underline{m}_b^2$.

are considered as extreme vertices). We endow the space of such $K$-pairs-vectors with the Euclidean metric.[52] Denote the surplus generated by $\Gamma_n^*$ by $s_n^*$.

Consider *any* convergent subsequence of $\{\Gamma_n^*\}$, and denote its limit by $\Gamma_\infty^*$. We show that $\Gamma_\infty^*$ is the optimal posted price mechanism under the uniform distribution.[53] Since any convergent subsequence of $\{\Gamma_n^*\}$ converges to the same mechanism, so does the sequence $\{\Gamma_n^*\}$.

We first show that $\Gamma_\infty^*$ is an intermediation mechanism under the uniform distribution (i.e., satisfies credibility and budget balance), and its surplus $s_\infty^*$ is the limit of surpluses $s_n^*$ of the mechanisms along the subsequence. To see this, note first that the credibility condition involves a finite set of weak inequalities between conditional expectations, each of which is computed over intervals for the buyer and seller induced by the $K$-pairs-vector, and thus continuous in the $K$-pairs-vector and in the distribution. Thus, credibility is satisfied also in the limit. Next, note that the function $\psi$ (defined in equation (2)) is continuous in the same variables, so having $\psi = 0$ for all the mechanisms along the subsequence implies $\psi = 0$, i.e., budget balance is satisfied, also in the limit. Finally, the ex ante surplus is also continuous in the same variables so the surpluses $s_n^*$ of the mechanisms along the subsequence converge to $s_\infty^*$.

Denote by $s_n^{pp}$ the surplus generated by the optimal posted price under $F_n$. Since $s_n^* \geq s_n^{pp}$ (by Proposition 5) and $s_n^{pp} \to s_U^{pp}$ (where $s_U^{pp}$ is the surplus of the optimal posted price mechanism under the uniform distribution) then $s_\infty^* \geq s_U^{pp}$. But, since $\Gamma_\infty^*$ is an intermediation mechanism under the uniform distribution, then Part A of the proof precludes $s_\infty^* > s_U^{pp}$. Therefore, $s_\infty^* = s_U^{pp}$. And, since the optimal posted price mechanism is the unique optimal intermediation mechanism under the uniform distribution (by Part A of the proof), then it must be that $\Gamma_\infty^*$ is that mechanism.

### *Proof of Lemma A.6*

First, suppose that $m^k \in \text{TLB}$ and $m^{k+1} \in \text{TLB}$ for some $k \in \{1, \dots, K-1\}$. Then, by definition, $\underline{m}_s^{k+1} < \underline{m}_b^k$ and $\underline{m}_s^{k+2} < \underline{m}_b^{k+1}$ and, therefore, $\mathbb{E}[v_b : v_b \in m_b^k] > \mathbb{E}[v_s : v_s \in m_s^{k+1}]$. This implies that trade is beneficial when the buyer reports $m_b^k$ and the seller reports $m_s^{k+1}$, which is inconsistent with $p(m_s^{k+1}, m_b^k) = 0$ due to credibility.

Next, suppose that $m^k \notin \text{BLB}$ and $m^{k+1} \notin \text{BLB}$ for some $k \in \{1, \dots, K-1\}$. Then, by definition, $\underline{m}_s^k \geq \underline{m}_b^k$ and $\underline{m}_s^{k+1} \geq \underline{m}_b^{k+1}$ and, therefore, $\mathbb{E}[v_b : v_b \in m_b^k] \leq \mathbb{E}[v_s : v_s \in m_s^k]$. This implies that trade is not beneficial when the buyer reports $m_b^k$ and the seller reports $m_s^k$, which is inconsistent with $p(m_s^k, m_b^k) = 1$ due to credibility.

### *Proof of Lemma A.7*

Suppose $m^k \in \text{TLB}$ and $m^{k-j} \in \text{BLB}$ for some $1 \leq j < k-1$. Then

$$\phi(m_s^k, m_b^k) = (\underline{m}_s^{k+1} - \underline{m}_b^k) \cdot \frac{|m_b^k|}{\bar{v}_b - \underline{v}_b} \cdot \frac{|m_s^k|}{\bar{v}_s - \underline{v}_s} < 0$$

[52]More precisely, given two $K$-pairs vectors, $v = ((v_b^1, v_s^1), \dots (v_b^K, v_s^K))$ and $w = ((w_b^1, w_s^1), \dots (w_b^K, w_s^K))$, we define the distance between $v$ and $w$ to be $\sqrt{(v_b^1 - w_b^1)^2 + (v_s^1 - w_s^1)^2 + \cdots + (v_b^K - w_b^K)^2 + (v_s^K - w_s^K)^2}$.

[53]A mechanism that is represented by a $K$-pairs-vector is said to be the optimal posted price if its elements are $(\underline{v}_b, \underline{v}_s)$, $(x^*, x^*)$ and $(\bar{v}_b, \bar{v}_s)$ (perhaps replicated).

and

$$\sum_{z=1}^{j}\phi\left(m_s^{k-z}, m_b^k\right) = \sum_{z=1}^{j}\left(\underline{m}_s^{k+1} - \underline{m}_b^{k-z}\right) \cdot \frac{\left|m_b^k\right|}{\bar{v}_b - \underline{v}_b} \cdot \frac{\left|m_s^{k-z}\right|}{\bar{v}_s - \underline{v}_s}$$
$$\geq \left(\underline{m}_s^{k+1} - \underline{m}_b^{k-1}\right) \cdot \frac{\left|m_b^k\right|}{\bar{v}_b - \underline{v}_b} \cdot \frac{\underline{m}_s^k - \underline{m}_s^{k-j}}{\bar{v}_s - \underline{v}_s},$$

where the inequality follows from the fact that $\underline{m}_b^{k-z} \leq \underline{m}_b^{k-1}$ for every $z \geq 1$. We will now show that $(\underline{m}_s^{k+1} - \underline{m}_b^{k-1}) \cdot (\underline{m}_s^k - \underline{m}_s^{k-j}) > -(\underline{m}_s^{k+1} - \underline{m}_b^k) \cdot (\underline{m}_s^{k+1} - \underline{m}_s^k)$.

To see this, note first that $\underline{m}_s^{k+1} - \underline{m}_b^{k-1} > \underline{m}_s^{k+1} - \underline{m}_s^k$, because $m^k \in \text{TLB}$ implies $m^{k-1} \notin \text{TLB}$ (according to Lemma A.6) and, therefore, $\underline{m}_b^{k-1} < \underline{m}_s^k$. Next, note that $\underline{m}_s^k - \underline{m}_s^{k-j} > \underline{m}_b^k - \underline{m}_s^{k+1}$, because $p(m_s^k, m_b^{k-1}) = 0$ and, therefore, by credibility, $\underline{m}_s^k - \underline{m}_b^{k-1} \geq \underline{m}_b^k - \underline{m}_s^{k+1}$, and since $m^{k-j} \in \text{BLB}$ then $\underline{m}_s^{k-j} < \underline{m}_b^{k-j} \leq \underline{m}_b^{k-1}$ and, therefore, $\underline{m}_s^k - \underline{m}_s^{k-j} > \underline{m}_b^k - \underline{m}_s^{k+1}$. This completes the proof for the first part of the lemma. The proof of the the second part is similar and, therefore, omitted.

## *Proof of Lemma A.8*

We divide the analysis into three cases:

*Case I*: Suppose that $K > 3$. For any extreme rectangle $m^k \in \text{TLB}$ with index $k \geq 3$, we associate $m^k$ with the rectangle $(m_s^{k-1}, m_b^k)$ if $m^{k-1} \in \text{BLB}$, and with the rectangles $(m_s^{k-1}, m_b^k)$ and $(m_s^{k-2}, m_b^k)$ otherwise (note that according to Lemma A.6 if $m^{k-1} \notin \text{BLB}$, then it must be that $m^{k-2} \in \text{BLB}$). For any extreme rectangle $m^k \in \text{TLB}$ with index $k \leq 2$, we associate $m^k$ with the rectangle $(m_s^k, m_b^{k+1})$ if $m^{k+2} \in \text{BLB}$, and with the rectangles $(m_s^k, m_b^{k+1})$ and $(m_s^k, m_b^{k+2})$ otherwise (according to Lemma A.6 if $m^{k+2} \notin \text{BLB}$ then it must be that $m^{k+3} \in \text{BLB}$).[54] Since by Lemma A.6 there are no two consecutive extreme rectangles that belong to TLB, then we associate each extreme rectangle with a distinct group of rectangles. According to Lemma A.7, the sum of the function $\phi$ over the elements of each group is positive, and thus it must be that $\psi$ is strictly positive, which contradicts $\Gamma$ being an intermediation mechanism.

*Case II*: Suppose that $K = 2$. Since $\psi(p) \leq 0$, it must be that either $m^1 \in \text{TLB}$ or $m^2 \in \text{TLB}$, but not both (since according to Lemma A.6 there are no two consecutive rectangles with a top-left corner below the equi-type line). Suppose $m^1 \in \text{TLB}$. If $m^3 \in \text{BLB}$, then according to Lemma A.6 we have that $\phi(m_s^k, m_b^k) + \phi(m_s^k, m_b^{k+1}) > 0$, which contradicts $\psi(p) \leq 0$. Therefore, it must be that $m^3 \notin \text{BLB}$, or equivalently $\underline{m}_s^3 > \bar{v}_b$. The argument for the case in which $m^2 \in \text{TLB}$ is similar.

*Case III*: Suppose that $K = 3$. If either $m^1 \in \text{TLB}$ or $m^3 \in \text{TLB}$, then applying the argument of Case I shows that $\psi(p) > 0$. Suppose that $m^2 \in \text{TLB}$ (and therefore $m^1 \notin \text{TLB}$ or $m^3 \notin \text{TLB}$). If $m^1 \in \text{BLB}$, then $\phi(m_s^2, m_b^2) + \phi(m_s^1, m_b^2) > 0$, and if $m^4 \in \text{BLB}$, then $\phi(m_s^2, m_b^2) + \phi(m_s^2, m_b^3) > 0$, which both contradict $\psi(p) \leq 0$. It must therefore be that $m^1 \notin \text{BLB}$ and $m^4 \notin \text{BLB}$.

[54] Note that by definition $\underline{m}_b^{K+1} = \bar{v}_b$ and, therefore, $m^{K+1} \in \text{BLB}$ if and only if $\underline{m}_s^{K+1} < \bar{v}_b$.

## References

Akbarpour, Mohammad and Shengwu Li (2020), "Credible auctions: A trilemma." *Econometrica*, 88, 425–467. [1659]

Austen-Smith, David (1993), "Interested experts and policy advice: Multiple referrals under open rule." *Games and Economic Behavior*, 5, 3–43. [1658]

Ausubel, Lawrence M. and Raymond J. Deneckere (1989), "A direct mechanism characterization of sequential bargaining with one-sided incomplete information." *Journal of Economic Theory*, 48, 18–46. [1660]

Baliga, Sandeep, Luis C. Corchon, and Tomas Sjostrom (1997), "The theory of implementation when the planner is a player." *Journal of Economic Theory*, 77, 15–33. [1659]

Battaglini, Marco (2002), "Multiple referrals and multidimensional cheap talk." *Econometrica*, 70, 1379–1401. [1658]

Bénabou, Roland and Jean Tirole (2002), "Self-confidence and personal motivation." *The Quarterly Journal of Economics*, 117, 871–915. [1657]

Bester, Helmut and Roland Strausz (2000), "Imperfect commitment and the revelation principle: The multi-agent case." *Economics Letters*, 69, 165–171. [1659, 1663, 1665]

Bester, Helmut and Roland Strausz (2001), "Contracting with imperfect commitment and the revelation principle: The single agent case." *Econometrica*, 69, 1077–1098. [1657, 1659]

Bester, Helmut and Roland Strausz (2007), "Contracting with imperfect commitment and noisy communication." *Journal of Economic Theory*, 136, 236–259. [1657, 1659, 1680]

Blume, Andreas, Oliver J. Board, and Kohei Kawamura (2007), "Noisy talk." *Theoretical Economics*, 2, 395–440. [1658]

Börgers, Tilman (2015), *An Introduction to the Theory of Mechanism Design*. Oxford University Press. [1669, 1695]

Börgers, Tilman and Peter Norman (2009), "A note on budget balance under interim participation constraints: The case of independent types." *Economic Theory*, 39, 477–489. [1668, 1669]

Carrillo, Juan D. and Thomas Mariotti (2000), "Strategic ignorance as a self-disciplining device." *Review of Economic Studies*, 67, 529–544. [1657]

Chakravorty, Bhaskar, Luis C. Corchón, and Simon Wilkie (2006), "Credible implementation." *Games and Economic Behavior*, 57, 18–36. [1659]

Chatterjee, Kalyan and William Samuelson (1983), "Bargaining under incomplete information." *Operations Research*, 31, 835–851. [1680]

Crawford, Vincent and Joel Sobel (1982), "Strategic information transmission." *Econometrica*, 50, 1431–1451. [1658, 1666, 1672]

Cremer, Jacques (1995), "Arm's length relationships." *Quarterly Journal of Economics*, 110, 275–295. [1657]

Dewatripont, Mathias and Eric Maskin (1990), "Contract renegotiation in models of asymmetric information." *European Economic Review*, 34, 311–321. [1660]

Doval, Laura and Vasiliki Skreta (2020), "Mechanism design with limited commitment." Working paper. [1657, 1659, 1680]

Evans, Robert and Sonje Reiche (2008), "Imperfect commitment and the revelation principle: The multi-agent case with transferable utility." *Economics Letters*, 99, 611–614. [1657, 1659, 1663]

Fey, Mark and Kristopher Ramsay (2009), "Mechanism design goes to war: Peaceful outcomes with interdependent and correlated types." *Review of Economic Design*, 13, 233–250. [1658]

Forges, Francoise (1986), "An approach to communication equilibria." *Econometrica*, 54, 1375–1385. [1658]

Gerardi, Dino, Richard McLean, and Andrew Postlewaite (2008), "Aggregation of expert opinions." *Games and Economic Behavior*, 65, 339–371. [1658]

Gilligan, Thomas W. and Keith Krehbiel (1989), "Asymmetric information and legislative rules with a heterogeneous committee." *American Journal of Political Science*, 33, 459–490. [1658]

Golman, Russell, David Hagmann, and George Loewenstein (2017), "Information avoidance." *Journal of Economic Literature*, 55, 96–135. [1657]

Goltsman, Maria, Johannes Hörner, Gregory Pavlov, and Francesco Squintani (2009), "Mediation, arbitration and negotiation." *Journal of Economic Theory*, 144, 1397–1420. [1658]

Hörner, Johannes, Massimo Morelli, and Francesco Squintani (2015), "Mediation and peace." *The Review of Economic Studies*, 82, 1483–1501. [1658]

Ivanov, Maxim (2010), "Communication via a strategic mediator." *Journal of Economic Theory*, 145, 869–884. [1658]

Green Jerry R., and Laffont Jean-Jacques (1987), "Posterior implementability in a two-person decision problem." *Econometrica*, 55, 69–94. [1678]

Krishna, Vijay (2009), *Auction Theory*. Academic Press. [1687]

Krishna, Vijay and John Morgan (2001), "Asymmetric information and legislative rules: Some amendments." *American Political Science Review*, 95, 435–452. [1658]

Krishna, Vijay and John Morgan (2008), "Contracting for information under imperfect commitment." *RAND Journal of Economics*, 39, 905–925. [1657, 1658, 1666]

Laffont, Jean-Jacques and Jean Tirole (1988), "The dynamics of incentive contracts." *Econometrica*, 56, 1153–1175. [1659]

McAdams, David and Michael Schwarz (2007), "Credible sales mechanisms and intermediaries." *The American Economic Review*, 97, 260–276. [1659]

McAfee, Preston R. and Daniel Vincent (1997), "Sequentially optimal auctions." *Games and Economic Behavior*, 18, 246–276. [1659]

Myerson, Roger (1979), "Incentive compatibility and the bargaining problem." *Econometrica*, 47, 61–73. [1660]

Myerson, Roger (1986), "Multistage games with communication." *Econometrica*, 54, 323–358. [1658]

Myerson, Roger and Mark Satterthwaite (1983), "Efficient mechanisms for bilateral trading." *Journal of Economic Theory*, 29, 265–281. [1656, 1660, 1670, 1673]

Neeman, Zvika and Gregory Pavlov (2012), "Ex-post renegotiation-proof mechanism design." *Journal of Economic Theory*, 148, 473–501. [1660]

Skreta, Vasiliki (2006), "Sequentially optimal mechanisms." *The Review of Economic Studies*, 73, 1085–1111. [1659]

Skreta, Vasiliki (2015), "Optimal auction design under non-commitment." *Journal of Economic Theory*, 159, 854–890. [1659]

Sobel, Joel (2013), "Giving and receiving advice." In *Advances in Economics and Econometrics* (D. Acemoglu, M. Arellano, and E. Dekel, eds.), 305–341, Cambridge University Press. Chapter 10. [1658]

Tirole, Jean (1986), "Procurement and renegotiation." *Journal of Political Economy*, 94, 235–259. [1659]

Vartiainen, Hannu (2013), "Auction design without commitment." *Journal of the European Economic Association*, 11, 316–342. [1659]

Wolinsky, Asher (2002), "Eliciting information from multiple experts." *Games and Economic Behavior*, 41, 141–160. [1658]

Zheng, Charles Z. (2002), "Optimal auction with resale." *Econometrica*, 70, 2197–2224. [1659]

---

Co-editor Thomas Mariotti handled this manuscript.

Manuscript received 28 March, 2019; final version accepted 15 February, 2021; available online 25 February, 2021.