# Bilateral Trade With a Benevolent Intermediary*

Ran Eilat† Ady Pauzner‡

March 28, 2019

**Abstract**

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 some trades they believe to be beneficial – a commitment that is used by mechanisms to achieve ex-ante optimality. The intermediation game is equivalent to a mechanism design problem with an additional "credibility" constraint, requiring that every outcome be interim-optimal, conditional on available information. Consequently, an interesting information trade-off arises, whereby acquiring fine information makes the trading decision more responsive to the parties' valuations, while coarse information allows more flexibility to credibly deny beneficial trades. We investigate how such intermediaries communicate with the parties and make decisions, and derive some properties of optimal intermediaries.

## 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, divorces proceedings and real-estate transactions. While intermediaries may be motivated by various goals, such as maximizing their own profit or building up a reputation, helping the parties to materialize potential gains from trade is often their main motivation. 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.

---

*This paper supersedes an earlier working paper entitled "Mediators and Bilateral Trade".

†Department of Economics, Ben Gurion University. eilatr@bgu.ac.il

‡The Eitan Berglas School of Economics, Tel-Aviv University. pauzner@post.tau.ac.il

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 these policies turn out to be sub-optimal when information is revealed. Such commitment power can be achieved, for example, by delegating the execution to a separate institution or to a hard-wired mechanism. 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 act against the interests of the parties in some cases. Such a commitment is often used in optimal mechanisms to incentivize the parties to reveal private information, thereby helping to maximize ex-ante social surplus.

For concreteness, consider the canonical bilateral trade problem in Myerson and Satterthwaite (1983). This simplified environment, which we also adopt in this paper, 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 of the object 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) show 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 achievable outcomes. 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 environment 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. Our model shows how such an intermediary can circumvent this limitation by restricting 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 into the same information set, the intermediary can sometimes withhold beneficial trade and thereby incentivize the agents to reveal their private information. The downside is, 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.

To formally capture the idea that the intermediary cannot commit to future actions we model intermediation in the bilateral trade context as a three-player game – between the intermediary, the buyer and the seller. The intermediary is the first to move and decides on the message sets for the other two agents. The buyer and the seller then simultaneously choose a message from their permissible set, or choose to quit. The messages are costless (so talking is cheap) and used only to convey information regarding the agents' preferences. If no agent quits, the intermediary updates his beliefs and makes a decision – whether the object is traded and at what price – in a way that maximizes the expected social surplus given his beliefs.

The communication phase in our model is one-shot – 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 trade-off in which we are interested – the benefit versus the harm of possessing finer information when commitment is imperfect (particularly in a multi-agent environment). Adding more stages of communication may indeed expand the set of achievable outcomes, since this would allow the intermediary to (sometimes) have more freedom in controlling the flow of information, but would not change the basic intuitions illuminated by the simpler model.

In section 2, we present the model and provide an equivalence result, according to which the surplus-maximizing equilibrium in the intermediation game is equivalent (in terms of the social choice function it induces) to the equilibrium of the optimal mechanism among those that satisfy a property we call credibility – that every outcome of the mechanism is interim-optimal conditional on the available information.[1] We refer to such mechanisms as *intermediation mechanisms*. We can therefore study intermediation games by analyzing intermediation mechanisms, using tools of mechanism design.

It is well-known that in mechanism design with imperfect commitment, such as in our case, the revelation principle does not necessarily hold. Thus, it is *not* without loss of generality to assume that each agent truthfully reports his type in equilibrium. In section 3, we restore a weak version of the revelation principle that applies in our setting: We

[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.

show that it is without loss of generality to assume that in intermediation mechanisms each agent's message set consists of intervals that partition the agent's type space and that in equilibrium each type truthfully reports the interval to which he belongs. This follows from a property that holds in our framework (but not in all imperfect-commitment environments), namely that it is without loss to assume that each agent's equilibrium strategy is pure (that is, no type randomizes).[2]

In section 4, we explore some properties of optimal intermediation mechanisms. We show that the allocation rule of such mechanisms must attain a special "step-form" that follows from the credibility requirement and has a convenient graphical representation. We then use this representation to provide conditions for the allocation rule and message sets of an intermediation mechanism, in order for the mechanism to be budget-balanced.

In section 5, we study the bounds on the surplus that intermediation mechanisms can achieve. We show that an intermediation mechanism is *strictly* less efficient than a mechanism with full commitment power and can alway do at least as well as a simple posted price. We then solve for the optimal intermediation mechanism when the agents' values are uniformly distributed and show that in this case imperfect commitment is extremely harmful – the intermediary can do no better than a simple posted price. We conclude by discussing the link between credibility and budget balance and show, by means of an example, that when the designer's resource constraint is relaxed (i.e. subsidizing trade is allowed) the negative consequences of imperfect commitment are less severe.

**Related Work.** Communication in our model is cheap talk, i.e., the information transmitted by the agnets is costless and non-verifiable. An extensive body of literature that followed the seminal paper by Crawford and Sobel (1982) on cheap-talk 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 equilibria outcomes, and Gerardi, McLean and Postlewaite (2008) who consider a receiver who can commit to distort the outcome, relative to his optimal rule, in order to incentivize the senders to reveal information.[3] 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

---

[2] c.f. Bester and Strausz (2000) who show, in different environment with multiple agents and imperfect commitment, that allowing the agents to randomize can sometimes be strictly beneficial for the designer.

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

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 expand the set of equilibria in games. In more recent contributions, Goltsman et al. (2009) compare different 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 are Blume, Board and Kawamura (2007), Fey and Ramsey (2009) and Ivanov (2010).[4]

Our paper is closely related to the literature on mechanism design with imperfect commitment. In the context of implementation problems with symmetric information, Baliga, Corchon and Sjostrom (1997) consider, as we do, a utility-maximizing designer who cannot commit to outcomes. In their setting, however, a fully revealing equilibrium always exists due to the symmetry of information, and the main concern is ruling out unwanted equilibria. Chakravorty, Corchon and Wilkie (2006) show that, in exchange economies, if a designer is restricted to off-equilibrium outcomes that must be optimal for at least some preference profiles, then a broad set of social choice functions is not Nash-implementable. In contrast, we consider a setting in which information is asymmetric across agents and assume that off-equilibrium outcomes are those preferred by the designer.[5]

A number of papers have studied the problem of multi-period 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 in case the item remains unsold.[6] 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 (2018) consider a different type of imperfect commitment model in which an auctioneer who communicates sequentially and privately with the buyers can deviate from the predetermined rules only if the deviation is undetectable. In contrast, the designer in our

[4]In a political science context, Kydd (2003, 2006) uses game-theoretic models to study how mediators can help parties avoid costly conflicts by strategically sharing information about the cost of fighting of one player (Kydd 2003) or the trustworthiness of the players (Kydd 2006).

[5]We make this assumption to rule out unreasonable equilibria that rely on "threats" by the agents to coordinate on bad outcomes. See the the description of the game and our solution concept in section 2.

[6]Zheng (2002) considers 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.

model is not concerned with whether or not his deviations can be detected.

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 ask for another round of bids, but in contrast to our model, the seller cannot affect the strategy space of the buyers. Vartiainen (2013) considers 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. Crucially, the designer in our model cannot use a communication device and therefore cannot add noise to the agents' reports. Another key distinction between our paper and the above-mentioned body of literature is that we assume the designer's objective to be welfare maximization 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.[7] 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 (2018) provide 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 the current paper, we do not allow the designer to use a communication device and the (coarse) revelation result we develop relies on the credibility property of the mechanism.

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

[7]Bester and Strausz (2001) show 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. In a different paper, Bester and Strausz (2000) show that this result does not extend to the case of multiple agents, like the one in our model.

# 2 The intermediation game and mechanism

In this section we describe the model. We first define the three-player game between the buyer, the seller and the intermediary. Next, we 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. We then provide an equivalence result according to which the outcome of the optimal intermediation mechanism is the same as that of the intermediation game. Therefore, rather than looking for equilibria of the game we can instead analyze the problem of a mechanism designer tasked with devising the optimal intermediation mechanism. This allows us to then employ some (but not all) of the tools of mechanism design. Finally, we briefly discuss the non-trivial connection between individual rationality and voluntary participation in problems of mechanism design with imperfect commitment, such as the one we analyze.

## 2.1 The intermediation game

Intermediation in the bilateral-trade problem can be 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 non-zero even if the object is not transferred). The value of the object for agent $i \in \{s, b\}$ is $v_i$ and the agents are risk neutral: given that 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 not to participate, 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. We assume that $F_i$ admits a density $f_i$ that is continuously differentiable and bounded away from zero over the interval $V_i$. For non-triviality, we assume that the intersection $V_s \cap V_b$ is non-empty.

The timing of the game is as follows:

**Stage 0** The agents decide whether or not to participate in intermediation. If one of the agents opts out, the game terminates without trade or payments.

**Stage 1** The intermediary specifies a finite 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 $m_i \in M_i$. If one of the messages is "out", then the game terminates without trade or payments.

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

We refer to the game starting at stage 2 as the *reporting sub-game*.

*Opting-out.* The ability of the agents to opt out at stages 0 and 2 captures the idea of interim voluntary participation, i.e., an agent can choose to opt out as long as he has not started to talk.[9] Note that we do not allow the agents to opt-out after the intermediary makes a decision. That is, the intermediary's decision binds the agents and we do not require ex-post voluntary participation.[10]

*Equilibrium.* The solution concept is Perfect Bayesian Equilibrium (PBE). We impose the refinement that the equilibrium chosen in the reporting sub-game is the best for the intermediary, i.e., the one that maximizes the expected social surplus.[11] We adopt this refinement in order to rule out unreasonable equilibria in which the intermediary is "forced" to choose suboptimal message sets, simply because the agents would otherwise coordinate on a bad equilibrium in the reporting sub-game (e.g. the babbling equilibrium).

## 2.2 Intermediation mechanisms

A *trade mechanism* $\Gamma$ consists of two finite sets of messages, $M_s$ for the seller and $M_b$ for the buyer, and an outcome function that determines, for each message pair $m = (m_s, m_b) \in M = M_s \times M_b$, an outcome $(p(m), t(m))$, where the allocation rule $p : M \to [0, 1]$

[8] The assumption that the messages sets are finite simplifies the exposition but is not necessary - all the results can be obtained without it at the cost of greater complexity and without additional insights.

[9] Removing stage 0, while maintaining the ability to opt-out at stage 2, does not change the set of equilibria in the game. However, removing the ability to opt-out at stage 2 trivializes the game, since in this case the intermediary would always seek to implement the first-best outcome in the sub-game that begins in period 1. But then the only equilibrium in the complete game would be the one that implements a posted price.

[10] An alternative modeling choice could be to allow the intermediary to only suggest an outcome and allow the agents to reject it. Both modeling choices lead to similar general insights, and here we analyze the simpler one, which does not require an additional stage in which the agents play again.

[11] 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.

determines the probability that the object is traded and the transfer rule $t : M \to \mathbb{R}$ determines the monetary transfer from the buyer to the seller (note that we implicitly impose ex-post budget balance).

Given an allocation rule $p$ and transfer rule $t$, the utility of type $v_s$ of the seller when the message pair $m = (m_s, m_b)$ is reported is $u_s(v_s; m) = -p(m) v_s + t(m)$ while the utility of type $v_b$ of the buyer is $u_b(v_b; m) = p(m) v_b - t(m)$. The social surplus is $W((v_s, v_b), m) = (v_b - v_s) p(m)$.

A strategy for agent $i$ is a function $\sigma_i : V_i \to \Delta(M_i)$ that maps each of the agent's types to a distribution over messages, while a Bayesian Nash Equilibrium (BNE) is a profile of strategies $\sigma = (\sigma_s, \sigma_b)$ such that each one 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".

Given trade mechanism $(\Gamma, \sigma)$, we denote by $\bar{u}_i(v_i, m_i)$ the expected utility of type $v_i$ of agent $i$ who reports the message $m_i$ in equilibrium and by $\bar{p}_i(m_i)$ his expected probability of trade:

$$\begin{aligned} \bar{u}_i(v_i, m_i) &= \mathbb{E}_{m_{-i}} u_i(v_i; m_i, m_{-i}) \\ \bar{p}_i(m_i) &= \mathbb{E}_{m_{-i}}\, p(m_i, m_{-i}) \end{aligned}$$

where the probability weight of each $m_{-i}$ is induced by agent $-i$'s equilibrium strategy $\sigma_{-i}$.[12] We require trade mechanisms to satisfy *individual rationality*:

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

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

Given trade mechanism $(\Gamma, \sigma)$ and message profile $m$ which is reported with positive probability in equilibrium, the *interim surplus* induced by $m$ is defined as:

$$W_I(m) = \mathbb{E}_{v_s, v_b}[v_b - v_s | m] \cdot p(m)$$

where $E_{v_i}(v_i|m)$ is the posterior mean of $v_i$ given that agent $i$, who plays according to

[12]The probability that message $m_{-i}$ is sent by agent$-i$ in equilibrium is $\mathbb{E}_{v_{-i}} \sigma_{-i}(v_{-i})[m_{-i}]$, where $\sigma_{-i}(v_{-i})[m_{-i}]$ is the probability that type $v_{-i}$ reports $m_{-i}$ when playing according to the strategy $\sigma_i(v_{-i})$. In the case of a pure strategies equilibrium, in which the agents do not randomize, $\bar{u}_i(v_i; m_i) = \mathbb{E}_{v_{-i}} u_i(v_i; m_i, \sigma_{-i}(v_{-i}))$

strategy $\sigma_i$, chooses message $m_i$. The *ex-ante surplus* is given by:

$$W_{EA} = \mathbb{E}_m W_I(m)$$

where the probability weight of each $m$ is induced by the equilibrium strategies. In the case of a pure strategies equilibrium, in which the agents do not randomize, we have that $W_{EA} = \mathbb{E}_{v_s, v_b}[(v_b - v_s) \cdot p(\sigma_s(v_s), \sigma_b(v_b))]$.

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 is said to be *credible* if, for any message profile $m$, and unless one of the agents opts out, the allocation rule maximizes the interim social surplus. An *intermediation mechanism* is a credible trade mechanism:

**Definition 1 (Credibility and Intermediation Mechanisms)** *Trade mechanism* $(\Gamma = \langle M, p, t\rangle, \sigma)$ *is credible if* $p(m)$ *maximizes the interim surplus* $W_I(m)$ *for any* $m = (m_1, m_2)$ *that is reported with positive probability in equilibrium, unless either* $m_1$ *or* $m_2$ *is an opt-out message.*[13]
*A trade mechanism that satisfies credibility is an* ***intermediation mechanism****.*

Note that credibility does not restrict the allocation rule off the equilibrium path, where Bayes' rule cannot be used to update beliefs, and does not apply if the agents opt out. For a discussion of the necessity of the opt-out messages in our model and their relation to credibility and individual rationality, see section 2.2.3.

An *optimal intermediation mechanism* maximizes the ex-ante surplus over all intermediation mechanisms:

**Definition 2 (Optimality)** *Intermediation mechanism* $(\Gamma, \sigma)$ *is optimal if* $W_{EA}(\Gamma, \sigma) \geq W_{EA}(\Gamma', \sigma')$ *for any intermediation mechanism* $(\Gamma', \sigma')$*.*

The designer is tasked with devising the optimal intermediation mechanism. We will now formally described his problem.

---

[13] When maximizing over the possible allocation decisions, we hold the equilibrium strategies fixed. That is, $p(m)$ is the maximizer of $W_I(m) = E_{v_s, v_b}[v_b - v_s | m] \cdot p'$ over all $p' \in [0, 1]$, where the left multiplier is held fixed (i.e., the conditional expectations are computed according to the equilibrium profile $\sigma$ of the original trade mechanism).

### 2.2.1 The designer's problem

The designer's problem is as follows: Find a trade mechanism $\Gamma = \langle M, p, t \rangle$ and a profile of strategies $\sigma = (\sigma_s, \sigma_b)$ that maximize the ex-ante social welfare $W_{EA} = \mathbb{E}_m \mathbb{E}_{v_s, v_b} [v_b - v_s | m] \cdot p(m)$, subject to three constraints:

1. *Incentive-compatibility* – given $\Gamma$, the strategy $\sigma_i$ is optimal:

$$\bar{u}_i (v_i, m_i) \geq \bar{u}_i (v_i, m_i') \tag{IC}$$

   for all $v_i \in V_i$ and $m_i \in supp(\sigma_i(v_i))$, and for every $m_i' \in M_i$,

2. *Individual rationality* – given $\Gamma$ and $\sigma$, each type $v_i$ who follows $\sigma_i$ is not worse off than if he had not participated in $\Gamma$:

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

   for all $v_i \in V_i$ and all $m \in supp(\sigma_i(v_i))$,

3. *Credibility* – given $\Gamma$ and $\sigma$, the allocation rule $p$ maximizes the interim surplus:

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

   for every profile of messages $(m_s, m_b) \in M_s \times M_b$ that is reported with positive probability, unless either $m_s$ or $m_b$ is an opt-out message.

### 2.2.2 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 of distributions over outcomes $X = [0,1] \times \mathbb{R}$.

We say that a social choice function $scf$ is *implementable* by the game (by the reporting sub-game) if there exists an equilibrium in the game (in the reporting sub-game) such that for any profile of types $v \in V$ the probability distribution it induces over outcomes is $scf(v)$. Similarly, we say that $scf$ is implementable by an intermediation mechanism if, for any profile of types $v \in V$, 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 $scf$:*

1. *$scf$ is implementable by the reporting sub-game starting with message set $M$ if and only if it is implementable by an intermediation mechanism with message set $M$.*

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

The formal proof is relagated to the appendix (as are all other proofs in the paper, except in straightforward cases). 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 we imposed, which selects the intermediary-optimal equilibrium in any sub-game.

#### 2.2.3 Voluntary participation, individual rationality and opt-out messages

In many real-world situations, individuals can decide whether or not to participate in a mechanism (for example, an individual can decide not to participate in intermediation if he expects the outcome to be unfavorable). In standard mechanism design, in which the designer has full commitment power, this property of *voluntary participation* is usually captured by imposing the requirement of *individual rationality* on the mechanism, i.e., each type of each agent must have at least one action that is as good for him (in expectations) as his outside option. It is then without loss of generality to assume that all types of all agents participate.

The equivalence between voluntary participation and individual rationality relies on the designer's full commitment power, namely his ability to choose any mapping between actions and outcomes in the mechanism. For instance, the designer can always add an action and commit, that if the agent picks it, the outcome will be the same as if he had not participated. This equivalence no longer holds in mechanisms with imperfect commitment. In our context, for example, requiring the outcome to be interim-optimal may conflict with giving the agents zero payoffs if the interim beliefs imply that trade is beneficial. Thus, requiring that all types of both agents obtain non-negative utility within the mechanism (i.e. individual rationality) might be too restrictive and may narrow the set of equilibria relative to the case in which agents are explicitly allowed not to participate. To capture voluntary participation (which is the "real-world" property in which we are interested) without restricting the set of equilibria, we impose individual rationality but allow the mechanism to use opt-out messages. An opt-out message is exempt from the credibility requirement, but must induce a no-trade and no-payments outcome. In this way, the designer always has the ability to provide the agents with actions that are as good as their outside options.

## 3 Partition-direct representation

In conventional 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 true 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.

In this section, we develop a coarse version of the revelation argument that applies to intermediation mechanisms: we show that it is without loss of generality to assume that each agent's message set partitions his type space into intervals, and that in equilibrium each agent reports the interval to which its type belongs.[14] We refer to a mechanism that satisfies these properties as a *partition-direct* mechanism.

To prove this result, we show that due to the credibility constraint, intermediation mechanisms satisfy a property we call message-monotonicity, whereby for any two messages of agent $i$ (with one exception to be discussed below), one of the messages induces higher trade probability than the other for *any* message of agent $-i$. We then use this property to show that it is without loss to assume that almost all types of both agents employ pure strategies (that is, they do not randomize in equilibrium).[15] Consequently, it is possible to partition each player's set of types according to the messages they send in equilibrium and to rename each message to be the set of types that send it. By the single-crossing property of the preferences, the set of types that pool to send each message is convex, hence messages are intervals.

[14] Type $v_i$ belongs to the interval $[a, b]$ if $v_i \in [a, b]$.

[15] Note that agents using pure strategies is not a straightforward result 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) show 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. In particular, this means that some types randomize.

## 3.1 Pure-strategy equilibrium

Suppose that $(\Gamma = \langle M, p, t \rangle, \sigma)$ is an intermediation mechanism. Our first lemma asserts it must satisfy *message monotonicity*: if two messages of agent $i$ induce different posterior mean type for agent $i$, then reporting one of the messages necessarily leads to (weakly) higher trade probability than the other, *for any message of agent* $-i$:

**Lemma 1 (Message Monotonicity)** *Let* $(\Gamma = \langle M, p, t \rangle, \sigma)$ *be an intermediation mechanism, and assume that* $m_i \in M_i$ *and* $m_i' \in M_i$ *are two messages of agent* $i$ *such that* $\mathbb{E}_{v_i}[v_i|m_i] \neq \mathbb{E}_{v_i}[v_i|m_i']$*. If* $p(m_i, m_{-i}) > p(m_i', m_{-i})$ *for some* $m_{-i} \in M_{-i}$ *then* $p\left(m_i, m_{-i}'\right) \geq p\left(m_i', m_{-i}'\right)$ *for all* $m_{-i}' \in M_{-i}$.

To see the intuition of this result, note that since types are independent, the posterior belief regarding the type of agent $i$ in equilibrium depends only on the message that is sent by agent $i$ (and not on the message that is sent by agent $-i$). When $\mathbb{E}_{v_i}[v_i|m_i] \neq \mathbb{E}_{v_i}[v_i|m_i']$, credibility implies that $p(m_i, m_{-i}) > p(m_i', m_{-i})$ if and only if the mean type of agent $i$ that sends $m_i$ is greater than the mean type that sends $m_i'$. However, credibility then also implies that $p\left(m_i, m_{-i}'\right) \geq p\left(m_i', m_{-i}'\right)$ for all other $m_{-i}'$.[16]

Having established that intermediation mechanisms are message-monotonic, we now proceed to show that if some agent $i$ has two messages $m_i', m_i''$ such that $\bar{p}_i(m_i') = \bar{p}_i(m_i'')$, then one of the messages, $m_i'$ or $m_i''$, can be eliminated – there exists another intermediation mechanism with a smaller message set in which each type of each agent gets the same expected payoff as in the original mechanism, and therefore the ex-ante surplus of both mechanisms is the same:

**Lemma 2** *Suppose that* $(\Gamma = \langle M, p, t \rangle, \sigma)$ *is an intermediation mechanism and that* $\bar{p}_i(m_i') = \bar{p}_i(m_i'')$ *for some two messages* $m_i'$ *and* $m_i''$ *of agent* $i$*. Then, there is an intermediation mechanism* $\left(\hat{\Gamma} = \left\langle \hat{M}, \hat{p}, \hat{t} \right\rangle, \hat{\sigma}\right)$ *with* $\left|\hat{M}_i\right| < |M_i|$ *such that:*

$$\bar{u}_j^{(\Gamma,\sigma)}(v_j, m_j) = \bar{u}_j^{\left(\hat{\Gamma},\hat{\sigma}\right)}(v_j, \hat{m}_j)$$

*for every* $v_j \in V_j$ *and* $m_j \in supp[\sigma_j(v_j)]$ *and* $\hat{m}_j \in supp[\hat{\sigma}_j(v_j)]$ *and for each agent* $j \in \{b, s\}$.

[16] Note that the argument does not apply if $\mathbb{E}_{v_b}[v_b|m_b] = \mathbb{E}_{v_b}[v_b|m_b'] = \mathbb{E}_{v_s}[v_s|m_s] = \mathbb{E}_{v_s}[v_s|m_s']$ for some four messages $m_b, m_b' \in M_b$ and $m_s, m_s' \in M_s$. This is because credibility does not pin down the allocation rule when the surplus from trade is zero. It might therefore be that, for example, $p(m_b, m_s) = p(m_b', m_s') = 1$ while $p(m_b', m_s) = p(m_b, m_s') = 0$.

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 that $m_i'$ 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 $m_{-i}$). In both cases, we can merge the two messages, $m_i'$ and $m_i''$, and update agent $i$'s strategy so that all types who sent $m_i'$ and $m_i''$ will send the merged message. We then update the allocation rule so that the expected probability of trade for each message of agent $-i$ is unchanged and adjust the transfer rule accordingly to support the equilibrium. By doing so we keep the equilibrium payoff of each type of each agent unchanged. Finally, we show that merging the messages does not violate credibility, thus if the original mechanism was an intermediation mechanism, then so is the modified one.

By Lemma 2, we can restrict attention to *minimal* intermediation mechanisms (in terms of the size of the message sets) in which, for each agent, there are no two messages that induce the same expected probability of trade. That is, given an intermediation mechanism $(\Gamma = \langle M, p, t \rangle, \sigma)$ we can assume that for each agent $i \in \{b, s\}$:

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

In a minimal intermediation mechanism, it must be the case that almost all types of both agents employ pure strategies (that is, they do not randomize):

**Lemma 3** *In any minimal intermediation mechanism* $(\Gamma = \langle M, p, t \rangle, \sigma)$*, almost all types of both agents do not randomize: For each agent* $i$*, and for almost all types* $v_i \in V_i$*, the strategy* $\sigma_i(v_i)$ *puts all the probability mass on a single message* $m_i \in M_i$ *and probability mass zero on all others.*

The proof is straightforward: In a minimal intermediation mechanism, there cannot be two distinct types of agent $i$ who are indifferent (and who possibly randomize) between the same two messages $m_i$ and $m_i'$.[17] Since the message set for each agent is finite, there are only finitely many types who randomize in a minimal intermediation mechanism.

## 3.2 A coarse-revelation result

Given a minimal intermediation mechanism (in which, according to Lemma 3, almost all types employ pure strategies), we can partition each agent's set of types according to the

[17] To see this, note that if two buyer types $v_b$ and $v_b'$ are indifferent between two distinct messages, $m_b$ and $m_b'$, then $\bar{p}_b(m_b) \cdot v_b - \bar{t}_b(m_b) = \bar{p}_b(m_b') \cdot v_b - \bar{t}_b(m_b')$ and $\bar{p}_b(m_b) \cdot v_b' - \bar{t}_b(m_b) = \bar{p}_b(m_b') \cdot v_b' - \bar{t}_b(m_b')$, where $\bar{t}_b(m_b) = \mathbb{E}_{m_s}\, t(m_s, m_b)$. Since $v_b \neq v_b'$, it must be the case that $\bar{p}_b(m_b) = \bar{p}_b(m_b')$, which is a contradiction of (MIN). The argument for the seller is analogous.

messages they send in equilibrium, and rename each message to be the set of types that send it.[18] Note that in the modified mechanism the elements of each message set $M_i$ form a partition of the type space $V_i$. Moreover, since the set of types that send each message in the original mechanism is convex,[19] each element $m_i \in M_i$ in the modified mechanism is an interval. Clearly, in the modified mechanism it is a best response for each type of each agent to report "truthfully", that is, to report the message to which it "belongs". We can therefore restrict attention to mechanisms in which messages are intervals 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 a strategy of agent $i$ 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 called a *truth-telling* equilibrium. We thus define:

**Definition 3 (Partition-Direct)** *An intermediation mechanism is* partition-direct *if: (i) for each agent $i$, the message set $M_i$ consists of intervals that partition the type-space $V_i$, and (ii) truth-telling is an equilibrium.*

This leads to the following proposition:[20]

**Proposition 2** *Given any intermediation mechanism, there exists a partition-direct intermediation mechanism, such that the expected payoff for all types of each agent is the same in both mechanisms, and both mechanisms achieve the same ex-ante social surplus.*

We thus obtain a weaker version of the revelation result: When looking for the optimal intermediation mechanism, it is without loss of generality to restrict our attention to partition-direct intermediation mechanisms.

---

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

[19] This is a straightforward implication of the single-crossing property of the agents' preferences. To see this, suppose that two buyer types $v_b^l \in V_b$ and $v_b^h \in V_b$ send the message $m_b' \in M_b$ in equilibrium and that $v_b^l < v_b^h$. Suppose further, by way of contradiction, that some type $v_b'' \in \left(v_b^l, v_b^h\right)$ sends the message $m_b'' \neq m_b'$ in equilibrium, so that the set of types that send $m_b'$ is not convex. Then, it must be the case that $\bar{p}_b\left(m_b''\right) \cdot v_b'' - \bar{t}_b\left(m_b''\right) \geq \bar{p}_b\left(m_b'\right) \cdot v_b'' - \bar{t}_b\left(m_b'\right)$, where $\bar{t}_b\left(m_b\right) = \mathbb{E}_{m_s}\, t\left(m_s, m_b\right)$. But since the mechanism is minimal then $\bar{p}_b\left(m_b'\right) \neq \bar{p}_b\left(m_b''\right)$, and therefore either $v_b^h$ or $v_b^l$ prefer sending the message $m_b''$ to sending the message $m_i'$, which is a contradiction to $m_b'$ being their optimal message. The proof for the seller types is analogous.

[20] The proof is given in the first paragraph of this sub-section (noting that by Lemma 2 it is without loss of generality to restrict attention to *minimal* intermediation mechanisms)

Given message set $M_i$ whose elements are intervals, we denote the upper and lower bounds of each interval $m_i \in M$ by $\bar{m}_i$ and $\underline{m}_i$, respectively. We also denote $\Pr(m_i) = F_i(\bar{m}_i) - F_i(\underline{m}_i)$, so that $\Pr(m_i)$ is the probability that the message $m_i$ is reported by agent $i$ with message set $M_i$ in a partition-direct intermediation mechanism. Finally, we refer to the message that pools the lowest types of the buyer as the *buyer's most reluctant message*, and to the message that pools the highest types of the seller as the *seller's most reluctant message*. Note that if agent $i$ has an opt-out message, it must be his most reluctant one.[21]

# 4 Optimal intermediation mechanisms

In light of the coarse revelation result in section 3, we can look for the optimal intermediation mechanism within the class of partition-direct intermediation mechanisms. Throughout this section we therefore assume that message sets consist of intervals that partition the agents' type space.

We first define an auxiliary property of allocation rules, which is slightly stronger than credibility. We say that an allocation rule $p$ *induces only beneficial trade*, with respect to some message set $M$, if $p$ imposes trade *if and only if* the interim surplus from trade is *strictly* positive.[22] We show that such an allocation rule takes a special form which we refer to as "step-form".

We next consider an auxiliary type of intermediation mechanism whose budget is required to be balanced only ex-ante (and not necessarily ex-post) and show that a well-known result from the mechanism design literature, that establishes an equivalence between the ex-ante and ex-post budget-balance notions, also holds in the case of intermediation mechanisms.

Finally, we introduce a function $\psi(M, p)$, defined over message-sets and allocation rules that induce only beneficial trade, and show that $\psi(M, p)$ is equal to the minimal (ex-ante) budget deficit required to sustain incentive compatibility and individual rationality in any trade mechanism with messages $M$ and allocation rule $p$. We show that $\psi(M, p) = 0$ is a sufficient condition for the existence of an intermediation mechanism $\langle M, p, t\rangle$,

[21] This is because an opt-out message induces zero expected probability of trade. Since in equilibrium it must be that the expected probability of trade for each agent is increasing in his eagerness to trade (higher buyer types trade with higher probability and vice-versa for the seller), then an opt-out message must pool the types that are most reluctant to trade for each agent.

[22] Note that credibility requires that trade takes place when the interim surplus is strictly positive and that there is no trade when the interim surplus is strictly negative. It does not impose restrictions in the case that it is exactly zero.

and provide a simple geometric interpretation of $\psi$. We also show that in an *optimal* intermediation mechanism $\langle M, p, t\rangle$, it must be the case that $p$ induces only beneficial trade with respect to $M$ and that $\psi(M, p) = 0$. Thus, in order to find the optimal intermediation mechanism we can restrict our search to $(M, p)$ satisfying $\psi(M, p) = 0$, and then look for the one that produces the highest ex-ante surplus.

## 4.1 Only beneficial trade and step-form trading rules

Given a partition-direct intermediation mechanism $(\Gamma = \langle M, p, t\rangle, \sigma)$, credibility implies that the allocation decision $p(m_s, m_b)$ 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 mean type in the interval $m_b$ (evaluated according to the distribution $F_b$) is strictly larger than the mean type in the interval $m_s$ (evaluated according to the distribution $F_s$), then $p(m_s, m_b) = 1$; if it is strictly smaller, then $p(m_s, m_b) = 0$. 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 will now define a stronger, auxiliary property, which requires that when the expected surplus from trade is zero, there can be no trade. This definition is useful because the message set $M$ then pins down the allocation rule $p$, up to only one degree of freedom – whether or not the most reluctant message of each agent is an opt-out message (which is exempt from credibility, but must impose no trade and no payments).

**Definition 4 (Only Beneficial Trade)** *An allocation rule $p$ induces only beneficial trade with respect to message set $M$ if*

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

*for every $(m_s, m_b) \in (M_s, M_b)$, with one exception: if $m_i$ is agent $i$'s most reluctant message, then it can be that $p(m_i, m_{-i}) = 0$ for all $m_{-i}$.*

**Remark** *In the next section, we will show that the allocation rule must induce only beneficial trade in an optimal intermediation mechanism. While such a result is straightforward in standard mechanism design (where eliminating non-beneficial trade saves on rents which can then be used to sustain incentive compatibility when beneficial trade is added), it is not obvious for mechanisms with imperfect commitment. For example, in the case of intermediation mechanisms, adding beneficial trade by changing the partition of the type*

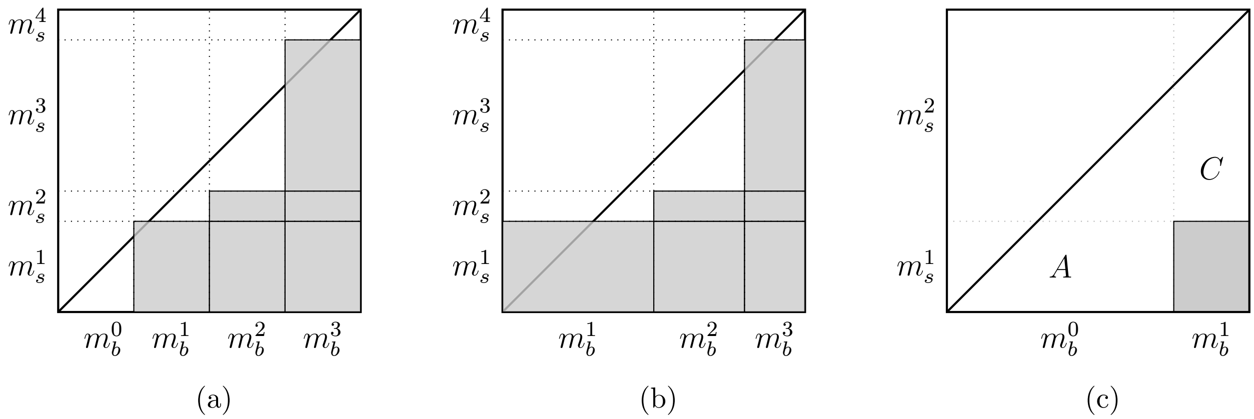


Figure 1: Allocation rules plotted on the agents' type space. The partition of the horizontal axis corresponds to the buyer's message set while the partition of the vertical axis corresponds to the seller's. The gray area corresponds to all the pairs of types who trade.

*space might violate the mechanism's credibility.*[23] *Thus, to show that non-beneficial trade can be eliminated we need to establish some additional results.*

When an allocation rule $p$ induces only beneficial trade with respect to message set $M$ it has a convenient graphical representation. If we plot the set of all the pairs of types who trade on the plane of types, this set has a special step-form: it is monotone in the agents' types, and its boundary consists of horizontal and vertical lines that correspond to the messages in $M_s$ and $M_b$ respectively.

As an example of this representation, consider Figures 1(a) and 1(b). In each, the seller types appear on the vertical axis (partitioned according to the messages in $M_s$), while the buyer types appear on the horizontal axis (partitioned according to $M_b$). The gray area in each figure corresponds to all the pairs of types that trade in equilibrium according to the plotted allocation rule, and the diagonal corresponds to the pairs in which the types are equal. Note that since intermediation mechanisms are minimal (by property (MIN) above), there are no two buyer messages with the same "height" of trade, and no two seller messages with the same "width".

For the analysis that follows, it will also be useful to denote the messages of each agent according to their order. Therefore we order the messages of each agent $i$ from lowest to highest and denote the $k^{th}$ message in $M_i$ by $m_i^k$. For convenience, we start the enumeration for the buyer from $k = 0$ if there is no trade when he reports the lowest interval, and from $k = 1$ otherwise. In this way, $m_b^1$ is the lowest buyer message for which

[23] Moreover, it is not even clear that the designer can eliminate non-beneficial trade and give the saved information rents to the agents as a lump sum, since this might change their participation decisions (see the discussion in section 2.2.3).

$\bar{p}_b(m_b) > 0$.[24] For the seller, the enumeration always starts from 1.

Since by property (MIN) we assume that for each agent there are no two messages that induce the same expected probability of trade, an allocation rule $p$ that induces only beneficial trade with respect to $M$ takes the following simple form:

$$p\left(m_s^k, m_b^l\right) = \begin{cases} 0 & k > l \\ 1 & k \leq l \end{cases} \tag{STEP-FORM}$$

for all $m_s^k \in M_s$ and $m_b^l \in M_b$.

## 4.2 Ex-ante budget balance

An intermediation mechanism, according to our definition, is ex-post budget-balanced: the monetary transfer $t$ goes directly from the buyer to the seller, and therefore the designer can never have a budget deficit. As a step towards solving the optimization problem, we consider an auxiliary form of intermediation mechanism in which the ex-post constraint is relaxed and replaced 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 non-negative. Beyond that, the auxiliary mechanism must also satisfy all the properties of an intermediation mechanism.

Formally, an *auxiliary intermediation mechanism* is $\Gamma' = \langle M, p, t_s, t_b \rangle$ where $t_i : M \to \mathbb{R}$ is the monetary transfer to agent $i$ (which may be negative). The definition of an opt-out message for agent $i$ is straightforwardly extended to become a message $m_i$ for which $t_s(m_s, m_b) = t_s(m_s, m_b) = p(m_s, m_b) = 0$ for all $m_{-i}$, and the definition for credibility remains unchanged (see Definition 1). The mechanism $\Gamma'$ is said to be *ex-ante budget-balanced* if:

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

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 (see, for example, Borgers and Norman, 2009). This equivalence result holds also in the case of intermediation mechanisms, but requires an adjustment in the proof:

**Lemma 4** *Suppose that $\Gamma' = \langle M, p, t_1, t_2 \rangle$ is a partition-direct auxiliary intermediation mechanism. If $\Gamma'$ is ex-ante budget-balanced, then there exists a transfer rule $t : M \to R$ such that $\Gamma = \langle M, p, t \rangle$ is a partition-direct intermediation mechanism.*

[24] See Figure 1(a) for an example in which enumeration for the buyer starts from 0, and Figure 1(b) for an example in which it starts from 1.

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 rates. The designer can therefore find a transfer rule $t$ that makes him pay 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 $M$ or the allocation rule $p$, credibility (and the property of inducing only beneficial trade) is maintained: if the ex-ante budget-balanced auxiliary mechanism satisfies it then so does the ex-post budget-balanced one. A modification of the proof is required to ensure that an opt-out message in $\Gamma'$ remains so in $\Gamma$ (i.e., if $m_i$ entails no transfers for agent $i$ under $\Gamma'$, then the same should be true under $\Gamma$).

## 4.3 Minimal budget and optimal intermediation mechanisms

Given message set $M$ and an allocation rule $p$ that induces only beneficial trade with respect to $M$, we define two functions that yield, for each message $m_i$ of agent $i$, the most reluctant type of the other agent that trades according to $p$:[25,26]

$$\begin{aligned} \omega_s(m_b) &= \sup\left(\max\left(m_s \in M_s : p(m_s, m_b) = 1\right)\right) \\ \omega_b(m_s) &= \inf\ \left(\min\left(m_b \in M_b : p(m_s, m_b) = 1\right)\right) \end{aligned}$$

When the intervals in $M_s$ and $M_b$ are enumerated according to the convention presented in Section 4.1, we have that $\omega_s\left(m_b^k\right) = \bar{m}_s^k$ and $\omega_b\left(m_s^k\right) = \underline{m}_b^k$. The functions $\omega_s(\cdot)$ and $\omega_b(\cdot)$ are given a geometric interpretation in the explanation of Figure 2(a) below.

We are now ready to define the function $\psi$:

$$\psi(M, p) = \sum_{(m_s, m_b)\ :\ p(m_s, m_b) = 1} \left(\omega_s(m_b) - \omega_b(m_s)\right) \cdot \Pr(m_s) \cdot \Pr(m_b) \qquad (2)$$

The value of $\psi(M, p)$ determines whether there exists a transfer rule $t$ such that $\langle M, p, t\rangle$ is an *intermediation mechanism*:

**Proposition 3** *Suppose that the allocation rule $p$ induces only beneficial trade with respect to the message set $M$. If $\psi(M, p) = 0$ then there is a transfer rule $t : M \to \mathbb{R}$ such that $\Gamma = \langle M, p, t\rangle$ is an intermediation mechanism. Conversely, if $\psi(M, p) > 0$, then there exists no $t : M \to \mathbb{R}$ such that $\langle M, p, t\rangle$ is an intermediation mechanism.*

[25] 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).

[26] $\max\left(m_s \in M_s : p(m_s, m_b) = 1\right)$ is the highest interval in $M_s$ for which $p(m_s, m_b) = 1$. Similarly, $\min\left(m_b \in M_b : p(m_s, m_b) = 1\right)$ is the lowest interval in $M_b$ for which $p(m_s, m_b) = 1$.

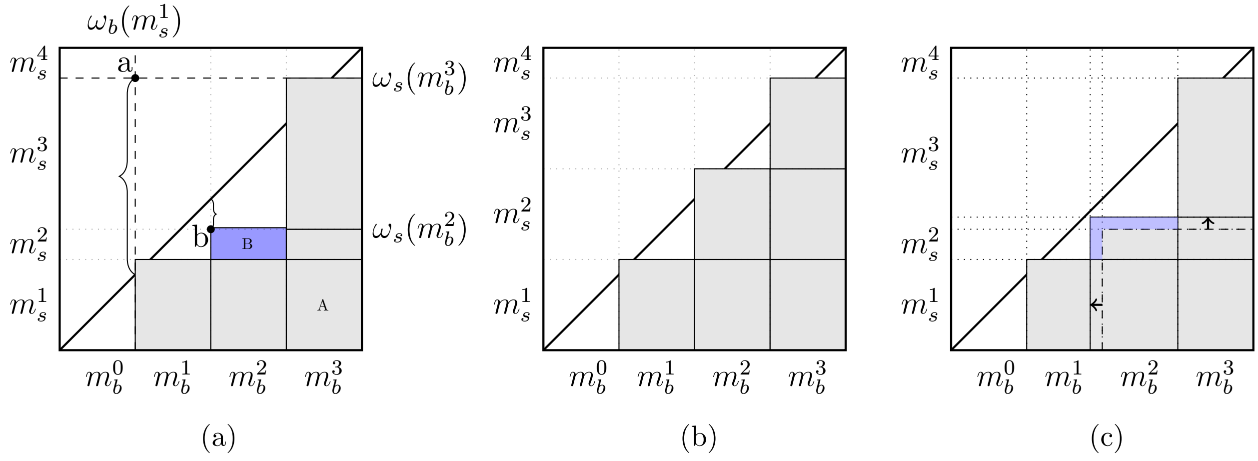


Figure 2: Allocation rules plotted on the agents' type-space. The left-hand panel illustrates the computation of $\psi$, the middle panel shows an allocation rule for which $\psi$ is positive, and the right-hand panel illustrates how an allocation rule with negative $\psi$ can be improved without violating credibility.

We prove the first part by constructing an ex-ante budget-balanced mechanism with message set $M$ and allocation rule $p$. We use the functions $\omega_s(\cdot)$ and $\omega_b(\cdot)$ to define the monetary transfers in the mechanism and then show that truth-telling by both agents is an equilibrium and that individual rationality is satisfied. The value of $\psi(M,p)$ is the (ex-ante) budget deficit entailed by the mechanism, and the fact that $\psi(M,p) = 0$ implies that the mechanism is ex-ante budget-balanced. Therefore, by Lemma 4, there also exists an equivalent ex-post budget-balanced intermediation mechanism and we are done. The second part follows from the fact that $\psi(M,p)$ is the minimal budget deficit that is entailed by any incentive-compatible and individually rational trade mechanism with message set $M$ and allocation rule $p$. Therefore, when $\psi(M,p) > 0$ a budget-balanced mechanism does not exist.

Figure 2(a) provides a geometric interpretation for the way in which $\psi$ can be computed. Rectangle $A$ corresponds to all the realizations in which the seller reports $m_s^1$ and the buyer reports $m_b^3$. We refer to $\Pr(m_s^1) \cdot \Pr(m_b^3)$ as the *probability mass of* $A$. Point $a$ corresponds to the rectangle $A$ and is determined as follows: its vertical coordinate is the highest seller type that trades when the buyer reports $m_b^3$, that is $\omega_s(m_b^3)$, and its horizontal coordinate is the lowest buyer type that trades when the seller reports $m_s^1$, that is $\omega_b(m_s^1)$. We refer to the vertical (or horizontal) distance between $a$ and the diagonal as the *information rent burden of* $A$. The value of $\psi$ is the sum, over all rectangles with trade, of the rectangle's probability mass multiplied by its information rent burden.

In Figure 2(a), all rectangles, except rectangle $B$, have a positive information rent burden (note that for rectangle $B$, point $b$ is below the diagonal). Therefore, the value of $\psi$ that corresponds to the allocation rule depicted in Figure 2(a) can be non-positive only if the product of rectangle $B$'s information rent burden and its probability mass is negative enough to compensate for all other elements in the sum, which are all positive.

More generally, according to property STEP-FORM, a rectangle's information-rent burden is negative if and only if it corresponds to intervals $m_s^k$ and $m_b^k$ for some (and same) $k$ *and* the rectangle's top left corner is strictly below the diagonal. It is easy to verify that there is no such rectangle in Figure 2(b). Therefore, there is no intermediation mechanism (or even a trade mechanism) that induces the trade depicted in Figure 2(b), and this is regardless of the prior distributions $F_s$ and $F_b$.

**Remark**. *Note that proposition 3 does not guarantee that an intermediation mechanism* $\Gamma = \langle M, p, t\rangle$ *exists when* $\psi(M,p) < 0$. *To see this, consider the example depicted in Figure 1(c) and assume that the prior distributions over types are uniform. It is easy to verify that the value of* $\psi(M,p)$ *in this example is indeed negative. However, any attempt to set transfers that maintain incentive compatibility in a mechanism* $\langle M, p, t\rangle$ *would necessitate non-zero transfers in the profiles of messages that correspond to rectangles* $A$ *or* $C$. *In this case either* $m_s^2$ *or* $m_b^0$ *could not serve as an opt-out message (which must have zero transfers), and credibility would imply that there must be trade in either* $A$ *or* $C$.

We conclude this section with two necessary conditions for the value of $\psi(M,p)$ and for the allocation rule $p$ in *optimal* intermediation mechanisms and explain how to improve on any intermediation mechanism that violates them, *without violating credibility.*

**Proposition 4** *Suppose* $\Gamma = \langle M, p, t\rangle$ *is an optimal intermediation mechanism. Then:*
*(i) The allocation rule* $p$ *induces only beneficial trade with respect to* $M$.
*(ii)* $\psi(M,p) = 0$.

Figure 2(c) shows the type of improvement that can be made when an intermediation mechanism has a negative value of $\psi$. By the previous discussion, there must exist $k \geq 1$ and two messages $m_s^k$ and $m_b^k$ such that the rectangle corresponding to $(m_s^k, m_b^k)$ has a top left corner below the diagonal (in Figure 2(c) this is $k = 2$). We then *gradually* increase the upper bound of $m_s^k$ and decrease the lower bound of $m_b^k$ (that is, we gradually bring the top left corner of the rectangle closer to the diagonal) and update $p$ to induce only beneficial trade with respect to the modified message set $M$, until either $\psi$ equals zero or $\bar{m}_s^k = \underline{m}_b^k$ (so that the top left corner of the rectangle touches the diagonal). Note that by doing so we *increase* the ex-ante social surplus since we are adding trade in beneficial realizations. Note also that when we update $p$ there is no pair of messages $(m_s, m_b)$ for which the trade probability increases (though it might decrease in some cases),[27] and

[27] To see the intuition of why there is no pair of messages $(m_s, m_b)$ for which the trade probability

therefore if $\psi$ increases, then it increases "continuously". If $\psi(M,p)=0$ we are done, because by Proposition 3 there exists an intermediation mechanism, with the modified message set and allocation rule, that provides a higher surplus than $\Gamma$. Otherwise, $\psi(M,p)$ is still negative and we reiterate the process with another rectangle.

To see why the (credible) allocation rule $p$ must induce only beneficial trade with respect to $M$, note that otherwise we could eliminate trade that is not beneficial (so that the modified allocation rule would induce only beneficial trade with respect to $M$) and thereby decrease $\psi$. Once $\psi$ is negative, we apply the first argument of the proof.

We can therefore look for the optimal intermediation mechanism among the set of intermediation mechanisms $(M,p,t)$ with allocation rule $p$ that satisfies only beneficial trade and for which $\psi(M,p)=0$.

# 5 Analysis

In this section, we apply the tools developed earlier to study the bounds on the surplus that intermediaries can achieve. We show that while an intermediary does at least as well as a posted price (in terms of social surplus), under mild conditions he is *strictly* less efficient than a mechanism with full commitment power. For the case of uniform prior distributions, we solve for the optimal mechanism and show that imperfect commitment is extremely harmful, since in this case the intermediary cannot do better than the lower bound. We conclude by discussing the link between credibility and budget balance and show by example that when the designer's resource constraint is relaxed (i.e. subsidizing trade is allowed) the consequences of imperfect commitment are less severe.

## 5.1 Infeasibility of the second-best outcome: an upper bound

It is obvious that the social surplus that can be attained by the optimal intermediation mechanism is bounded from above by that of the optimal (conventional) mechanism. But is this bound achievable? That is, can an intermediation mechanism sometimes do just as well as the conventional one? We show that for regular distributions the answer is negative – an intermediation mechanism does *strictly* worse than a conventional mechanism. To prove this result we rely on the characterization of the optimal mechanism provided by Myerson and Satterthwaite (1983).

increases, consider the example depicted in Figure 2(c) in which we increase $\bar{m}_s^2$ and decrease $\underline{m}_b^2$. Note that by doing so we decrease the mean buyer type in $m_b^1$ and $m_b^2$ and increase the mean seller type in $m_s^2$ and $m_s^3$. Thus, for all possible pairs of messages, other than $\left(m_s^2, m_b^2\right)$, trade is (weakly) less beneficial and therefore when we update $p$ the trade probability cannot increase.

A trade mechanism is said to be optimal if it maximizes the ex-ante social surplus among all feasible trade mechanisms. Myerson and Satterthwaite (1983) show that if the distributions $F_s$ and $F_b$ are regular (that is, the virtual valuations of the buyer and the seller are increasing in types),[28] then the optimal mechanism implements the following allocation rule:

$$p^*(v_s, v_b) = \begin{cases} 1 & \text{if } \; B(v_s, v_b) \leq 0 \\ 0 & \quad \text{otherwise} \end{cases} \tag{3}$$

where $B(v_s, v_b) = \left(v_s + \alpha \frac{F_s(v_s)}{f_s(v_s)}\right) - \left(v_b - \alpha \frac{1-F_b(v_b)}{f_b(v_b)}\right)$ for some $\alpha \in (0, 1]$, and $v_s$ and $v_b$ are the reported types of the seller and buyer, respectively.[29]

The allocation rule $p^*$ cannot be implemented by an intermediation mechanism. To see the intuition, consider the (continuous) function $v_s^*(v_b)$ that maps for each buyer type $v_b \in V_b$ the cutoff seller type below which all seller types trade with $v_b$ according to the allocation rule $p^*$, and above which they don't. Such a cutoff type exists because the function $B(\cdot, \cdot)$ is continuous, increasing in its first parameter and decreasing in its second parameter. The function $v_s^*(v_b)$ is continuous and strictly increasing, at least over some non-empty interval $[v_b', v_b''] \subseteq V_b$, and satisfies $v_s^*(v_b) < v_b$.[30,31] The fact that $v_s^*(v_b)$ is strictly increasing implies that each buyer type $v_b \in [v_b', v_b'']$ trades with a different set of seller types, and therefore, to implement $p^*$ the mechanism has to separate the buyer types $[v_b', v_b'']$ and seller types $[v_s^*(v_b'), v_s^*(v_b'')]$ in equilibrium. But then, for any $v_b \in (v_b', v_b'')$ and small enough $\varepsilon > 0$ we have that the seller type $v_s^*(v_b) + \varepsilon$ does not trade with the buyer type $v_b$, even though both types reveal themselves, and $v_s^*(v_b) + \varepsilon < v_b$. This means that beneficial trade is knowingly denied under the allocation rule $p^*$, and therefore $p^*$ is inconsistent with credibility. We thus obtain:

**Proposition 5 (Upper bound)** *Suppose that the virtual valuation of the buyer, $v_b - \frac{1-F_b(v_b)}{f_b(v_b)}$, is increasing on $[\underline{v}_b, \bar{v}_b]$, and the virtual valuation of the seller, $v_s + \frac{F_s(v_s)}{f_s(v_s)}$, is increasing on $[\underline{v}_s, \bar{v}_s]$. Then, the ex-ante social surplus attained by the intermediary is strictly lower than that attained by the optimal trade mechanism.*

---

[28] The virtual valuation of the buyer is given by $v_b - \frac{1-F_b(v_b)}{f_b(v_b)}$ and that of the seller by $v_s + \frac{F_s(v_s)}{f_s(v_s)}$.

[29] Note that $\alpha = 0$ corresponds to the efficient outcome which is unattainable.

[30] In the proof, we show that there exist two buyer types $v_b' < v_b''$ for which $v_s^*(v_b')$ and $v_s^*(v_b'')$ are in the interior of $V_s$ . Then, for all types $v_b \in [v_b', v_b'']$ the function $v_s^*(v_b)$ can be implicitly defined by $B(v_s^*(v_b), v_b) = 0$. The function $v_s^*$ is continuous because $B$ is continuous, and it is strictly increasing on $[v_b', v_b'']$ because $B$ is strictly increasing in its first parameter and strictly decreasing in its second. Finally, it satisfies $v_s^*(v_b) < v_b$ because $B(x, x) > 0$ for all $x \in V_b \cap V_s$.

[31] For example, if $F_s$ and $F_b$ are both uniform distributions over $[0, 1]$, then $v_s^*(v_b) = v_b - 0.25$ for $v_b \in [0.25, 1]$ and $v_s^*(v_b) = 0$ otherwise.

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

**Proposition 6** *For any price $x \in V_s \cap V_b$, there is an intermediation mechanism $(\Gamma, \sigma)$ that implements the posted price $x$.*

The proof is immediate and goes by construction. Consider the message set $M = M_s \times M_b$ such that $M_i$ partitions agent $i$'s type space into two messages: $M_i = \{[\underline{v}_i, x], [x, \bar{v}_i]\}$, and the allocation rule $p$ and transfer rule $t$ that impose trade at price $x$ if the seller reports $m_s = [\underline{v}_s, x]$ and the buyer reports $[x, \bar{v}_b]$, and no trade or transfer otherwise. Thus, $m_s = [x, \bar{v}_s]$ and $m_b = [\underline{v}_b, x]$ 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:

**Corollary 1 (Lower bound)** *The ex-ante social surplus attained by the intermediary is weakly higher than that attained by the optimal posted price.*

## 5.3 The case of uniform distributions

In this section, we study the case of uniform distributions which can be solved analytically. We show how the credibility constraint increases the tension between surplus maximization and budget balance. In the uniform case, this tension leads to a negative result: the intermediary can do no better than the posted-price mechanism. A formal proof is given in the appendix; here we sketch the intuition of the core argument using an example. Note that while the exact calculations and the final result rely on the uniform priors assumption, the insights provided by this example carry over to the general case.

Let $F_s$ and $F_b$ be uniform and assume, by contradiction, that the intermediary could 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 the three representative examples depicted in Figure 3.[32] We will show that there is no intermediation mechanism that is ex-ante budget-balanced and has an allocation rule that corresponds to these examples. In what follows, the term *extreme vertex* will refer to the top left corner of a rectangle

[32] There is one additional case which is not represented in these examples and is discussed separately in the proof.

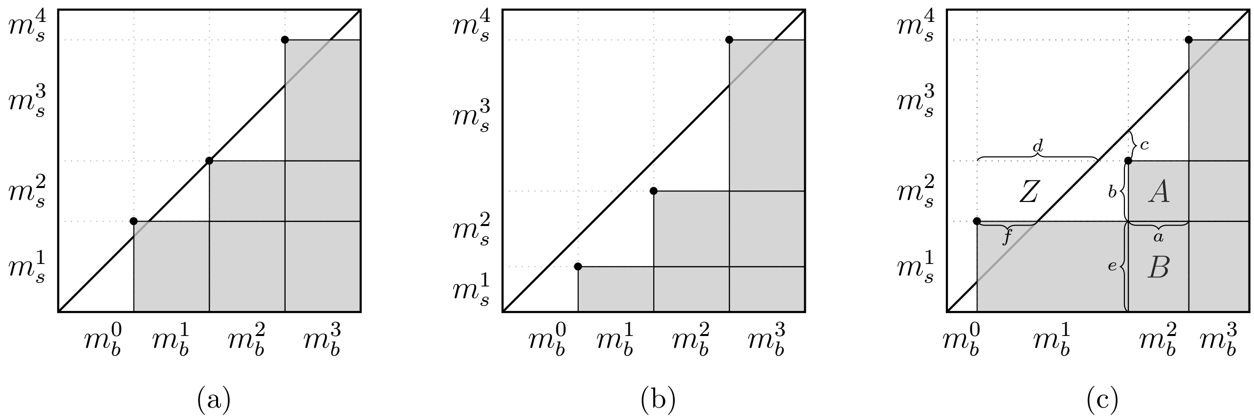


Figure 3: Allocation rules plotted on the agents' type space. If the agents' types are uniformly distributed, then none of the above representative allocation rules can be part of an intermediation mechanism.

that corresponds to a pair of messages $\left(m_s^k, m_b^k\right)$ for some $k$. In Figure 3, all the extreme vertices are marked by a small black dot.

Consider first the case in which $M$ and $p$ are such that all the extreme vertices are on, or above, the diagonal, as depicted in Figure 3(a). In this case, the information-rent burden of all the rectangles is positive, and therefore the value of $\psi\left(M, p\right)$ is strictly positive. Proposition 3 implies that $M$ and $p$ cannot be part of a (budget-balanced) intermediation mechanism.

Next, consider the case in which $M$ and $p$ are such that there are two consecutive extreme vertices below the diagonal, as depicted in Figure 3(b).[33] In this case, the allocation rule $p$ violates credibility: since the types are uniformly distributed, the mean buyer type that reports $m_b^1$ is strictly larger than the mean seller type that reports $m_s^2$. Therefore, imposing no trade when the seller reports $m_s^2$ and the buyer reports $m_b^1$, i.e. $p\left(m_s^2, m_b^1\right) = 0$, is inconsistent with credibility.

Finally, consider the case in which $M$ and $p$ are such that there is an extreme vertex above the diagonal followed by an extreme vertex below the diagonal, as depicted in Figure 3(c). We will show that because of credibility, the sum of the probability mass of rectangle $A$ multiplied by its information rent burden ("the negative summand"), plus the probability mass of rectangle $B$ multiplied by its information rent burden ("the positive summand") is positive. This implies that $\psi\left(M, p\right)$ is positive and therefore $M$ and $p$ cannot be part of a (budget-balanced) intermediation mechanism.

To see this, note that the "negative summand" is given by $\frac{b}{\bar{v}_s - \underline{v}_s} \frac{a}{\bar{v}_b - \underline{v}_b} \cdot (-c)$ whereas the positive summand is given by $d \cdot \frac{e}{\bar{v}_s - \underline{v}_s} \frac{a}{\bar{v}_b - \underline{v}_b}$, where $a, b, c, d$ and $e$ are the lengths of the

[33] The same argument applies if one of the vertices is on the diagonal and the other is strictly below it.

segments as marked in Figure 3(c). By credibility, the fact that there is no trade when the agents report $(m_s^2, m_b^1)$ implies that the rectangle $Z$ has more mass above the diagonal than below it, and therefore it must be that $f \geq c$ and $d \geq b + c > b$. Since $e > f$ we have that $d \cdot \frac{e}{\bar{v}_s - \underline{v}_s} \frac{a}{\bar{v}_b - \underline{v}_b} > c \cdot \frac{b}{\bar{v}_s - \underline{v}_s} \frac{a}{\bar{v}_b - \underline{v}_b}$.

The above computation illustrates the following more general property (that always holds when the agents have more than 3 intervals in their message sets, and in some cases, as in the example, also when they have 3 or less intervals): Given $M$ and $p$, for every rectangle that contributes a negative summand to $\psi$ (such as $A$ in the example) there are (distinct) adjacent rectangles that contribute positive summands (such as rectangle $B$ in the example) with strictly greater absolute value. By generalizing the above arguments and dealing separately with the cases of a small number of intervals (by proving directly that posting a price yields a higher expected surplus), we obtain the following result:

**Proposition 7** *Assume that $F_s$ and $F_b$ are uniform. Then, the intermediary implements the optimal posted price $x^*$, where*

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

## 5.4 The information-precision tradeoff

A trade mechanism consists of three components: a message set $M$, an allocation rule $p$ and a transfer rule $t$. The previous results make it possible to describe the difference between the problem faced by the designer of a (conventional) trade mechanism and that faced by the designer of an intermediation mechanism. Thus, for the conventional mechanism designer, the message set for each agent is given (and is identical to the agent's type set, by the revelation principle), and his task is to find the optimal allocation rule subject to the budget-balance constraint.[34] For the intermediation mechanism designer the task is reversed – he must optimally choose the message sets for the agents, which in turn pin down the allocation rule to induce only beneficial trade (by Proposition 3), subject to the budget-balance constraint.

Myerson and Satterthwaite (1983) show that, under mild conditions, achieving the efficient outcome is impossible. Thus, if the agents trade whenever the buyer values the

[34] The transfer rule $t$ that supports truth-telling is pinned down by $M$ and $p$ up to a constant, as long as the average trade probability for each agent is monotone in his type.

object more than the seller, then the social surplus generated by trade is not sufficiently large to cover the information rents that are required to incentivize the agents to truthfully report their types. To achieve the best feasible outcome, the conventional mechanism designer has to finely craft the allocation rule (see Equation 3) so as to impose trade in realizations for which the ratio between the expected social surplus they generate and the information rents they imply is above some threshold. Trade in all other realizations is denied, even if it generates positive social surplus, in order to save on the information rents.

The intermediation mechanism designer faces the same problem of insufficient resources. In contrast to the conventional designer, however, his ability to separate between realizations that provide high and low surplus-to-rent ratios is limited, since he communicates in a coarse language. This can easily be illustrated in the case of uniform distributions. Thus, consider the example depicted in Figure 3(c) and suppose that the designer wants to deny trade in realizations that correspond to the rectangle $Z$ and are below the diagonal (which generate a positive surplus but provide a low surplus-to-rent ratio). To credibly accomplish this, the designer must pool these realizations with others that are above the diagonal, where the surplus is negative, into one information set, so that trade on average is not profitable. For example, he can do so by pooling all buyer types in the interval $m_b^1$ and seller types in the interval $m_s^2$, thereby creating the information set $Z$. Note, however, that in this case the buyer types in $m_b^1$ are pooled together even when the seller reports $m_s^1$, which implies that trade takes place above the diagonal, namely in realizations that create a *negative* surplus.[35]

When the resource constraints are relaxed, the gap between the social surplus achieved by intermediation mechanisms and conventional mechanisms narrows. To see this, consider the case in which both agents' types are distributed uniformly over $[0, 1]$, and suppose that the designer can subsidize trade using an external source of money. As a benchmark, consider the case in which the subsidy is 0, which corresponds to Figure 4(a). By Proposition 7, the optimal intermediation mechanism implements the posted price $\frac{1}{2}$, whereas by Equation (3) the optimal conventional mechanism imposes trade whenever the buyer's value exceeds the seller's by more than 0.25 (as occurs in all the realizations below the dashed line in Figure 4(a)). A computation shows that the intermadiation mechanism achieves 89% of the surplus generated by the optimal conventional mechanism.[36]

[35] Alternatively, the designer could decrease the lower bound of $m_b^1$ so as to ensure that $\mathbb{E}_{v_b}\left[v_b|m_b^1\right] \leq \mathbb{E}_{v_b}\left[v_s|m_s^1\right]$ and then set $p\left(m_s^1, m_b^1\right) = 0$ which is consistent with credibility. In this case, however, the higher buyer types in $m_b^1$ stop trading, and realizations with high surplus-to-rent ratio are denied.

[36] The expected social surplus for the intermediation mechanism is $\int_{0.5}^{1}\int_{0}^{0.5}\left(v_b - v_s\right)dv_s dv_b = \frac{1}{8}$ and

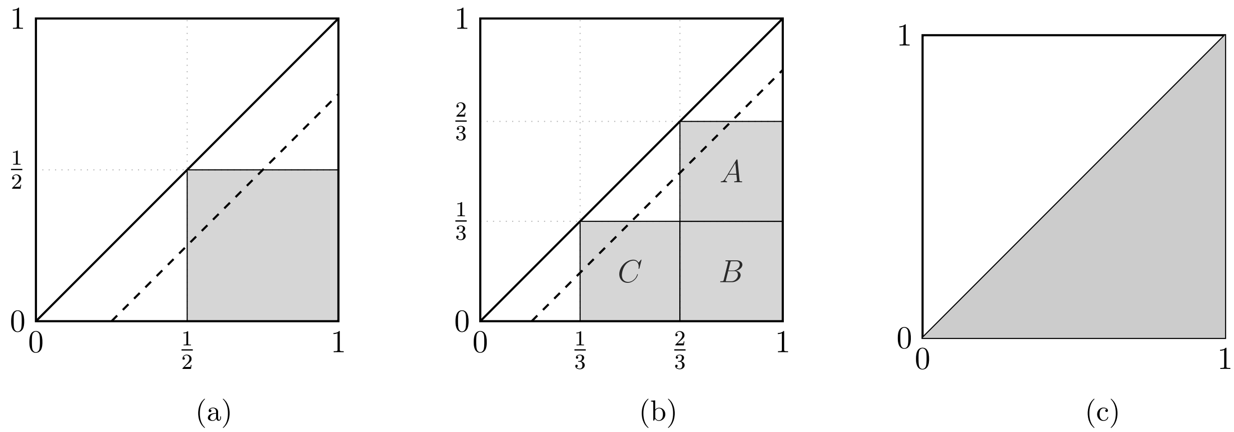


Figure 4: Bilateral trade with uniform distributions and a subsidy. The left-hand and right-hand panels show the allocation rules of the optimal intermediation mechanism with subsidy levels 0 and $\frac{1}{6}$, respectively. The middle panel shows a feasible (but not necessarily optimal) allocation rule for an intermediation mechanism with a subsidy of $\frac{1}{27}$.

Now consider the case in which the (expected) subsidy is $\frac{1}{27}$, which corresponds to Figure 4(b). It is not difficult to verify that there exists an intermediation mechanism with the depicted allocation rule (although note that this is not the optimal intermediation mechanism but rather just a feasible one).[37] The dashed line corresponds to the optimal conventional mechanism in which trade takes place whenever the buyer's valuation is higher than the seller's by more than ˜0.17.[38] The intermadiation mechanism now achieves 96% of the surplus generated by the optimal conventional mechanism.

Finally, consider a subsidy of $\frac{1}{6}$. This subsidy is sufficiently large that both the conventional mechanism and the intermediation mechanism can attain the first-best (efficient) outcome, which is depicted in Figure 4(c). In this case, the ratio between the surpluses generated by the mechanisms is exactly 1.

Comparing the surplus gain due to subsidizing trade in an intermediation mechanism to that in a conventional mechanism shows that a subsidy is more effective in the former case. These findings are summarized in Table 1. A subsidy of $\frac{1}{27}$ increases the surplus of

for the conventional mechanism is $\int_{0.25}^{1}\int_{0}^{v_b-0.25}(v_b-v_s)\,dv_s dv_b=\frac{9}{64}$

[37] One way to support the allocation rule depicted in Figure 4(b) with transfers such that the agents report truthfully is the following: In rectangle $A$, the price is $\frac{5}{9}$, in rectangle $B$ it is $\frac{4}{9}$ and in rectangle $C$ it is $\frac{3}{9}$. 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 $\frac{1}{27}$).

[38] From Equation 3, we know that trade takes place whenever $v_b - v_s > \beta$ for some $\beta \in (0,1)$. We then solve $\int_{\beta}^{1}\int_{0}^{v_b-\beta}(v_b-v_s)\,dv_s dv_b+\frac{1}{27}=\int_{\beta}^{1}\left(\int_{\beta}^{v_b}(x-\beta)\,dx\right)dv_b+\int_{0}^{1-\beta}\left(\int_{v_s}^{1-\beta}(1-\beta-x)\,dx\right)dv_s$, 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 $\beta =$ ˜0.17.

| | **Surplus** ($\times 100$) | | | **Surplus gain due to subsidy** ($\times 100$) | | **Ratio of surplus gain to subsidy** | |
|---|---|---|---|---|---|---|---|
| **Subsidy** | Intermed. | Mech. | Ratio | Intermed. | Mech. | Intermed. | Mech. |
| 0 | 12.5 | 14.06 | 89% | | | | |
| $\frac{1}{27}$ | 14.81 | 15.39 | 96% | 2.31 | 1.33 | 62% | 36% |
| $\frac{1}{6}$ | 16.67 | 16.67 | 100% | 4.17 | 2.6 | 25% | 16% |

Table 1: **Subsidizing trade for intermediaries and mechanisms.** All values are computed for the case in which agents' valuations are uniformly distributed over $[0, 1]$. The first column is the level of the subsidy. The second and third columns are the expected social surpluses (a lower bound for the intermediary with subsidy 1/27), and the fourth column is the ratio between the two. The fifth and sixth columns are the surplus gain, relative to the case of a zero subsidy. The seventh and eighth columns are the ratio between the surplus gain and the subsidy.

the optimal mechanism by $1.33 \times 10^{-2}$ (surplus gain to subsidy ratio of 36%) while 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%) while that of the intermediation mechanism increases 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.

# Appendix

## Proof of Proposition 1

**Part I**: Suppose that some social choice function $scf$ is implementable in the reporting sub-game that starts with message set $M$. Denote the equilibrium strategies of the agents (in stage 2) by $\sigma' = (\sigma'_s, \sigma'_b)$ and the strategy of the intermediary (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 $i$.

Consider the trade mechanism $(\Gamma = \langle M, p, t\rangle, \sigma')$ where $p(m) = p'(m)$ and $t(m) = t'(m)$ whenever $(p', t')$ are defined (that is, if neither $m_s$ nor $m_b$ in the report $m = (m_s, m_b)$ are the message labeled "out"), and $p(m) = t(m) = 0$ 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 $m = (m_1, m_2)$ on the equilibrium path, and unless either $m_1$ or $m_2$ is an opt-out message (implying $p(m) = t(m) = 0$), 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 $scf$.

Conversely, suppose that $scf$ is implemented by an intermediation mechanism $\Gamma = ((M, p, t), \sigma')$. Label every opt-out message in the mechanism by "out". Then, having the agents' play $\sigma'_1, \sigma'_2$ in stage 2 and the intermediary play $(p, t)$ in stage 3 is an equilibrium of the sub-game that starts with $M$: For any $m$ on the equilibrium path, the intermediary's strategy is optimal since $\Gamma$ satisfies credibility, and for any $m$ off the equilibrium path there exists a belief that makes $p(m)$ optimal[39]; and $\sigma'_1, \sigma'_2$ are an equilibrium in stage 2. Thus, $scf$ is implemented by the reporting sub-game.

**Part II**: By part 1, the set of social choice functions that are implementable by all intermediation mechanisms is the same as the set of social choice functions that are implementable by all possible reporting sub-games. Note that: (1) the optimal intermediation mechanism is the one that solves $\max_{(\langle M,p,t\rangle,\sigma)} W_{EA}$ ,where $(\langle M, p, t\rangle, \sigma)$ is any intermediation mechanism, and that (2) by the refinement that the intermediary-optimal equilibrium is played in any sub-game, the outcome of the entire game is the

[39]Note that off-equilibrium beliefs are unconstrained and that types spaces of the buyer and the seller overlap. Thus, for example, $p(m) = 1$ is supported by the belief that the buyer is of a type higher than that of the seller.

one that solves $\max_M \max_{(\sigma,p,t)_M} W_{EA}$, where $(\sigma, p, t)_M$ is any equilibrium in the reporting sub-game starting with $M$. Clearly, both the direct maximization and the two-step maximization yield the same $scf$.

## Proof of Lemma 1

Let $\Gamma = \langle M, p, t\rangle$ be an intermediation mechanism, and suppose 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 $m_b \in M_b$. By credibility it must be the case that either $m_s'$ is an opt-out message or $\mathbb{E}_{v_s,v_b}[(v_b - v_s) \,|\, (m_s, m_b)] > \mathbb{E}_{v_s,v_b}[(v_b - v_s) \,|\, (m_s', m_b)]$. In the first case we are done, because if $m_s'$ is an opt-out message, then $p(m_s', m_b') = 0$ for all $m_b' \in M_b$, by definition. Otherwise $\mathbb{E}_{v_s}[v_s|m_s] < \mathbb{E}_{v_s}[v_s|m_s']$ since types are independent, and therefore $\mathbb{E}_{v_s,v_b}[(v_b - v_s) \,|\, (m_s, m_b')] > \mathbb{E}_{v_s,v_b}[(v_b - v_s) \,|\, (m_s', m_b')]$ for any $m_b' \neq m_b$. By credibility, we then have that $p(m_s, m_b') \geq p(m_s', m_b')$, where the inequality is weak because $m_b'$ can be an opt-out message. The proof for the buyer side is similar.

## Proof of Lemma 2

Suppose that $(\Gamma = \langle M, p, t\rangle, \sigma)$ is an intermediation mechanism in which some agent $i$ has two messages $m_i'$ and $m_i''$ such that $\bar{p}_i(m_i') = \bar{p}_i(m_i'')$. Suppose further that both $m_i'$ and $m_i''$ are sent by at least one type of agent $i$ in equilibrium (that is $\{v_i \mid m_i' \in supp[\sigma_i(v_i)]\}$ and $\{v_i \mid m_i'' \in supp[\sigma_i(v_i)]\}$ are non-empty), since otherwise we just drop the message that is not sent by any type and we are done.

We will construct a new intermediation mechanism $\left(\hat{\Gamma}, \hat{\sigma}\right)$ in which $m_i'$ and $m_i''$ are merged, and all types of both agents expect the same payoff in both mechanisms. To avoid confusion in evaluating conditional expectations, we will add a superscript to the expected value operator indicating 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 $\sigma$).

We begin by making two useful observations:

1. Message-monotonicity of $\Gamma$ implies that either $\mathbb{E}^{\sigma}_{v_i}[v_i|m_i'] = \mathbb{E}^{\sigma}_{v_i}[v_i|m_i'']$ or $p(m_i', m_{-i}) = p(m_i'', m_{-i})$ for every $m_{-i} \in M_{-i}$ (otherwise it cannot be that $\bar{p}_i(m_i') = \bar{p}_i(m_i'')$).

2. Denote by $\bar{t}_i(m_i) = \mathbb{E}^{\sigma}_{m_{-i}} t(m_i, m_{-i})$ the expected monetary transfer from the buyer to the seller when agent $i$ sends a message $m_i$. Since both $m_i'$ and $m_i''$ are being used in equilibrium, and since $\bar{p}_i(m_i') = \bar{p}_i(m_i'')$, then it must be the case that $\bar{t}_i(m_i') = \bar{t}_i(m_i'')$.

Consider a new mechanism, denoted as ($\hat{\Gamma} = \left\langle \hat{M}, \hat{p}, \hat{t} \right\rangle, \hat{\sigma}$), which is identical to $(\Gamma, \sigma)$ (the "original mechanism"), with the following three modifications:
(i) The messages $m_i'$ and $m_i''$ in the original mechanism are replaced by a single message $\hat{m}_i$ in the new mechanism. Therefore $\left|\hat{M}_i\right| < |M_i|$.
(ii) When agent $i$ sends the message $\hat{m}_i$ in the new mechanism $\hat{\Gamma}$, the monetary transfer between the buyer and the seller and the probability of trade are set to be equal to the weighted sum of those induced by $m_i'$ and $m_i''$ in the original mechanism:

$$\begin{aligned}
\hat{t}(\hat{m}_i, m_{-i}) &= \frac{\Pr(m_i')}{\Pr(m_i') + \Pr(m_i'')} t(m_i', m_{-i}) + \frac{\Pr(m_i'')}{\Pr(m_i') + \Pr(m_i'')} t(m_i'', m_{-i}) \\
\hat{p}(\hat{m}_i, m_{-i}) &= \frac{\Pr(m_i')}{\Pr(m_i') + \Pr(m_i'')} p(m_i', m_{-i}) + \frac{\Pr(m_i'')}{\Pr(m_i') + \Pr(m_i'')} p(m_i'', m_{-i})
\end{aligned}$$

for all $m_{-i} \in M_{-i}$, where $\Pr(m_i)$ is the probability that message $m_i$ is sent in the equilibrium of the original mechanism (that is, when agent $i$ plays according to the strategy $\sigma_i$ in $\Gamma$).
(iii) All types who sent $m_i'$ and $m_i''$ 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 $\bar{\hat{p}}(\hat{m}_i) = \bar{p}(m_i') = \bar{p}(m_i'')$ and $\bar{\hat{t}}(\hat{m}_i) = \bar{t}(m_i') = \bar{t}(m_i'')$. 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: $\bar{u}_i^{(\Gamma,\sigma)}(v_i, m_i) = \bar{u}_i^{\left(\hat{\Gamma},\hat{\sigma}\right)}(v_i, \hat{m}_i)$ for every $v_i \in V_i$ and $m_i \in supp[\sigma_i(v_i)]$ and $\hat{m}_i \in supp[\hat{\sigma}_i(v_i)]$ and for each agent $i$.

It remains to verify that $(\hat{\Gamma}, \hat{\sigma})$ satisfies credibility. To do so it suffices to check that $\hat{p}(\hat{m}_i, m_{-i})$ equals 1 (equals 0) when $\mathbb{E}^{\hat{\sigma}}_{v_s,v_b}(v_b - v_s|\hat{m}_i, m_{-i})$ is positive (negative). This is because for all other message pairs $(m_s, m_b) \in M_s \times M_b$, the expected gains from trade, conditional on $(m_s, m_b)$ being reported, and the allocation decision $p(m_s, m_b)$ is identical in $(\Gamma, \sigma)$ and $\left(\hat{\Gamma}, \hat{\sigma}\right)$.

For any $m_{-i}$, denote $a_1 \equiv \mathbb{E}^{\sigma}_{v_s,v_b}[v_b - v_s|(m_i', m_{-i})]$ and $a_2 \equiv \left[\mathbb{E}^{\sigma}_{v_s,v_b}(v_b - v_s|(m_i'', m_{-i}))\right]$ and without loss of generality assume that $a_1 \geq a_2$. Note that $\mathbb{E}^{\hat{\sigma}}_{v_s,v_b}(v_b - v_s|\hat{m}_i, m_{-i}) \in (a_1, a_2)$, since all the types who sent $m_i'$ and $m_i''$ in the original mechanism $\Gamma$ send $\hat{m}_i$ in the new mechanism $\hat{\Gamma}$.

Suppose first that $a_1 = a_2$. If $a_1 > 0$, then by credibility of $\Gamma$ we have that $p(m_i', m_{-i}) = p(m_i'', m_{-i}) = 1$, and therefore $\hat{p}(\hat{m}_i, m_{-i}) = 1$, as required. Similarly, if $a_1 < 0$, then $p(m_i', m_{-i}) = p(m_i'', m_{-i}) = 0$ and therefore $\hat{p}(\hat{m}_i, m_{-i}) = 0$, as required. If $a_1 = 0$, then any value of $\hat{p}(\hat{m}_i, m_{-i})$ in consistent with credibility. Therefore $(\hat{\Gamma}, \hat{\sigma})$ is credible.

Suppose alternatively that $a_1 > a_2$. By observation 1 above (and by the way in which $\hat{p}(\hat{m}_i, m_{-i})$ is constructed), it must be the case that $p(m_i', m_{-i}) = p(m_i'', m_{-i}) = \hat{p}(\hat{m}_i, m_{-i})$ for all $m_{-i}$. It must also be the case that either $a_1 > a_2 \geq 0$ or $0 \geq a_1 > a_2$ (because if $a_1 > 0 > a_2$ then $p(m_i', m_{-i}) = p(m_i'', m_{-i})$ is not consistent with credibility of the original mechanism $\Gamma$). We then have that: If $a_1 > a_2 \geq 0$, then $p(m_i', m_{-i}) = p(m_i', m_{-i}) = 1$ by credibility of $\Gamma$, and therefore $\hat{p}(\hat{m}_i, m_{-i}) = 1$, as required; Similarly, if $a_2 < a_1 \leq 0$, then $p(m_i', m_{-i}) = p(m_i', m_{-i}) = 0$ by credibility of $\Gamma$, and therefore $\hat{p}(\hat{m}_i, m_{-i}) = 0$, as required. Therefore, $(\hat{\Gamma}, \hat{\sigma})$ is credible.

## Proof of Lemma 4

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

We begin by defining the transfer rule $t(m_s, m_b)$ as follows:

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

where $d(m_s, m_b) = t_s(m_s, m_b) + t_b(m_s, m_b)$. 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, conditional on both agents reporting truthfully, the expected payment for each agent from sending any message $m_i \in M_i$ in both mechanisms is the same: (i) $\mathbb{E}_{m_b} t(m_s, m_b) = \mathbb{E}_{m_b} t_s(m_s, m_b)$ for every message $m_s \in M_s$, and (ii) $\mathbb{E}_{m_s} t(m_s, m_b) = -\mathbb{E}_{m_s} t_b(m_s, m_b)$ for every message $m_b \in M_b$. Therefore, since truth-telling is an equilibrium in $\Gamma$ it is also an equilibrium in $\Gamma'$.

Next, suppose that $m_i'$ is an opt-out message for agent $i$ in $\Gamma$, that is $p(m_i', m_{-i}) =$

$t_i(m_i', m_{-i}) = 0$ for all $m_{-i}$, and therefore $\mathbb{E}_{m_{-i}} t_i(m_i', m_{-i}) = \mathbb{E}_{m_{-i}} t(m_i', m_{-i}) = 0$. Note that it could be the case that $t(m_i', m_{-i})$, 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 pick any message $m_i'' \neq m_i'$, which is not an opt-out message in $\Gamma$ and update the transfer rule $t$ as follows: for all $m_{-i}$ we increase $t(m_i'', m_{-i})$ by $\frac{\Pr(m_i')}{\Pr(m_i'')} \cdot t(m_i', m_{-i})$ and set $t(m_i', m_{-i})$ to be zero. Note that since $\mathbb{E}_{m_{-i}} t(m_i', m_{-i}) = 0$, the monetary transfer that each agent expects following each of his possible messages is unchanged. Note also that $m_i'$ is now an opt-out message (zero probability for trade and zero transfers for all $m_{-i}$). Since the message set, and the allocation rule, and the set of opt-out messages is the same in $\Gamma$ and $\Gamma'$, then $\Gamma'$ is credible. We then have that $\Gamma' = \langle M, p, t \rangle$ is a partition-direct intermediation mechanism, as desired.

## Proof of Proposition 3

Suppose that $M = M_s \times M_b$ is a message set such that $M_i$ consists of intervals that partition $V_i$ for each agent $i$, and $p$ is an allocation rule that satisfies only beneficial trade with respect to $M$. We will first show that when $\psi(M, p) = 0$ 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 auxiliary intermediation mechanism (as defined in section 4.2). Then, by Lemma 4, there exists $t : M \to R$ such that $\langle M, p, t \rangle$ is an (ex-post budget-balanced) intermediation mechanism.

We define the following two transfer rules : $t_s(m_s, m_b) = \omega_s(m_b)$ and $t_b(m_s, m_b) = -\omega_b(m_s)$ if $p(m_s, m_b) = 1$, and $t_s(m_s, m_b) = t_b(m_s, m_b) = 0$ otherwise, where $t_i(m_s, m_b)$ is the monetary transfer (which may be negative) to agent $i$ when the pair of messages $(m_s, m_b)$ is reported.

The mechanism $\Gamma = \langle M, p, t_s, t_b \rangle$, along with its truth-telling equilibrium, is an auxiliary intermediation mechanism. To see this, note that credibility of $\Gamma$ is satisfied because $p$ satisfies only beneficial trade with respect to $M$. Individual rationality is satisfied because, for every message $m_s \in M_s$, a buyer of type $v_b$ who (truthfully) reports $m_b$ such that $v_b \in m_b$ pays $\omega_b(m_s)$ *if the object is traded*, which is lower than $v_b$ (recall that $\omega_b(m_s)$ is by definition the *lowest* buyer type that trades with $m_s$). 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 by the message of agent $-i$ 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 non-profitable price. Finally, the expected budget deficit of the mechanism is given by:

$$\sum_{(m_s,m_b)\in M} [t_s(m_s,m_b) + t_b(m_s,m_b)] \cdot \Pr(m_s) \cdot \Pr(m_b)$$

which is exactly $\psi(M,p)$. Thus, when $\psi(M,p) = 0$ the auxiliary intermediation mechanism $\Gamma$ is ex-ante budget-balanced, and therefore, by Lemma 4, there exists $t$ such that $\Gamma' = \langle M, p, t\rangle$ is an intermediation mechanism.

Consider now the case in which $\psi(M,p) > 0$, which implies that the mechanism $\Gamma$ defined above creates a budget deficit. A well-known result in mechanism design (see e.g. Krishna (2010)) 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 who is most reluctant to trade of that agent ($\bar{u}_s(\bar{v}_s)$ 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 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 money to the seller or collects (in expectation) more money from the buyer. In other words, any other mechanism creates a (weakly) higher budget deficit. Thus, when $\psi(M,p) > 0$ there exists no (budget-balanced) intermediation mechanism with message set $M$ and allocation rule $p$.

## Proof of Proposition 4

Suppose first, by way of contradiction, that $\Gamma = (M,p,t)$ is an optimal (partition-direct) intermediation mechanism, in which the allocation rule $p$ satisfies only beneficial trade with respect to $M$, but $\psi(M,p) < 0$. We will show that we can modify $M$ and $p$ to induce more (and only) beneficial trade, in a way that increases $\psi(M,p)$ to a zero. Then, by Proposition 3, there exists an intermediation mechanism that achieves a higher surplus than $\Gamma$, a contradiction.

Since $\psi(M,p) < 0$ there must be some $k \geq 1$ for which $\underline{m}_b^k > \bar{m}_s^k$. To see this, note that for any two intervals, $m_s^l \in M_s$ and $m_b^k \in M_b$, we have that $p\left(m_s^l, m_b^k\right) = 1$ if and only if $l \leq k$ (see property STEP-FORM in section 4.1). Therefore there must be some $k \geq l \geq 1$ such that $\omega_b\left(m_s^l\right) > \omega_s\left(m_b^k\right)$, or equivalently $\underline{m}_b^l > \bar{m}_s^k$. Since by definition $\underline{m}_b^k > \underline{m}_b^l$ then $\underline{m}_b^k > \bar{m}_s^k$. If there is more than one index $k$ with this property, pick the lowest one.

Consider the following modification for $M$ and $p$ (illustrated in Figure 2(c), in which

$k = 2$): increase $\bar{m}_s^k$ by $\Delta$ and decrease $\underline{m}_b^k$ by $\Delta$ and update the allocation rule $p$ to satisfy only beneficial trade with respect to the modified set $M$. The size of $\Delta > 0$ is the lowest value for which either $\bar{m}_s^k = \underline{m}_b^k$ or $\psi(M, p) = 0$ (we discuss below why such a value of $\Delta$ exists).

Note first that the modification increases the total expected surplus from trade. To see this, it is useful to think about the modification in two steps: First, the effect on the total surplus when we only change $\underline{m}_b^k$ and $\bar{m}_s^k$ (but do not update $p$) is positive, because trade is added in realizations in which it is beneficial (recall that $\underline{m}_b^k > \bar{m}_s^k$). Next, the effect of updating $p$ to satisfy only beneficial trade cannot, by definition, decrease the expected surplus.

Next, note that when we change $\bar{m}_s^k$ and $\underline{m}_b^k$ as described above we increase the means of the intervals $m_s^k$ and $m_s^{k+1}$ and decrease those of the intervals $m_b^k$ and $m_b^{k-1}$. Consequently, there is no pair of intervals $(m_s, m_b)$, other than $\left(m_s^k, m_b^k\right)$, for which trade on average became more beneficial, and therefore when we update the allocation rule, the probability for trade $p(m_s, m_b)$ does not increase for any $(m_s, m_b)$. It follows that *if* the modification increases $\psi(M, p)$, then $\psi(M, p)$ increases continuously in $\Delta$, so it must be the case that there exists $\Delta$ such that either $\bar{m}_s^k = \underline{m}_b^k$ or $\psi(M, p) = 0$.

If after the modification $\psi(M, p) = 0$, then we are done, because according to Proposition 3 there exists an intermediation mechanism with the modified message set $M$ and allocation rule $p$ that generates higher ex-ante surplus than $\Gamma$, a contradiction. If $\psi(M, p) < 0$, we reiterate the process. Note that since the number of messages for each agent is finite, the process must stop.

Suppose now, by way of contradiction, that $\Gamma = (M, p, t)$ is an optimal intermediation mechanism with an allocation rule $p$ that satisfies credibility, but does *not* satisfy only beneficial trade with respect to $M$. That is, there exist (one or more) pairs of intervals, $m_s$ and $m_b$, such that the means of $m_s$ and $m_b$ are equal, but $p(m_s, m_b) > 0$. By modifying the mechanism and setting $p(m_s, m_b) = 0$ for all such pairs, we do not change the mechanism's ex-ante expected surplus, and we make $p$ satisfy only beneficial trade. It must be the case, however, that after the modification $\psi(M, p) < 0$, since otherwise the original mechanism $\Gamma$ could not have been budget balanced. This contradicts the optimality of $\Gamma$ by the first part of the proof.

## Proof of Proposition 5

We begin by showing that there exist two distinct buyer types, $v_b'$ and $v_b''$, both in the interior of $V_b$ and two seller types,$v_s'$ and $v_s''$, both in the interior of $V_s$, such that $B(v_s', v_b') =$

$B(v_s'', v_b'') = 0$. This is because $B(\underline{v}_s, \bar{v}_b) < 0$ and $B(x,x) > 0$ for any $x \in V_s \cap V_b$ and therefore, by the intermediate value theorem, for any $x \in V_s \cap V_b$, there is a point $(v_b, v_s)$ on the line segment $v_s = \underline{v}_s + \frac{x - \underline{v}_s}{\bar{v}_b - x} \cdot (\bar{v}_b - v_b)$ that connects the point $(\underline{v}_s, \bar{v}_b)$ with the point $(x,x)$ for which $B(v_b, v_s) = 0$.

Next, we define the function $v_s^*(v_b)$ implicitly as follows: $B(v_s^*(v_b), v_b) = 0$ for all $v_b \in [v_b', v_b'']$. Since $B(v_s, v_b)$ is continuous, strictly increasing in $v_s$, strictly decreasing in $v_b$ and has bounded partial derivatives (because $f_s$ and $f_b$ are bounded away from zero) then $\frac{dv_s^*(v_b)}{dv_b}$ is strictly positive and finite over $[v_b', v_b'']$. Recall that $v_s^*(v_b)$ is the highest seller type with which the buyer type $v_b$ trades under $p^*$. Since $v_s^*(v_b)$ is strictly increasing then each buyer $v_b \in [v_b', v_b'']$ trades with a different subset of seller types. It follows that in order to implement $p^*$ the designer must separate all the buyer types in $[v_b', v_b'']$ and all the seller types in $[v_s^*(v_b'), v_s^*(v_b'')]$, that is, to know the agents' exact types in equilibrium.

Finally, note that because $B(x,x) > 0$ for all $x \in V_s \cap V_b$, and because $B(v_s^*(v_b), v_b) = 0$ for all $v_b \in [v_b', v_b'']$, then it must be the case that $v_s^*(v_b) < v_b$ for all $v_b \in [v_b', v_b'']$. Therefore, pick any buyer type $v_b \in [v_b', v_b'']$ and note that there exists small enough $\varepsilon > 0$ such that $v_s^*(v_b) + \varepsilon < v_b$ but $p^*(v_s^*(v_b) + \varepsilon, v_b) = 0$. This implies that $p^*$ is inconsistent with credibility, and therefore an intermediation mechanism cannot implement the allocation rule of the second-best (conventional) mechanism.

## Proof of Proposition 7

Suppose that $F_b$ and $F_s$ are uniform. Given a partition-direct intermediation mechanism $\Gamma = (M, p, t)$ we denote by $K$ the index of the highest interval of the buyer, that is $K = \max\{k : m_b^k \in M_b\}$.[40] We also denote $\phi(m_s^l, m_b^k) = (\underline{m}_s^{k+1} - \underline{m}_b^l) \cdot \frac{\underline{m}_b^{k+1} - \underline{m}_b^k}{\bar{v}_b - \underline{v}_b} \cdot \frac{\underline{m}_s^{l+1} - \underline{m}_s^l}{\bar{v}_s - \underline{v}_s}$,[41] and since $p(m_s^l, m_b^k) = 1$ if and only if $k \geq l$, then we have that $\psi(M, p) = \sum_{k=1}^{K} \sum_{l=1}^{k} \phi(m_s^l, m_b^k)$.

The proof has two parts. 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 5(a) (and formally characterized by Lemma 7). The second part 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

[40] Recall that the enumeration starts from $k = 0$ if there is no trade when the buyer reports the lowest interval, and from $k = 1$ otherwise (see Section 4.1)

[41] Recall that by definition $\bar{m}_i^l = \underline{m}_i^{l+1}$ for all $l$ and for each agent $i$.

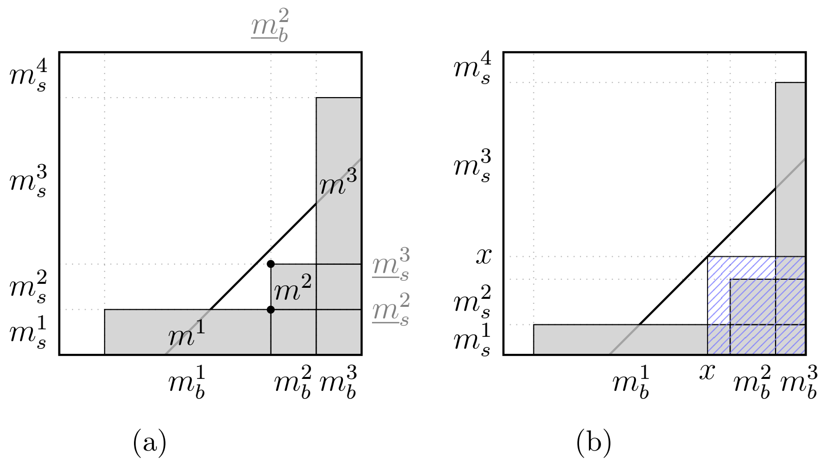


Figure 5

case that $K = 1$, and since the optimal posted price generates the highest surplus among all intermediation mechanisms with $K = 1$, we get the desired result.

**Part I : The core argument**

We begin with several definitions that will allow us to give a geometric interpretation of the proof. For any $k \in \{1, ..., K\}$, we denote $m^k = \left(m_s^k, m_b^k\right)$ and refer to each $m^k$ as an *extreme rectangle* of trade. The rectangles $m^k$ and $m^{k+1}$ are called *consecutive* extreme rectangles. We refer to $\left(\underline{m}_s^{k+1}, \underline{m}_b^k\right)$ as the *top-left corner* of $m^k$ and to $\left(\underline{m}_s^k, \underline{m}_b^k\right)$ as the *bottom-left corner* of $m^k$. 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.[42] Finally, we define $TLB$ to be the set of extreme rectangles with top-left corner below the diagonal, and $BLB$ to be the set of extreme rectangles with bottom-left corner below the diagonal:[43]

$$\begin{aligned} TLB &= \left\{\left(m_s^k, m_b^k\right) \in M : \underline{m}_s^{k+1} < \underline{m}_b^k\right\} \\ BLB &= \left\{\left(m_s^k, m_b^k\right) \in M : \underline{m}_s^k < \underline{m}_b^k\right\} \end{aligned}$$

In the example illustrated in Figure 5(a), the top-left and bottom-left corners of the extreme rectangle $m^2$ are marked with small black circles. 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 $BLB$.

Our first lemma establishes that, due to credibility, there are no two consecutive

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

[43] We say that a point $(x, y)$ is below the diagonal if $x > y$.

extreme rectangles with their top-left corner below the diagonal and no two consecutive extreme rectangles with their bottom-left corner above the diagonal:

**Lemma 5** *For any $k \in \{1, ..., K-1\}$:*

1. *If $m^k \in TLB$ then $m^{k+1} \notin TLB$;*
2. *If $m^k \notin BLB$ then $m^{k+1} \in BLB$.*

**Proof.** First, suppose that $m^k \in TLB$ and $m^{k+1} \in TLB$ for some $k \in \{1, ..., K-1\}$. 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 $\mathbb{E}\left[v_b : v_b \in m_b^k\right] > \mathbb{E}\left[v_s : v_s \in m_s^{k+1}\right]$. This implies that when the buyer reports $m_b^k$ and the seller reports $m_s^{k+1}$ trade is beneficial, which is inconsistent with $p\left(m_s^{k+1}, m_b^k\right) = 0$ due to credibility.

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

If $K > 1$ and $\psi(M, p) = 0$, then there is at least one extreme rectangle $m^k$ that belongs to the set $TLB$. This is because extreme rectangles with their top-left corner below the diagonal are the only ones that contribute negative summands to $\psi$ (i.e. they are the only rectangles with trade for which the value of $\phi$ is negative). Our next lemma asserts, however, 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 diagonal, then there are positive terms in $\psi$ that outweigh the negative value that is contibuted by $\phi\left(m^k\right)$. The geometric interpretation and intuition of this result are outlined in the main text.

**Lemma 6** *For any $k \in \{1, ..., K\}$ and any $m^k \in TLB$:*

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

**Proof**. Suppose $m^k \in TLB$ and $m^{k-j} \in BLB$ for some $1 \leq j < k-1$. We then have that:

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

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 $\left(\underline{m}_s^{k+1} - \underline{m}_b^{k-1}\right) \cdot \left(\underline{m}_s^{k} - \underline{m}_s^{k-j}\right) > -\left(\underline{m}_s^{k+1} - \underline{m}_b^{k}\right) \cdot \left(\underline{m}_s^{k+1} - \underline{m}_s^{k}\right)$.

To see this, note first that $\underline{m}_s^{k+1} - \underline{m}_b^{k-1} > \underline{m}_s^{k+1} - \underline{m}_s^{k}$. This is because $m^k \in TLB$ implies $m^{k-1} \notin TLB$ (according to Lemma 5), and therefore $\underline{m}_b^{k-1} < \underline{m}_s^{k}$. Next, note that that $\underline{m}_s^{k} - \underline{m}_s^{k-j} > \underline{m}_b^{k} - \underline{m}_s^{k+1}$. This is because $p\left(m_s^k, m_b^{k-1}\right) = 0$ 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 $m^{k-j} \in BLB$ 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.■

The third lemma characterizes the structure of any intermediation mechanism with $K > 1$ and $\psi(M,p) \leq 0$. The intuition is as follows: unless the mechanism satisfies the specific structure (which is qualitatively illustrated in Figure 5(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 $m^k = \left(m_s^k, m_b^k\right)$ and the other elements of the set is *positive*. This would be a contradiction to $\psi(M,p) \leq 0$.

**Lemma 7** *If $\Gamma = (M,p,t)$ is an intermediation mechanism, then $K \leq 3$. Furthermore, if $K > 1$ it must be that either:*
*1. $K = 2$ and $m^1 \in TLB$ and $m^3 \notin BLB$,[44] or*
*2. $K = 2$ and $m^2 \in TLB$ and $m^1 \notin BLB$, or*
*3. $K = 3$ and $m^2 \in TLB$ and $m^1 \notin BLB$ and $m^4 \notin BLB$.*

**Proof**. *Case I*: Suppose that $K > 3$. For any extreme rectangle $m^k \in TLB$ with index $k \geq 3$, we associate $m^k$ with the rectangle $\left(m_s^{k-1}, m_b^k\right)$ if $m^{k-1} \in BLB$, and with the rectangles $\left(m_s^{k-1}, m_b^k\right)$ and $\left(m_s^{k-2}, m_b^k\right)$ otherwise (note that according to Lemma 5 if $m^{k-1} \notin BLB$, then it must be that $m^{k-2} \in BLB$). For any extreme rectangle $m^k \in TLB$ with index $k \leq 2$, we associate $m^k$ with the rectangle $\left(m_s^k, m_b^{k+1}\right)$ if $m^{k+2} \in BLB$, and with the rectangles $\left(m_s^k, m_b^{k+1}\right)$ and $\left(m_s^k, m_b^{k+2}\right)$ otherwise (according to Lemma 5 if $m^{k+2} \notin BLB$ then it must be that $m^{k+3} \in BLB$).[45] Since by Lemma 5 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 6 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 is contradicts $\Gamma$ being an intermediation mechanism.

---

[44] 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 $\bar{v}_s > \bar{v}_b$.

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

*Case II*: Suppose that $K = 2$. Since $\psi(M,p) \leq 0$ it must be that either $m^1 \in TLB$ or $m^2 \in TLB$, but not both (since according to Lemma 5 there are no two consecutive rectangles with a top-left corner below the diagonal). Suppose $m^1 \in TLB$. If $m^3 \in BLB$, then according to Lemma 5 we have that $\phi\left(m_s^k, m_b^k\right)+\phi\left(m_s^k, m_b^{k+1}\right) > 0$, which contradicts $\psi(M,p) \leq 0$. Therefore it must be that $m^3 \notin BLB$, or equivalently $\underline{m}_s^3 > \bar{v}_b$. The argument for the case in which $m^2 \in TLB$ is similar.

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

**Part 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 that 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 7. 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 5(a).[46] The proof for the other two cases with $K = 2$ is similar.

Since $\Gamma$ is optimal it must be that:

$$\underline{m}_b^1 + \underline{m}_b^2 = \underline{m}_s^2 + \underline{m}_s^3 \tag{4}$$

and

$$\underline{m}_b^2 + \underline{m}_b^3 = \underline{m}_s^3 + \underline{m}_s^4, \tag{5}$$

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 non-beneficial trade.[47] 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

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

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

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 5(b), is then given by $S(x) \equiv \frac{x-\underline{v}_s}{\bar{v}_s-\underline{v}_s} \cdot \frac{\bar{v}_b-x}{\bar{v}_b-\underline{v}_b} \cdot \left(\frac{\bar{v}_b+x}{2} - \frac{x+\underline{v}_s}{2}\right)$, 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 $\Gamma$.[48] 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 we have that 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 $\Gamma$.

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:

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

where the second inequality follows from $m^2 \in TLB$ (and therefore $\underline{m}_s^3 < \underline{m}_b^2$). Plugging

[48] To see this, note that the second summand is increasing in $x$ in the range $x \in \left[\underline{m}_s^3 + \left|m_b^3\right|, \underline{m}_b^2 - \left|m_s^1\right|\right]$, and is thus (weakly) greater than $\frac{(\underline{m}_s^3+|m_b^3|)-\underline{m}_s^3}{\bar{v}_s-\underline{v}_s} \cdot \frac{\bar{v}_b-\underline{m}_b^2}{\bar{v}_b-\underline{v}_b} \cdot \left(\frac{\bar{v}_b+\underline{m}_b^2}{2} - \frac{(\underline{m}_s^3+|m_b^3|)+\underline{m}_s^3}{2}\right)$ which is equal to $\frac{|m_s^3|}{\bar{v}_s-\underline{v}_s} \cdot \frac{|m_b^3|}{\bar{v}_b-\underline{v}_b} \cdot \left(\frac{|m_b^2|+|m_b^3|}{2}\right)$ by equation 5 and since $\bar{v}_b = \bar{m}_b^3$. The expected surplus generated by the trade of buyers in $m_b^3$ and sellers in $m_s^3$ is $\frac{|m_b^3|}{\bar{v}_b-\underline{v}_b} \cdot \frac{|m_s^3|}{\bar{v}_s-\underline{v}_s} \cdot \left(\frac{\underline{m}_b^3+\bar{v}_b}{2} - \frac{\underline{m}_s^3+\underline{m}_s^4}{2}\right)$, 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 5.

in (4) and (5) we get that: [49]

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

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

**In sum,** we showed in the first part of the proof 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 part, 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. This posted price can be implemented by an intermediation mechanism (with $K=1$) by Proposition 6.

Thus, it must be that $K=1$ in the optimal intermediation mechanism. By Proposition 3, the optimal intermediation mechanism with $K=1$ must implement some posted price (because otherwise $\psi(M,p)\neq 0$), and among all posted prices the one that maximizes the social surplus is:

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

which is the maximizer of $\int_x^{\bar{v}_b}\int_{\underline{v}_s}^x(v_b-v_s)\,dv_s dv_b$ among all $x\in V_b\cap V_s$.

---

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

# References


1. **Akbarpour, M. and Li, S**. (2018), "Credible Mechanisms," *Working paper.*

2. **Austen-Smith, D.** (1993), "Interested experts and policy advice: Multiple referrals under open rule," *Games and Economic Behavior*, 5(1), 3-43.

3. **Baliga, S., Corchon, L. and Sjostrom, T.** (1997), "The theory of implementation when the planner is a player," *Journal of Economic Theory*, 77(1), 15-33.

4. **Battaglini, M.** (2002), "Multiple referrals and multidimensional cheap talk", *Econometrica*, 70(4), 1379-1401.

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

6. ——— (2001), "Contracting with Imperfect Commitment and the Revelation Principle: The Single Agent Case," *Econometrica*, 69(4), 1077-98.

7. ——— (2007), "Contracting with imperfect commitment and noisy communication," *Journal of Economic Theory*, 136, 236–259.

8. **Blume A., Board, O.J., and Kawamura, K.** (2007), "Noisy talk," *Theoretical Economics*, 2 (4), 395-440.

9. **Chakravorty, B., Corchón, L. C. and Wilkie, S.** (2006), "Credible implementation," *Games and Economic Behavior*, 57, 18–36.

10. **Crawford, V. and Sobel, J.** (1982), "Strategic information transmission," *Econometrica*, 50(6), 1431-51.

11. **Dewatripont, M. and Maskin, E.** (1990), "Contract renegotiation in models of asymmetric information," *European Economic Review*, 34, 311-321.

12. **Doval, L. and Skreta, V.** (2018), "Mechanism design with limited commitment," *Working paper.*

13. **Fey, M. and Ramsay, C.** (2009): "Mechanism design goes to war: peaceful outcomes with interdependent and correlated types," *Review of Economic Design*, 13(3), 233– 250.

14. **Forges, F.** (1986), "An approach to communication equilibria," *Econometrica*, 54,1375-1385.

15. **Gerardi, D., McLean, R. and Postlewaite, A.** (2008), "Aggregation of expert opinions," *Games and Economic Behavior*, 65(2), 339-371.

16. **Gilligan, T. and Krehbiel, K.** (1989), "Asymmetric information and legislative rules with a heterogeneous committee," *American Journal of Political Science*, 33(2), 459-490

17. **Goltsman, M., Hörner, J., Pavlov, G. and Squintani, F.** (2009), "Mediation, arbitration and negotiation," *Journal of Economic Theory*, 144(4), 1397-1420.

18. **Hörner, J., Morelli, M. and Squintani, F.** (2015), "Mediation and Peace," *The Review of Economic Studies*, 82(4), 1483–1501.

19. **Ivanov, M.** (2010), "Communication via a strategic mediator," *Journal of Economic Theory*, 145 (2) 869–884.

20. **Krishna, V. and Morgan, J.** (2001), "Asymmetric information and legislative rules: Some amendments," *American Political Science Review*, 95(2), 435-452.

21. **Kydd, A.** (2003), "Which side are you on? Bias, credibility, and mediation," *American Journal of Political Science*, 47(4), 597-611.

22. ——— (2006), "When can mediators build trust?," *American Political Science Review*, 100(3), 449-462.

23. **Laffont, J.-J. and Tirole, J.** (1988), "The Dynamics of Incentive Contracts," *Econometrica*, 56(5), 1153-1175.

24. **McAdams, D. and Schwartz, M.** (2007), "Credible Sales Mechanisms and Intermediaries," *The American Economic Review*, 97(1), 260-276.

25. **McAfee, P.R., and Vincent, D.** (1997), "Sequentially Optimal Auctions," *Games and Economic Bahaviour*, 18, 246-276.

26. **Myerson, R.** (1986), "Multistage games with communication," *Econometrica*, 54,323–358.

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

28. **Neeman, Z. and Pavlov, G.** (2012), "Ex-post renegotiation-proof mechanism design", *Journal of Economic Theory*, 148, 473–501.

29. **Skreta, V.** (2006), "sequentially optimal mechanisms", *The Review of Economic Studies*, 73 (4), 1085-1111.

30. ——— (2015), "Optimal auction design under non-commitment," *Journal of Economic Theory*, 159, 854-890.

31. **Sobel, J.** (2013). Giving and receiving advice. In: Acemoglu, D., Arellano, M., Dekel, E. (Eds.), Advances in Economics and Econometrics. Cambridge University Press, pp. 305–341, Chap. 10.

32. **Tirole, J.** (1986), "Procurement and Renegotiation," *Journal of Political Economy*, 94(2), 235-259.

33. **Vartiainen, H.** (2013), "Auction design without commitment," *Journal of the European Economic Association*, 11, 316-342.

34. **Wolinsky, A.** (2002), "Eliciting information from multiple experts," *Games and Economic Behavior*, 41(1), 141-160.

35. **Zheng, C. Z.** (2002), "Optimal auction with resale," *Econometrica*, 70, 2197–2224.