Bilateral Bargaining with a Biased Intermediary*

Yusuke Yamaguchi†

October 25, 2024

Abstract

Bilateral bargaining is often mediated by an *intermediary*. In many cases, an intermediary shares some interest with one of the two parties in the negotiation and cannot commit to and enforce his decision. This paper studies the implication of having such a biased intermediary without commitment and enforcement power on bargaining outcome. I consider a bilateral trade model à la Myerson and Satterthwaite (1983) with binary valuations in which the intermediary offers a price to the traders. I focus on the set of communication equilibria of the game, which characterizes what the players can achieve if they are allowed any preplay and intraplay communication. I show that even a tiny bias is detrimental to ex-post efficiency and that a biased intermediary does not help achieve it. I also characterize the second-best equilibrium and show that the expected social surplus is weakly decreasing in the degree of the intermediary's bias.

Keywords: Bargaining, intermediary, bias, mediator, communication equilibrium.

JEL Classification: C78, D82.

---

*I acknowledge funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 714693).

†Toulouse School of Economics, Toulouse, France. yusuke.yamaguchi@tse-fr.eu.

1 Introduction

Bilateral bargaining is often mediated by an *intermediary*, who communicates with the two parties in the negotiation and offers them a resolution. Such mediation obfuscates the private information inherent in offers and counter-offers that would be directly revealed absent an intermediary. For this reason, mediation is believed to mitigate strategic incentives in communication and help achieve efficient outcomes.

In many cases, however, an intermediary is not benevolent and shares some interest with one of the two parties in the negotiation. He opportunistically behaves and tries to achieve the outcomes he desires, which would harm efficiency if his objective is not aligned with it. He is not omnipotent either in the sense that he cannot commit to and enforce his decision. The lack of commitment and enforcement power hinders his ability to mediate the bargaining. As an example, consider a real estate agent. He facilitates transactions by communicating with a seller and a buyer and offering them a price. Since he gets a portion of a sale price as a commission, he has the incentive to realize a higher sale price, being *biased* toward the seller.[1] Normally, he does not in advance commit to the price he offers, and trade occurs only when the seller and the buyer accept his offer.

The purpose of this paper is to study the implication of having a *biased intermediary without commitment and enforcement power* on bargaining outcome. To draw a connection to the mechanism design literature, I study a game played by a seller, a buyer, and an intermediary in a bilateral trade setting reminiscent of Myerson and Satterthwaite (1983) with binary valuations. The seller owns a good and wants to trade it with the buyer. Each trader's valuation of the good is his private information and is independently and binary distributed. There is always gain from trade except for the pair of high-type seller and low-type buyer. If there is trade, the seller's payoff is the price minus his valuation and the commission to the intermediary; the buyer's payoff is his valuation minus the price and the commission to the intermediary; the intermediary's payoff is the sum of the commissions paid by the traders. If there is no trade, all players get zero.

[1] In the U.S., it has been standard that the seller pays the commissions to his agent *and* the buyer's agent; he pays 5% of a sale price, which is equally split between the seller's agent and the buyer's agent. The rule has recently changed so that a seller can choose not to pay for a buyer's agent and that a buyer needs to negotiate with his agent about the commission. See `https://www.nytimes.com/2024/08/16/realestate/home-buyers-sellers-rules.html` for details.

The *underlying game* proceeds as follows: first, the seller and the buyer observe the realizations of their types (valuations); second, the intermediary offers a price to the traders; and third, the traders simultaneously respond to the offer. If both traders accept it, trade occurs at that price; if at least one of them rejects it, no trade occurs. I am interested in what the three players can achieve in equilibrium of the underlying game *if they are allowed any preplay and intraplay communication*.

To obtain the bound on the achievable outcomes, I augment the game by adding a *fictitious* mediator, who collects type reports from the traders and recommends an action to each player. The *mediated bargaining game* proceeds as follows: first, the mediator publicly commits to a *mediation plan*, a function from type reports to recommendations; second, the seller and the buyer observe the realizations of their types and privately report them to the mediator; third, the mediator privately recommends a price to the intermediary; fourth, the intermediary offers a price to the traders; fifth, the mediator privately recommends responses to the offer to the traders; and last, the traders simultaneously respond to the offer. If both traders accept it, trade occurs at that price; if at least one of them rejects it, no trade occurs.

The set of incentive-compatible mediation plans, known as the set of *communication equilibrium (CE)* due to Forges (1986), characterizes the set of achievable outcomes in the underlying game with some preplay and intraplay communication. In my model, however, some CE are not robust to the players' trembles. Thus, I introduce a refinement *acceptable CE* and use it as the equilibrium concept. Roughly speaking, acceptable CE requires that the mediator always recommends weakly dominant responses—accept an offer if and only if it guarantees a nonnegative payoff and reject it otherwise.

The first main result is that even a tiny bias is detrimental to ex-post efficiency and that a biased intermediary does not help achieve it. Specifically, I derive a necessary and sufficient condition under which ex-post efficiency can be achieved in acceptable CE. This condition is (i) more stringent than that in the case of *unbiased* intermediary (when his payoff is independent of trading price) and (ii) the same as that in the case of no intermediary (a mediated seller-offer bargaining game). [2] To prove this result, I first obtain a necessary condition for acceptable CE,

[2]See Section D.1.1 for the unbiased case and Section D.2 for the no-intermediary case.

which requires that in any acceptable CE the mediator always recommends the intermediary to offer either of the highest acceptable prices (one for each type) of the buyer. Intuitively, given the above refinement of CE, any price that guarantees the buyer a nonnegative payoff will be accepted. Then, the intermediary has the incentive to offer one of the highest acceptable prices. Combined with the definition of ex-post efficiency, this necessary condition pins down a candidate mediation plan for an ex-post efficient acceptable CE. Verifying the players' IC constraints, I show that ex-post efficiency can be achieved if and only if the prior probability of the high-type buyer is small enough. This is because ex-post efficiency requires the intermediary to offer a low price for some reports. To deter the intermediary from deviating to a high price, which will be accepted only by the pair of high-type traders, and the low-type seller from misreporting his type, they must believe the buyer is likely enough to be of low-type. However, if the intermediary is unbiased, he can be incentivized to offer prices below the highest acceptable prices, which enlarges the possibility for ex-post efficiency.

As the second main result, I characterize the second-best outcome and show that a smaller bias is socially desirable as the expected social surplus is weakly decreasing in the degree of bias. The necessary condition for acceptable CE discussed above implies that the players' IC constraints can be written as linear inequalities in the probabilities the mediator recommends a certain price for each report. As such, the SB can be obtained by solving the corresponding linear program. The solution to this program implies that there is a threshold for the ratio of the players' net payment/revenue at which the SB outcome changes its regime. A simple calculation shows that the expected social surplus decreases if that ratio exceeds the threshold.

1.1 Related literature

Three features of the intermediary—bias, no commitment, and no enforcement— distinguish the present paper from other work in the literature. Intermediary, broadly defined, has been studied in the literature on mechanism design.[3] Classic papers such as Myerson and Satterthwaite (1983) and Matsuo (1989) can be considered studying the bargaining with an unbiased intermediary with commitment and enforcement power;[4] the principal maximizes the expected surplus; she can

[3]An intermediary in this literature is typically called a *principal*.

[4]The relationship with Matsuo (1989) is discussed in Appendix D.1.2.

commit to a mechanism; and the agents' acceptance is not necessary once they participate in a mechanism.[5] Some of the above three features have been separately studied in the literature. For example, Myerson (1982) studies a mechanism design problem when both hidden information and hidden action exist, which can be viewed as the principal having no enforcement power. Regarding the commitment assumption, there is also a growing literature on mechanism design with limited commitment (Bester and Strausz (2001, 2000, 2007); Doval and Skreta (2022); Lomys and Yamashita (2021), among others). Relatedly, Eilat and Pauzner (2021) study bilateral trade with a benevolent intermediary without commitment power.[6] Thus, the novelty of this paper is that the above three features are simultaneously present. Combined with no commitment, the intermediary's bias gives him the incentive to offer as high a price as possible, which is the driving force of the results.

Biased intermediary has been studied in the context of international relations (Kydd (2003, 2006)). Kydd (2003) shows that an intermediary is effective in reducing the probability of conflict only if he is biased. Roughly speaking, this is because in his model the intermediary needs to credibly communicate his private information to avoid conflict. He shows that such credible communication is possible only for a biased intermediary. In contrast to his result, a biased intermediary does not help achieve ex-post efficiency in the present paper. There are many differences in the models that could have led to opposite conclusions, but I believe the crucial difference is that the traders do not need to know the opponent's type to achieve ex-post efficiency. It is also worth mentioning that Kydd (2003) considers a fixed communication protocol whereas the present paper allows for any preplay and intraplay communication between the players.[7]

The present paper is also distinct from the previous work in the literature in that the intermediary is an active player of the game, who has a preference and actions to choose from. This clarifies the effect of the intermediary's bias on his facilitative role (facilitate agreement by obfuscating private information inherent in offers and counter-offers).[8] See Čopič and Ponsatí

[5] In this sense, "no enforcement" is equivalent to imposing ex-post IR constraint instead of ex-ante ones.

[6] They directly study the game between the traders and the intermediary, in which the intermediary offers a mechanism. Note that the intermediary in my model just offers a price and is not allowed to offer a mechanism. In this regard, Eilat and Pauzner (2021) allow more flexibility in the intermediary's actions.

[7] The relationship with Kydd (2003) is also discussed in Section 5.2.

[8] Gottardi and Mezzetti (2024) study not only on this facilitative role but also on the evaluative role of the intermediary; the intermediary can provide some guidance on the appropriate resolution based on his expertise.

(2008); Fanning (2021, 2023); Jarque, Ponsatí, and Sákovics (2003) for the facilitative role and also Ganguly and Ray (2012); Goltsman, Hörner, Pavlov, and Squintani (2009); Hörner, Morelli, and Squintani (2015) for the study on mediation.

Methodologically, I lean on the work of Forges (1986), Myerson (1986b), and Sugaya and Wolitzky (2021), who study multistage games with communication and prove the *communication revelation principle* for various equilibrium concepts. The communication revelation principle establishes that any equilibrium outcome of the underlying game with some preplay and intraplay communication is achievable as a CE of the mediated bargaining game. In particular, it is without loss of generality to focus on the direct communication between the players and the mediator in the mediated bargaining game, as described above.

The remainder of the paper is organized as follows. Section 2 describes the model and define equilibrium concepts. In Section 3, I present some preliminary results. Section 4 studies the bound on the expected social surplus achievable in equilibrium. In particular, I derive a necessary and sufficient condition for achieving ex-post efficiency and also characterize the second-best outcome. In Section 5, I discuss the implications of other payoff specifications and their relation to some papers in the literature. Section 6 concludes the paper. All proofs are relegated to Appendix A.

2 Model

There are three players: a seller, a buyer, and an intermediary. The seller owns a good and wants to trade it with the buyer. The traders' valuations of the good are binary and independently distributed, and their realizations are private information of the respective trader. Specifically, the seller's valuation of the good is high (s_H) with probability $\pi_S \in (0,1)$ and low (s_L) with probability $1-\pi_S$. The buyer's valuation of the good is high (b_H) with probability $\pi_B \in (0,1)$ and low (b_L) with probability $1-\pi_B$. Let $\Theta_S = \{s_L, s_H\}$ and $\Theta_B = \{b_L, b_H\}$ be the set of possible valuations of the seller and the buyer, respectively. With slight abuse of notation, I denote the seller with valuation $s \in \Theta_S$ (resp. the buyer with valuation $b \in \Theta_B$) by s (resp. b). I assume that there is always gain from trade except for the pair of traders (s_H, b_L); that is, $0 < s_L < b_L < s_H < b_H$. If (s, b) trade at price p, the seller gets $(1-\delta_S)p - s$, the buyer gets

$b - (1 + \delta_B)p$, and the intermediary gets $(\delta_S + \delta_B)p$, where $\delta_S, \delta_B \in (0, 1)$; if there is no trade, all the players get 0. This payoff specification is interpreted as each trader paying the commissions $\delta_S p$ and $\delta_B p$ to the intermediary.[9] Since the intermediary's payoff is increasing in price p, he shares some interest with the seller. In this sense, he is considered *biased* toward the seller.[10]

The *underlying game* I consider here proceeds as follows: (i) the seller and the buyer observe the realizations of their types (valuations); (ii) the intermediary offers a price to the traders; (iii) the traders simultaneously respond to the offer. If both traders accept it, trade occurs at that price; if at least one of them rejects it, no trade occurs. I am interested in what the three players can achieve in the underlying game *if they can freely communicate with each other at any point in the game* (sending messages, tossing a coin, using a correlation device, etc.) However, there are infinitely many possible forms of preplay and intraplay communication between the players, hence infinitely many games I need to consider. To circumvent this issue, I consider a mediated version of the underlying game, which is the subject of Section 2.1.

2.1 The mediated bargaining game

I augment the model by adding a *fictitious* mediator, who collects type reports from the traders between (i) and (ii) and later privately recommends action to each player. She carries out such mediation using a decision rule called a mediation plan. A *pure mediation plan* is a pair of functions (q, r), where $q\colon \Theta_S \times \Theta_B \to \mathbb{R}_+$ is a *price-recommendation* and $r\colon \Theta_S \times \Theta_B \times \mathbb{R}_+^2 \to \{Y, N\}^2$ is a *response-recommendation*.[11] For example, suppose that the mediator uses a pure mediation plan (q, r) and that the traders have reported $(\tilde{s}, \tilde{b}) \in \Theta_S \times \Theta_B$. Then, the mediator recommends the intermediary to offer $p = q(\tilde{s}, \tilde{b}) \in \mathbb{R}_+$. After observing the intermediary's offer $\tilde{p}$, she recommends $r(\tilde{p} \mid \tilde{s}, \tilde{b}, p) = \big(r_S(\tilde{p} \mid \tilde{s}, \tilde{b}, p), r_B(\tilde{p} \mid \tilde{s}, \tilde{b}, p)\big) \in \{Y, N\}^2$ to the traders, where r_S is the Seller's response (acceptance "Y" or rejection "N") to the offer and r_B is the Buyer's one.[12] Note that even though $r(\tilde{p} \mid \tilde{s}, \tilde{b}, p)$ specifies both traders' responses, the mediator privately

[9] I discuss alternative specifications in Section 5.

[10] If the intermediary's payoff is independent of price (for example, if he gets a fixed commission from trade), he is considered *unbiased*. I explore this case in Appendix D.1.1

[11] Strictly speaking, the communication revelation is established only for finite games. Hence, the intermediary's action space (the set of possible prices) must be finite to ensure the validity of the entire analysis. If I assume a finite set of possible prices, I can obtain qualitatively the same results. For the brevity of the exposition, I use the current continuous formulation.

[12] I denote $r(\tilde{s}, \tilde{b}, p, \tilde{p})$ by $r(\tilde{p} \mid \tilde{s}, \tilde{b}, p)$.

recommends each trader's response; trader $i \in \{S, B\}$ only observes $r_i(\tilde{p} \mid \tilde{s}, \tilde{b}, p) \in \{Y, N\}$. Let Q and R be the set of all price-recommendations and response-recommendations, respectively. For greatest generality, I allow the mediator to use a *mixed mediation plan* (or simply a *mediation plan*), a probability distribution $\mu \in \Delta(Q \times R)$, where $\Delta(X)$ denote the set of all probability distributions over the set X. Let $\mu(q, r)$ denote the probability that the mediator uses pure mediation rule (q, r).

I consider the *mediated bargaining game* that proceeds as follows:

1. The mediator publicly commits to a mediation plan $\mu \in \Delta(Q \times R)$ she uses;
2. The seller and the buyer privately observe the realizations of their types;
3. The traders privately report their types $\tilde{s} \in \Theta_S$ and $\tilde{b} \in \Theta_B$ to the mediator;
4. The mediator privately picks up a pure mediation plan $(q, r) \in Q \times R$ with probability $\mu(q, r)$ and then privately recommends a price $p = q(\tilde{s}, \tilde{b}) \in \mathbb{R}_+$ to the intermediary;
5. The intermediary offers a price $\tilde{p} \in \mathbb{R}_+$;
6. The mediator privately recommends traders' responses $r(\tilde{p} \mid \tilde{s}, \tilde{b}, p) \in \{Y, N\}^2$;
7. The traders simultaneously respond to the offer $\tilde{p}$ by either acceptance (Y) or rejection (N). If both traders accept it, trade occurs at the price $\tilde{p}$, and the payoffs are realized; if at least one of them rejects it, no trade occurs, and all players get a payoff of 0.

Remark 1. As I discussed in Introduction, on top of bias, the intermediary features no commitment and no enforcement. Indeed, he does not commit to the price he offers; he updates his belief about the traders' types upon receiving a recommendation and offers an optimal price for him given his belief. Moreover, the traders' acceptances are necessary for trade to occur at the offered price. In contrast to the intermediary, the mediator *can* commit to the mediation plan she uses.

2.2 Communication equilibrium and its refinement

Each player can manipulate a mediation plan by disobeying recommendations, and the traders can also manipulate it by misreporting their types. In an equilibrium mediation plan, which is defined shortly, none of such manipulations are profitable.

I use *communication equilibrium (CE)* of Forges (1986) as an equilibrium concept. In CE, no player could ex-ante expect to gain by manipulation (disobeying and misreporting). In other words, no player could expect to gain by manipulation after any event that is observable to him and has a *positive probability* of occurring in equilibrium.[13]

It is worth noting that there is no loss of generality coming from the form of communication in the mediated bargaining game (namely, Time 3, 4, and 6). In principle, I can consider more complex forms of communication between the players and the mediator. For example, the traders' message space can be larger than their type spaces, and the mediator can communicate more than recommended actions to the players. However, thanks to the communication revelation principle shown by Forges (1986), Myerson (1986b) and, Sugaya and Wolitzky (2021), any equilibrium outcome of the underlying game with some preplay and intraplay communication is achievable as a CE of the mediated bargaining game. As such, it suffices to characterize the set of CE outcomes.

To formally define CE, let $V(q,r)$ be the intermediary's ex-ante expected payoff when the mediator uses a pure mediation plan (q,r) and all players are honest and obedient to her:

$$V(q,r) = \sum_{(s,b)\in\Theta_S\times\Theta_B} \Pr(s,b)(\delta_S+\delta_B)q(s,b)\mathbf{1}_{\{r(q(s,b)\,|\,s,b,q(s,b))=(Y,Y)\}},$$

where $\Pr(s,b)$ is the prior probability that the traders' types are (s,b), and $\mathbf{1}_{\{\cdot\}}$ is the indicator function. Similarly, for trader $i\in\{S,B\}$, let $U_i(q,r)$ be i's ex-ante expected payoff when the mediator uses a pure mediation plan (q,r) and all players are honest and obedient to her:

$$U_S(q,r) = \sum_{(s,b)\in\Theta_S\times\Theta_B} \Pr(s,b)\big[(1-\delta_S)q(s,b)-s\big]\mathbf{1}_{\{r(q(s,b)\,|\,s,b,q(s,b))=(Y,Y)\}},$$
$$U_B(q,r) = \sum_{(s,b)\in\Theta_S\times\Theta_B} \Pr(s,b)\big[b-(1+\delta_B)q(s,b)\big]\mathbf{1}_{\{r(q(s,b)\,|\,s,b,q(s,b))=(Y,Y)\}}.$$

Manipulation by the intermediary can be represented by a function $\alpha:\mathbb{R}_+\to\mathbb{R}_+$. Let Σ_I be the set of all such functions. Manipulation by trader $i\in\{S,B\}$ can be represented by a pair of functions

[13]Myerson (1986b) considers a refinement of CE, *sequential communication equilibrium (SCE)*. SCE requires there being no profitable manipulation even after events that have *zero-probability* of occurring in equilibrium. As I show in Appendix B.1, every CE is an SCE in the mediated bargaining game. Thus, I work with CE in the remainder of the paper.

(β_i, γ_i), where $\beta_i \colon \Theta_i \to \Theta_i$ is a manipulation by report, and $\gamma_i \colon \Theta_i \times \mathbb{R}_+ \times \{Y, N\} \to \{Y, N\}$ is a manipulation by response. For $i \in \{S, B\}$, let Σ_i be the set of all such pairs of functions.

For any $\alpha \in \Sigma_I$, let $\alpha \circ q$ be such that $(\alpha \circ q)(s, b) = \alpha\big(q(s, b)\big)$ for all $(s, b) \in \Theta_S \times \Theta_B$; that is, $\alpha \circ q$ gives the intermediary's offer as a function of reports (s, b) when the mediator uses q and the intermediary manipulates it by α. For any $(\beta_i, \gamma_i) \in \Sigma_i$, let $q \circ \beta_i$ and $r \circ (\beta_i, \gamma_i)$ be such that, for all $(\theta_i, \theta_j) \in \Theta_i \times \Theta_j$ and $p \in \mathbb{R}_+$,[14]

$$
\begin{aligned}
&(q \circ \beta_i)(\theta_i, \theta_j) = q\big(\beta_i(\theta_i), \theta_j\big) \\
&\big(r \circ (\beta_i, \gamma_i)\big)(p \mid \theta_i, \theta_j) = \Big(\gamma_i\big(\theta_i, p, r_i(p \mid \beta_i(\theta_i), \theta_j, q(\beta_i(\theta_i), \theta_j))\big), r_j\big(p \mid \beta_i(\theta_i), \theta_j, q(\beta_i(\theta_i), \theta_j)\big)\Big)
\end{aligned}
$$

where $j \neq i$; that is, $q \circ \beta_i$ gives the intermediary's offer when the mediator uses q, trader i manipulates it by β_i, and the other trader j and the intermediary are honest and obedient to the mediator. Similarly, $r \circ (\beta_i, \gamma_i)$ gives the traders' responses when the mediator uses (q, r), trader i manipulates it by (β_i, γ_i), and the other trader j and the intermediary are honest and obedient to the mediator. To simplify the notation, let $(q, r) \circ (\beta_i, \gamma_i)$ denote $\big(q \circ \beta_i, r \circ (\beta_i, \gamma_i)\big)$. For any $\alpha \in \Sigma_I$ and $(\beta_i, \gamma_i) \in \Sigma_i$, the ex-ante expected payoffs of the intermediary and trader i when they use these manipulations, $V(\alpha \circ q, r)$ and $U_i\big((q, r) \circ (\beta_i, \gamma_i)\big)$, respectively, are naturally defined. With these notations, I can now define CE.

Definition 1. A mediation plan $\mu \in \Delta(Q \times R)$ is a *communication equilibrium (CE)* if, for all $\alpha \in \Sigma_I$,

$$
\sum_{(q,r) \in Q \times R} \mu(q, r) V(q, r) \geq \sum_{(q,r) \in Q \times R} \mu(q, r) V(\alpha \circ q, r), \tag{2.1}
$$

and for all all $i \in \{S, B\}$ and $(\beta_i, \gamma_i) \in \Sigma_i$,

$$
\sum_{(q,r) \in Q \times R} \mu(q, r) U_i(q, r) \geq \sum_{(q,r) \in Q \times R} \mu(q, r) U_i\big((q, r) \circ (\beta_i, \gamma_i)\big). \tag{2.2}
$$

Remark 2. In Appendix B.2, I show that the *ex-ante* IC constraints (2.1) and (2.2) are satisfied if

[14] Abusing the notation, I write $r \circ (\beta_i, \gamma_i)$ instead of $(q \circ r) \circ (\beta_i, \gamma_i)$ as I use the latter to denote the pair $\big(q \circ \beta_i, r \circ (\beta_i, \gamma_i)\big)$.

and only if the *interim* IC constraints are satisfied; that is, (i) no trader can gain by manipulation *after* he learns his type (at Time 3) and (ii) the intermediary cannot gain by manipulation *after* he receives a recommendation (at Time 5). Accordingly, I consider the players' interim IC constraints in the remainder of the paper as they are easier to handle than the ex-ante ones.

2.2.1 Acceptable CE

Without any restriction on admissible response-recommendations, it is easy to construct a CE that implements an ex-post efficient outcome since the mediator can punish the intermediary's deviations. Indeed, consider a mediation rule that, whenever the buyer has a higher valuation, recommends (i) the intermediary to offer a mutually acceptable price (a price that guarantees nonnegative payoffs to both traders) and (ii) the traders to accept an offer only when the offer is mutually acceptable and the intermediary has not deviated. In Appendix C, I provide a necessary and sufficient condition under which such a mediation plan constitutes a CE. However, such CE are not robust to the players' trembles. For example, if the intermediary by mistake offers a mutually acceptable price different from the recommended one, the traders can be better off by accepting this offer though they are recommended to punish the intermediary by rejecting it.

Such non-robust CE can be eliminated if the traders' are always recommended to accept mutually acceptable prices, which is reminiscent of a weakly dominant strategy. To formally define such response-recommendation, observe that the seller with valuation s gets a nonnegative payoff by accepting p if and only if $p \geq \underline{p_s} \equiv \frac{s}{1-\delta_S}$. Similarly, the buyer with valuation b gets a nonnegative payoff by accepting p if and only if $p \leq \overline{p_b} \equiv \frac{b}{1+\delta_B}$. I assume $\underline{p_{s_L}} \leq \overline{p_{b_L}}$ and $\underline{p_{s_H}} \leq \overline{p_{b_H}}$ so that there is a room for achieving ex-post efficiency.[15] Define $r^A \in R$ as follows. For all $\tilde{p} \in \mathbb{R}_+$ and all $(s,b) \in \Theta_S \times \Theta_B \setminus \{(s_H, b_L)\}$,

$$r^A(p \mid s, b, \tilde{p}) = \begin{cases} (N,Y) & \text{if } p \in \left[0, \underline{p_s}\right) \\ (Y,Y) & \text{if } p \in \left[\underline{p_s}, \overline{p_b}\right] \\ (Y,N) & \text{if } p \in (\overline{p_b}, +\infty) \end{cases}$$

[15]These conditions are reduced to $\frac{1+\delta_B}{1-\delta_S} \leq \min\left\{\frac{b_L}{s_L}, \frac{b_H}{s_H}\right\}$. If this does not hold, at least one of (s_L, b_L), (s_L, b_H), and (s_H, b_H) cannot agree to trade though there is gain from trade.

and for all $\tilde{p} \in \mathbb{R}_+$,

$$r^A(p \mid s_H, b_L, \tilde{p}) = \begin{cases} (N,Y) & \text{if } p \in \left[0, \overline{p_{b_L}}\right] \\ (N,N) & \text{if } p \in \left(\overline{p_{b_L}}, \underline{p_{s_H}}\right) \\ (Y,N) & \text{if } p \in \left[\underline{p_{s_H}}, +\infty\right). \end{cases}$$

That is, r^A neglects the intermediary's deviations and recommends each trader to accept an offer if and only if it guarantees him a nonnegative payoff.[16] Thus, each trader finds it optimal to follow r^A if he has honestly reported his type. I restrict attention to CE that only uses r^A as a response-recommendation.

Definition 2. A CE $\mu \in \Delta(Q \times R)$ is *acceptable* if $\mu\left(Q \times \{r^A\}\right) = 1$.[17]

3 Preliminary Results

In this section, I first provide a necessary condition for acceptable CE and use it to simplify the players' IC constraints.

It is straightforward to show that in any acceptable CE, the mediator recommends the intermediary to offer either $\overline{p_{b_L}}$ and $\overline{p_{b_H}}$ whenever there is gain from trade. To see this, suppose that the intermediary is recommended $p \notin \left\{\overline{p_{b_L}}, \overline{p_{b_H}}\right\}$. If $p < \overline{p_{b_L}}$ and he expects it to be accepted, then he can gain by deviating to $\overline{p_{b_L}}$ since it will also be accepted given the definition of r^A. Similarly, if $\overline{p_{b_L}} < p < \overline{p_{b_H}}$ and he expects it to be accepted, then he can gain by deviating to $\overline{p_{b_H}}$. Formalizing this argument, I can show the following lemma.

Lemma 1. *A mediation plan $\mu \in \Delta\left(Q \times \{r^A\}\right)$ is an acceptable CE only if it recommends the intermediary to offer either $\overline{p_{b_L}}$ or $\overline{p_{b_H}}$ whenever there is gain from trade. That is, for all $q \in \operatorname{supp}(\mu)$,*

$$q(s,b) \in \left\{\overline{p_{b_L}}, \overline{p_{b_H}}\right\} \textit{ for all } (s,b) \in \Theta_S \times \Theta_B \setminus \{(s_H, b_L)\} \tag{3.1}$$

[16]Since r^A does not depend on the recommended price $\tilde{p}$, abusing the notation, I hereafter omit $\tilde{p}$ from the arguments of r^A.

[17]This refinement has a flavor of *acceptable correlated equilibrium* of Myerson (1986a), hence the name.

Proof. See Appendix A.1. □

Lemma 1 implies that for any candidate for acceptable CE, the players' IC constraints can be written as functions of the total probability that the mediator recommends offering price $\overline{p_{b_H}}$ for each pair of reports. For any mediation plan $\mu \in \Delta\left(Q \times \{r^A\}\right)$ that satisfies (3.1), let

$$x_{HH}(\mu) = \sum_{q \in \left\{q' \in Q:\ q'(s_H, b_H) = \overline{p_{b_H}}\right\}} \mu(q),$$
$$x_{LH}(\mu) = \sum_{q \in \left\{q' \in Q:\ q'(s_L, b_H) = \overline{p_{b_H}}\right\}} \mu(q),$$
$$x_{LL}(\mu) = \sum_{q \in \left\{q' \in Q:\ q'(s_L, b_L) = \overline{p_{b_H}}\right\}} \mu(q),$$

where $\mu(q)$ denotes the probability that the mediator picks up pure mediation plan (q, r^A) when she uses μ. For example, $x_{HH}(\mu)$ is the total probability that μ recommends the intermediary to offer $\overline{p_{b_H}}$ when the reports are (s_H, b_H). These probabilities completely describe the price-recommendations for all reports $(s, b) \in \Theta_S \times \Theta_B \setminus \{(s_H, b_L)\}$. Note that by definition of r^A, traders (s_H, b_L) never trade if they are obedient to the mediator. Since the objective is to characterize the bound on the implementable outcomes, it suffices to recommend offering a high enough price for reports (s_H, b_L). This deters traders from double deviations (misreport and then disobey) and makes the players' IC constraints easier to satisfy. For this reason, I restrict attention to price-recommendations in $Q_x \subset Q$ defined below:

$$Q_x = \left\{q \in Q\colon q \text{ satifies (3.1) and } q(s_H, b_L) \in \left(\overline{p_{b_H}}, +\infty\right)\right\}.$$

Note that for any mediation plan $\mu \in \Delta\left(Q \times \{r^A\}\right)$ that satisfies (3.1), I can construct a mediation plan $\mu_x \in \Delta\left(Q_x \times \{r^A\}\right)$ that has weakly less stringent IC constraints than μ and induces the same distribution over the bargaining outcomes (whether trade occurs or not, and if it does, at which price) as μ if the players are honest and obedient to the mediator. To explain how to construct it, for any q in the support of such μ, let $x_q \in Q_x$ be an arbitrary price-recommendation such that $x_q(s, b) = q(s, b)$ for all $(s, b) \in \Theta_S \times \Theta_B \setminus \{(s_H, b_L)\}$. That is, x_q recommends the same prices as q for all reports $(s, b) \in \Theta_S \times \Theta_B \setminus \{(s_H, b_L)\}$ but recommends a high price in $\left(\overline{p_{b_H}}, +\infty\right)$

for reports (s_H, b_L). The mediation plan μ_x can be constructed by letting $\mu_x(x_q) = \mu(q)$ for all $q \in \text{supp}(\mu)$. Lemma 2 allows me to focus on such μ_x.

Lemma 2. *For any mediation plan $\mu \in \Delta\left(Q \times \{r^A\}\right)$, if μ is an acceptable CE, so is μ_x.*

Proof. See Appendix A.2. □

Remark 3. The proof of Lemma 2 implies that for any mediation plan in $\Delta\left(Q_x \times \{r^A\}\right)$ characterized by total probabilities x_{HH}, x_{LH}, and x_{LL}, the players' IC constraints are given by the following linear inequalities in these probabilities:

$$x_{HH} \geq x_{LH} \tag{IC-s_H}$$

$$x_{LL} \geq x_{LH} \tag{IC-b_H}$$

$$(1-\pi_B)\left[(1-\delta_S)\overline{p_{b_L}} - s_L\right](1-x_{LL}) \geq \pi_B(1-\delta_S)\left(\overline{p_{b_H}} - \overline{p_{b_L}}\right)(x_{HH}-x_{LH}) \tag{IC-s_L}$$

$$\pi_S\pi_B b_H x_{HH} + (1-\pi_S)\pi_B(b_H-b_L)x_{LH} \geq (1-\pi_S)(1-\pi_B)b_L x_{LL} \tag{IC-Int1}$$

$$(1-\pi_S)(1-\pi_B)b_L(1-x_{LL}) \geq \pi_S\pi_B b_H(1-x_{HH}) + (1-\pi_S)\pi_B(b_H-b_L)(1-x_{LH}) \tag{IC-Int2}$$

As such, if I am interested in a maximum (or a maximizer) of some objective function that is also linear in x_{HH}, x_{LH}, and x_{LL}, it can be found by solving a corresponding linear program.

4 Bound on the Expected Social Surplus

In this section, I take the expected social surplus (the expected value of the sum of the players' payoffs) as an objective function to be maximized.[18] Since the payments cancel out, the social surplus coincides with the trade surplus. Hence, for any $\mu \in \Delta\left(Q_x \times \{r^A\}\right)$ characterized by x_{HH}, x_{LH}, and x_{LL}, the expected social surplus is given by

$$\sum_{(s,b)\in\Theta_S\times\Theta_B} \Pr(s,b) \sum_{q\in Q_x} \mu(q)(b-s)\mathbf{1}_{\{r^A(q(s,b)\,|\,s,b)=(Y,Y)\}}$$
$$= \pi_S\pi_B(b_H-s_H)x_{HH} + (1-\pi_S)\pi_B(b_H-s_L) + (1-\pi_S)(1-\pi_B)(b_L-s_L)(1-x_{LL}). \tag{4.1}$$

[18]In Appendix E, I briefly discuss how to implement the ex-post efficient and the second-best outcomes in the game without the mediator.

Thus, the problem is reduced to maximizing (4.1) by choosing x_{HH}, x_{LH}, and x_{LL} in $[0, 1]$ subject to (IC-s_H), (IC-b_H), (IC-s_L), (IC-Int1), and (IC-Int2). I call this linear program (P).

4.1 Ex-post efficiency

I first explore when ex-post efficiency can be achieved in equilibrium. A mediation plan is *ex-post efficient* if, assuming the players are honest and obedient to the mediator, trade occurs if and only if the buyer has a higher valuation. If an ex-post efficient mediation plan is an acceptable CE, it is called *ex-post efficient* acceptable CE. Given the definition of r^A, a mediation plan $\mu \in \Delta\left(Q_x \times \{r^A\}\right)$ is ex-post efficient if and only if $x_{HH} = 1$ and $x_{LL} = 0$ so that it always recommends $\overline{p_{b_H}}$ to (s_H, b_H) and $\overline{p_{b_L}}$ to (s_L, b_L).[19] Moreover, $x_{LH} = 0$ is also necessary for such a mediation plan to be an acceptable CE. Indeed, if $x_{LH} > 0$ and hence price $\overline{p_{b_L}}$ is offered to (s_L, b_H) with probability $1 - x_{LH} < 1$, the high-type buyer b_H can gain by misreporting his type. In that case, he can secure positive payoff $b_H - (1 + \delta_B)\overline{p_{b_L}}$ with probability $1 - \pi_S$ while he gets the same payoff only with smaller probability $(1 - \pi_S)(1 - x_{LH})$ if he is honest and obedient to the mediator. I summarize these observations in the next lemma.

Lemma 3. *A mediation plan $\mu \in \Delta\left(Q_x \times \{r^A\}\right)$ is ex-post efficient if and only if $x_{HH} = 1$ and $x_{LL} = 0$. Moreover, such a mediation plan is an acceptable CE only if $x_{LH} = 0$.*

By Lemma 3, it suffices to check when $(x_{HH}, x_{LH}, x_{LL}) = (1, 0, 0)$ satisfy the players' IC constraints. Note that the high-type seller s_H and both types of buyers have no incentive for manipulation; if they honestly report their types, they find it optimal to follow r^A by definition; even if they misreport their types, they cannot trade at better terms. However, the low-type seller s_L may have the incentive to misreport his type to realize trade at a higher price $\overline{p_{b_H}}$, which is possible if the buyer is of high-type. To deter him from such manipulation, he must believe that the buyer is likely enough to be of low-type b_L. That is, the prior probability π_B of the high-type buyer must be small enough. The intermediary must not have the incentive to disobey recommendations either. If he is recommended $\overline{p_{b_H}}$, he learns that the traders are (s_H, b_H), and hence following the recommendation is optimal. However, if he is recommended $\overline{p_{b_L}}$, he may have the incentive to offer $\overline{p_{b_H}}$ to realize trade at a higher price, which occurs only when the

[19]When there is no confusion, I write x_{HH}, x_{LH}, and x_{LL} instead of $x_{HH}(\mu_x)$, $x_{LH}(\mu_x)$, and $x_{LL}(\mu_x)$.

buyer is of high-type. He does not have such an incentive if π_B is small enough. It turns out that the seller's threshold is more stringent, and I obtain the following proposition.

Proposition 1. *There exists an ex-post efficient acceptable CE if and only if the prior probability π_B of the high-type buyer satisfies $\pi_B \leq \frac{(1-\delta_S)b_L-(1+\delta_B)s_L}{(1-\delta_S)b_H-(1+\delta_B)s_L}$.*

Proof. See Appendix A.3. □

Let $\overline{\pi_B}$ denote the threshold in Proposition 1. It is intuitive that $\overline{\pi_B}$ is decreasing in δ_S, δ_S, and s_L; the higher the commission rates or the seller's reservation value, the smaller the possibility of efficient trade. In this sense, the intermediary's bias, or desire to realize trade at a higher price, is detrimental to ex-post efficiency.

4.2 Second-best outcome

In this subsection, I characterize the *second-best (SB)* outcome—implementable outcome that maximizes the expected social surplus when ex-post efficiency cannot be achieved. Acceptable CE that implements the SB outcome can be obtained as the solution to the linear program (P) when $\pi_B > \overline{\pi_B}$. The next proposition gives the characterization.

Proposition 2. *There exists a threshold $\overline{h}$ on the ratio $\frac{1+\delta_B}{1-\delta_S}$ such that*

1. *if $\frac{1+\delta_B}{1-\delta_S} \in \left(1, \overline{h}\right]$, then in the SB outcome, traders (s_H, b_H) do not trade with positive probability while (s_L, b_H) and (s_L, b_L) trade at price $\overline{p_{b_L}}$ with probability 1;*
2. *otherwise, in the SB outcome, traders (s_H, b_H) and (s_L, b_H) trade at price $\overline{p_{b_H}}$ with probability 1 while (s_L, b_L) do not trade.*

Proof. See Appendix A.4. □

As shown in the proof of Proposition 1, if $\pi_B > \overline{\pi_B}$ then $(x_{HH}, x_{LH}, x_{LL}) = (1, 0, 0)$ do not satisfy the low-type seller's IC constraint (IC-s_L) since he can gain by misreporting his type. Given the expression of the expected social surplus (4.1), the SB is likely to be achieved by keeping $x_{LH} = x_{LL} = 0$ and decreasing x_{HH} until (IC-s_L) binds—giving up some probability of trade between (s_H, b_H)—to deter s_L from misreporting. Indeed, the SB is characterized by

$\left(x^*_{HH}, x^*_{LH}, x^*_{LL}\right) = (\overline{x_{HH}}, 0, 0)$ if $h \le \overline{h}$, where $\overline{x_{HH}}$ is the value of x_{HH} at which (IC-s_L) binds when $x_{LH} = x_{LL} = 0$. Decreasing x_{HH} from 1 to $\overline{x_{HH}}$ is innocuous to all other constraints except for the intermediary's constraint (IC-Int2), which ensures he follows the recommendation of offering $\overline{p_{b_L}}$. When $\overline{p_{b_L}}$ is recommended, the smaller x_{HH}, the more he believes the traders are (s_H, b_H), who do not accept price $\overline{p_{b_L}}$. Thus, he has the incentive to disobey and offer $\overline{p_{b_H}}$ if x_{HH} is too small. The lower bound implied by (IC-Int2) is compatible with $\overline{x_{HH}}$ if ratio $h = \frac{1+\delta_B}{1-\delta_S}$ is small enough. To see why this, consider the low-type seller's incentive. Recall that he gets a payoff of $(1-\delta_S)\overline{p_b} - s_L$ if he trades at price $\overline{p_b}$, where $\overline{p_b} = \frac{b}{1+\delta_B}$; he pockets the price net of the commission $\delta_S \overline{p_b}$ to the intermediary. That is why there is a bound on ratio h, not on the absolute levels of the commission rates δ_S and δ_B. As ratio h increases, misreporting becomes more attractive to him, hence a smaller $\overline{x_{HH}}$.[20] Thus, there is an upper bound on ratio h.

If ratio h exceeds the threshold $\overline{h}$, then (IC-s_L) and (IC-Int2) cannot be satisfied at the same time at $(x_{HH}, x_{LH}, x_{LL}) = \left(\overline{x_{HH}}, 0, 0\right)$ because the intermediary has the incentive to disobey and offer $\overline{p_{b_H}}$. In this case, the SB involves the mediator always recommending him to offer $\overline{p_{b_H}}$. That is, the SB is $\left(x^*_{HH}, x^*_{LH}, x^*_{LL}\right) = (1, 1, 1)$. The following argument is not rigorous but gives some intuition into this case. For fixed δ_S and δ_B, ratio h is likely to exceed the threshold $\overline{h}$ when the prior probability π_B of high-type buyer is large since the threshold is decreasing in π_B. Consider the case where π_B is large enough. When $\overline{p_{b_L}}$ is recommended, the intermediary has the incentive to follow the recommendation if the traders are likely to be (s_L, b_L), who reject $\overline{p_{b_H}}$. The smaller the value of x_{LL}, the more he believes so. Indeed, the left-side of (IC-Int2) is decreasing in x_{LL}. By (IC-b_H), he has the largest incentive to follow the recommendation $\overline{p_{b_L}}$ when $x_{LL} = x_{LH}$. However, substituting this into (IC-Int2), it can be satisfied only when $x_{HH} = x_{LH} = x_{LL} = 1$.[21] In other words, the intermediary can no longer be incentivized to offer $\overline{p_{b_L}}$. Thus, in the only acceptable CE hence the SB, the mediator always recommends $\overline{p_{b_H}}$.

From Proposition 2, a simple calculation shows that the expected social surplus is weakly decreasing in h. On $h \in \left(1, \overline{h}\right)$, it is decreasing as the trade probability $\overline{x_{HH}}$ is decreasing in h. If $h \le \overline{h}$ increases to $h' > \overline{h}$ and the SB outcome changes its regime, the expected surplus

[20] His payoff from trade is decreasing in ratio h for whichever price $\overline{p_{b_L}}$ or $\overline{p_{b_H}}$. However, the equilibrium payoff decreases faster than the best payoff from misreporting.

[21] Remember I assume π_B is large enough.

decreases. This implies that changes in the commission rates that increase ratio h can lead to the loss of social surplus and harm society.

Corollary 1. *The expected social surplus is weakly decreasing in h. Specifically,*

1. *On* $\left(1, \overline{h}\right)$*, the expected social surplus is decreasing in ratio h.*

2. *If* $h \leq \overline{h}$ *increases to* $h' > \overline{h}$*, the expected social surplus weakly decreases.*

Proof. See Appendix A.5. $\square$

5 Discussion

In this section, I briefly discuss alternative payoff specifications and their implications.[22]

5.1 Intermediary as a market mechanism in Loertscher and Marx (2022)

Loertscher and Marx (2022) study a model of bilateral bargaining and model the market as a mechanism that maximizes the expected weighted average of firms' (sellers and buyers) payoffs. Motivated by their work, I can think of the intermediary playing a role of the market mechanism, especially the role of offering prices to the traders. Consider the following payoff specification:

- If traders (s, b) trade at price p,
  - the seller gets $p - s$;
  - the buyer gets $b - p$;
  - the intermediary gets $\lambda(p - s) + (1 - \lambda)(b - p)$;
- If no trade occurs, all players get 0.

The weight $\lambda \in [0, 1]$ represents the degree of the intermediary's bias toward the seller. A biased intermediary as in the main model can be captured by assuming $\lambda \in \left(\frac{1}{2}, 1\right]$, and $\lambda = \frac{1}{2}$ corresponds to the unbiased case. With this specification, I can obtain qualitatively the same results, namely:

[22]The details of the analyses are available upon request.

- An ex-post efficient acceptable CE exists if and only if the prior probability π_B of high-type buyer is small enough, $\pi_B \leq \hat{\pi}_B$;
- There exists a threshold $\overline{\lambda}$ on the degree of bias at which the SB outcome changes its regime, and the SB outcomes are those identified in Proposition 2;
- The expected trade surplus decreases if the bias changes from $\lambda \leq \overline{\lambda}$ to $\lambda' > \overline{\lambda}$.

The important difference is that the threshold $\hat{\pi}_B$ is independent of the bias λ, which implies ex-post efficiency can be achieved even when the intermediary's bias is extreme, say $\lambda = 1$.

This specification is also suitable for comparison with some work in the literature such as Matsuo (1989) since the intermediary's payoff is a straightforward extension of it. I discuss the relationship in Appendix D.1.2.

5.2 Intermediary in international relations as in Kydd (2003)

Kydd (2003) studies a biased intermediary in the context of international relations and shows that an intermediary is effective in reducing the probability of conflict only when he is biased. Motivated by his work, consider the following payoff specification:

- If traders (s, b) trade at price p,
  - the seller gets $p - s$;
  - the buyer gets $b - p$;
  - the intermediary gets $\lambda(p - s)$;
- If no trade occurs,
  - the seller and the buyer get 0;
  - the intermediary gets $-c$, where $c \in \mathbb{R}_+$.

The weight $\lambda \in [0, 1]$ represents the degree of the intermediary's bias as before The intermediary is biased toward the seller if $\lambda > 0$ and unbiased if $\lambda = 0$. Contrary to Kydd (2003), I can show that a biased intermediary is not effective in achieving ex-post efficiency. There are many differences in the two models that could have contributed to the opposite conclusions. However, I believe one of the important reasons is that traders do not need to know the opponent's type to achieve ex-post efficiency, whereas in Kydd (2003) the intermediary needs to credibly communicate his private information to reduce the probability of conflict, which is the reason why bias is necessary.[23]

There are many other payoff specifications under which qualitatively the same results can be obtained. What is important for the intermediary's preference is (i) he prefers trade to no trade; and (ii) his payoff is increasing in price. Combined with no commitment, this gives him the incentive to offer only the highest acceptable prices ($\overline{p_{b_H}}$ and $\overline{p_{b_L}}$ in the main model). Using this fact, I can obtain the thresholds for the prior probability or the parameter governing the intermediary's bias under which particular mediation plans constitute equilibrium.

6 Conclusion

This paper has studied the implication of having a biased intermediary in bilateral bargaining who cannot commit to and enforce his decision. I have shown two main results. First, the intermediary's bias is detrimental to ex-post efficiency, however small it is. Second, I characterize the SB outcome and show that the expected social surplus is weakly decreasing in the intermediary's bias. In this sense, a smaller bias is socially desirable.

The results suggest that a biased intermediary is not so useful in achieving efficient outcomes. However, intermediaries are ubiquitous and appreciated even when they are known to be biased as in the case of real estate agents. This discrepancy could be explained by, among others, (i) other roles of intermediaries and (ii) endogenous contracts between traders and intermediaries. Intermediaries' role in practice is not limited to offering prices as in the present model. Traders need to find counterparties before they engage in any bargaining. Intermediaries usually know markets better than traders and can help them in this respect. On top of offering a price, intermediaries may be able to provide some information not available to traders based on their past experiences. Also, traders in practice need to bargain with their intermediaries to determine the commission rates at the beginning of bargaining. These considerations may well turn over

[23] It is also worth mentioning that Kydd (2003) focuses on a fixed communication protocol whereas the present paper considers all admissible communication protocols, leveraging the communication revelation principle.

the conclusion, and even biased intermediaries could be proven to be useful. I plan to explore these extensions in future work.

Appendix

A Proofs

A.1 Proof of Lemma 1

I show the contraposition. Consider an arbitrary mediation plan $\mu \in \Delta(Q \times \{r^A\})$ such that there exists some $q \in \text{supp}(\mu)$ and some $(s,b) \in \Theta_S \times \Theta_B \setminus \{(s_H, b_L)\}$ satisfying $q(s,b) = p \notin \{\overline{p_{b_L}}, \overline{p_{b_H}}\}$, where $\text{supp}(\mu)$ is the support of μ. I show that the intermediary can gain by disobeying the recommendation if he is recommended p. When he is recommended p, his posterior belief that the traders are $(\tilde{s}, \tilde{b})$ is given by

$$\xi(\tilde{s}, \tilde{b} \mid p) = \frac{\Pr(\tilde{s}, \tilde{b}) \sum_{q \in \{q' \in Q:\, q'(\tilde{s}, \tilde{b}) = p\}} \mu(q)}{\sum_{(s,b) \in \Theta_S \times \Theta_B} \Pr(s,b) \sum_{q \in \{q' \in Q:\, q'(s,b) = p\}} \mu(q)},$$

where $\mu(q)$ denotes the probability that the mediator picks up pure mediation plan (q, r^A) when she uses μ. Note that by assumption at least one of $\xi(s_H, b_H \mid p)$, $\xi(s_L, b_H \mid p)$, and $\xi(s_L, b_L \mid p)$ is positive. Given these beliefs, if the traders' are honest and obedient to the mediator, the intermediary's expected payoff from offering $\tilde{p}$ is

$$\begin{cases} [\xi(s_H, b_H \mid p) + \xi(s_L, b_H \mid p)]\,(\delta_S + \delta_B)\tilde{p} & \text{if } \tilde{p} \in \left[\underline{p_{s_H}}, \overline{p_{b_H}}\right] \\ \xi(s_L, b_H \mid p)(\delta_S + \delta_B)\tilde{p} & \text{if } \tilde{p} \in \left(\overline{p_{b_L}}, \underline{p_{s_H}}\right) \\ [\xi(s_L, b_H \mid p) + \xi(s_L, b_L \mid p)]\,(\delta_S + \delta_B)\tilde{p} & \text{if } \tilde{p} \in \left[\underline{p_{s_L}}, \overline{p_{b_L}}\right] \\ 0 & \text{if } \tilde{p} \notin \left[\underline{p_{s_L}}, \overline{p_{b_H}}\right]. \end{cases}$$

From these payoffs, it is easy to see:

- if $p \notin \left[\underline{p_{s_L}}, \overline{p_{b_H}}\right]$, he can gain by offering some $\tilde{p} \in \left[\underline{p_{s_L}}, \overline{p_{b_H}}\right]$;
- if $p \in \left(\overline{p_{b_L}}, \underline{p_{s_H}}\right)$, he can gain by offering either $\overline{p_{b_L}}$ or $\overline{p_{b_H}}$;
- if $p \in \left[\underline{p_{s_H}}, \overline{p_{b_H}}\right)$ or $p \in \left[\underline{p_{s_L}}, \overline{p_{b_L}}\right)$, he can gain by offering either $\overline{p_{b_L}}$ or $\overline{p_{b_H}}$.

Thus μ is not a CE.

A.2 Proof of Lemma 2

Consider an arbitrary mediation plan $\mu \in \Delta\left(Q \times \{r^A\}\right)$ that satisfies (3.1) for all $q \in \mathrm{supp}(\mu)$. By construction all players get the same expected payoffs in μ and μ_x if they are honest and obedient to the mediator. Hence, it suffices to show that μ has weakly more room for profitable manipulations for all players than μ_x. Let $x_{HH} \equiv x_{HH}(\mu) = x_{HH}(\mu_x)$, $x_{LH} \equiv x_{LH}(\mu) = x_{LH}(\mu_x)$, and $x_{LL} \equiv x_{LL}(\mu) = x_{LL}(\mu_x)$. The IC constraints for the high-type seller, the low-type buyer, and the intermediary are the same in μ and μ_x. I show that the high-type buyer and the low-type seller have weakly more room for profitable manipulations, hence weakly more stringent IC constraints in μ than in μ_x.

Consider the high-type buyer b_H. Suppose he misreports his type in μ_x. Since $x_q(s_H, b_L) \in \left(\overline{p_{b_H}}, +\infty\right)$ for all $x_q \in \mathrm{supp}(\mu_x)$, he can secure a positive payoff of $b_H - (1+\delta_B)\overline{p_{b_L}}$ only when the seller is of low-type and price $\overline{p_{b_L}}$ is offered. Hence, his expected payoff from misreporting is at most

$$(1-\pi_S)(1-x_{LL})\left[b_H - (1+\delta_B)\overline{p_{b_L}}\right].$$

In μ, however, he may be able to get a positive payoff when the seller is of high-type; he can secure $b_H - (1+\delta_B)p$ if $p \in \left[\underline{p_{s_H}}, \overline{p_{b_H}}\right)$ is offered. Hence, his expected payoff from misreporting in μ is at most

$$\begin{aligned}&(1-\pi_S)(1-x_{LL})\left[b_H - (1+\delta_B)\overline{p_{b_L}}\right]\\ &+\pi_S \sum_{p\in[\underline{p_{s_H}}, \overline{p_{b_H}})} \sum_{q\in\{q'\in Q:\ q'(s_H,b_L)=p\}} \mu(q)[b_H - (1+\delta_B)p],\end{aligned}$$

where the second term is nonnegative. Thus, high-type buyer's IC constraint in μ is at least as stringent as that in μ_x.

Next, consider the low-type seller s_L. Suppose he misreports his type in μ_x. Since $x_q(s_H, b_L) \in \left(\overline{p_{b_H}}, +\infty\right)$ for all $x_q \in \mathrm{supp}(\mu_x)$, he can secure a positive payoff only when the buyer is of high-type; he can get $(1-\delta_S)\overline{p_{b_H}} - s_L$ if price $\overline{p_{b_H}}$ is offered and $(1-\delta_S)\overline{p_{b_L}} - s_L$ if $\overline{p_{b_L}}$ is offered. Hence, his expected payoff from misreporting is at most

$$\pi_B\left[x_{HH}(1-\delta_S)\overline{p_{b_H}} + (1-x_{HH})(1-\delta_S)\overline{p_{b_L}} - s_L\right].$$

In μ, however, he may be able to get a positive payoff when the buyer is of low-type; he can secure $(1-\delta_S)p - s_L$ if $p \in \left(\underline{p_{s_L}}, \overline{p_{b_L}}\right]$ is offered. Hence, his expected payoff from misreporting in μ is at most

$$\pi_B\left[x_{HH}(1-\delta_S)\overline{p_{b_H}} + (1-x_{HH})(1-\delta_S)\overline{p_{b_L}} - s_L\right] \\ + (1-\pi_B) \sum_{p \in (\underline{p_{s_L}}, \overline{p_{b_L}}]} \sum_{q \in \{q' \in Q:\ q'(s_H, b_L) = p\}} \mu(q)[(1-\delta_S)p - s_L],$$

where the second term is nonnegative. Thus, low-type seller's IC constraint in μ is at least as stringent as that in μ_x.

The above argument suffices for the proof of Lemma 2. However, I spell out the players' IC constraints in μ_x in order for the coming analysis.

High-type seller s_H If he is honest and obedient to the mediator, he gets $(1-\delta_S)\overline{p_{b_H}} - s_H$ when the buyer is of high-type and price $\overline{p_{b_H}}$ is offered and gets 0 otherwise. If he misreports his type, he can secure at most $(1-\delta_S)\overline{p_{b_H}} - s_H$ if the buyer is of high-type and $\overline{p_{b_H}}$ is offered. Hence, he has no incentive to manipulate μ_x if

$$\pi_B x_{HH}\left[(1-\delta_S)\overline{p_{b_H}} - s_H\right] \geq \pi_B x_{LH}\left[(1-\delta_S)\overline{p_{b_H}} - s_H\right] \iff x_{HH} \geq x_{LH}. \tag{IC-s_H}$$

High-type buyer b_H If he is honest and obedient to the mediator, he gets $b_H - (1+\delta_B)\overline{p_{b_L}}$ if the seller is of low-type and $\overline{p_{b_L}}$ is offered and gets 0 otherwise. Combined with the preceding argument, he has no incentive to manipulate μ_x if

$$(1-\pi_S)(1-x_{LH})\left[b_H - (1+\delta_B)\overline{p_{b_L}}\right] \geq (1-\pi_S)(1-x_{LL})\left[b_H - (1+\delta_B)\overline{p_{b_L}}\right] \\ \iff x_{LL} \geq x_{LH}. \tag{IC-b_H}$$

Low-type seller s_L If he is honest and obedient to the mediator, he gets $(1-\delta_S)\overline{p_{b_H}} - s_L$ when the buyer is of high-type and $\overline{p_{b_H}}$ is offered, gets $(1-\delta_S)\overline{p_{b_L}} - s_L$ when $\overline{p_{b_L}}$ is offered regardless of the buyer's type, and gets 0 otherwise. Combined with the preceding argument, he has no incentive to manipulate μ_x if

$$\begin{aligned}
&\pi_B x_{LH}\left[(1-\delta_S)\overline{p_{b_H}}-s_L\right]+\left[\pi_B(1-x_{LH})+(1-\pi_B)(1-x_{LL})\right]\left[(1-\delta_S)\overline{p_{b_L}}-s_L\right]\\
&\quad\geq \pi_B\left[x_{HH}(1-\delta_S)\overline{p_{b_H}}+(1-x_{HH})(1-\delta_S)\overline{p_{b_L}}-s_L\right]\\
\iff &(1-\pi_B)\left[(1-\delta_S)\overline{p_{b_L}}-s_L\right](1-x_{LL})\geq \pi_B(1-\delta_S)\left(\overline{p_{b_H}}-\overline{p_{b_L}}\right)(x_{HH}-x_{LH}).
\end{aligned} \tag{IC-s_L}$$

Low-type buyer b_L His expected payoff is 0 if he is honest and obedient to the mediator, and his expected payoff is at most 0 whatever manipulation he uses. Hence, his IC constraint is trivially satisfied.

Intermediary His posterior beliefs when he is recommended $\overline{p_{b_H}}$ are given by

$$\begin{aligned}
\xi(s_H,b_H\mid b_H)&=\frac{\pi_S\pi_B x_{HH}}{\pi_S\pi_B x_{HH}+(1-\pi_S)\pi_B x_{LH}+(1-\pi_S)(1-\pi_B)x_{LL}}\\
\xi(s_L,b_H\mid b_H)&=\frac{(1-\pi_S)\pi_B x_{LH}}{\pi_S\pi_B x_{HH}+(1-\pi_S)\pi_B x_{LH}+(1-\pi_S)(1-\pi_B)x_{LL}}\\
\xi(s_L,b_L\mid b_H)&=\frac{(1-\pi_S)(1-\pi_B)x_{LL}}{\pi_S\pi_B x_{HH}+(1-\pi_S)\pi_B x_{LH}+(1-\pi_S)(1-\pi_B)x_{LL}}\\
\xi(s_H,b_L\mid b_H)&=0.
\end{aligned}$$

By a similar argument as in the proof of Lemma 1, the intermediary finds it optimal to offer either $\overline{p_{b_L}}$ or $\overline{p_{b_H}}$. Hence, following the recommendation is optimal for him if he prefers offering $\overline{p_{b_H}}$ to $\overline{p_{b_L}}$:

$$\begin{aligned}
&\left[\xi(s_H,b_H\mid b_H)+\xi(s_L,b_H\mid b_H)\right](\delta_S+\delta_B)\overline{p_{b_H}}\\
&\quad\geq\left[\xi(s_L,b_H\mid b_H)+\xi(s_L,b_L\mid b_H)\right](\delta_S+\delta_B)\overline{p_{b_L}}\\
\iff&\left[\pi_S\pi_B x_{HH}+(1-\pi_S)\pi_B x_{LH}\right]b_H\\
&\quad\geq\left[(1-\pi_S)\pi_B x_{LH}+(1-\pi_S)(1-\pi_B)x_{LL}\right]b_L\\
\iff&\pi_S\pi_B b_H x_{HH}+(1-\pi_S)\pi_B(b_H-b_L)x_{LH}\geq(1-\pi_S)(1-\pi_B)b_L x_{LL}.
\end{aligned}\tag{IC-Int1}$$

Similarly, his posterior beliefs when he is recommended $\overline{p_{b_L}}$ are given by

$$\xi(s_H, b_H \mid b_L) = \frac{\pi_S\pi_B(1-x_{HH})}{\pi_S\pi_B(1-x_{HH}) + (1-\pi_S)\pi_B(1-x_{LH}) + (1-\pi_S)(1-\pi_B)(1-x_{LL})}$$
$$\xi(s_L, b_H \mid b_L) = \frac{(1-\pi_S)\pi_B(1-x_{LH})}{\pi_S\pi_B(1-x_{HH}) + (1-\pi_S)\pi_B(1-x_{LH}) + (1-\pi_S)(1-\pi_B)(1-x_{LL})}$$
$$\xi(s_L, b_L \mid b_L) = \frac{(1-\pi_S)(1-\pi_B)(1-x_{LL})}{\pi_S\pi_B(1-x_{HH}) + (1-\pi_S)\pi_B(1-x_{LH}) + (1-\pi_S)(1-\pi_B)(1-x_{LL})}$$
$$\xi(s_H, b_L \mid b_L) = 0.$$

By the same argument as in the previous case, following the recommendation is optimal for him if he prefers offering $\overline{p_{b_L}}$ to $\overline{p_{b_H}}$:

$$\begin{aligned}
&[\xi(s_L, b_H \mid b_L) + \xi(s_L, b_L \mid b_L)]\,(\delta_S+\delta_B)\overline{p_{b_L}} \\
&\quad \geq [\xi(s_H, b_H \mid b_L) + \xi(s_L, b_H \mid b_L)]\,(\delta_S+\delta_B)\overline{p_{b_H}} \\
\iff &[(1-\pi_S)\pi_B(1-x_{LH}) + (1-\pi_S)(1-\pi_B)(1-x_{LL})]\,b_L \\
&\quad \geq [\pi_S\pi_B(1-x_{HH}) + (1-\pi_S)\pi_B(1-x_{LH})]\,b_H \\
\iff &(1-\pi_S)(1-\pi_B)b_L(1-x_{LL}) \geq \pi_S\pi_B b_H(1-x_{HH}) + (1-\pi_S)\pi_B(b_H-b_L)(1-x_{LH}).
\end{aligned} \tag{IC-Int2}$$

When he is recommended $p \notin \{\overline{p_{b_L}}, \overline{p_{b_H}}\}$, he believes that the traders are (s_L, b_H) with probability 1. Since they accept no offer, following the recommendation is always optimal for him.

A.3 Proof of Proposition 1

If $(x_{HH}, x_{LH}, x_{LL}) = (1,0,0)$, then (IC-$s_H$), (IC-$b_H$), and (IC-Int1) are clearly satisfied. (IC-s_L) is reduced to

$$\begin{aligned}
&(1-\pi_B)\left[(1-\delta_S)\overline{p_{b_L}} - s_L\right] \geq \pi_B(1-\delta_S)\left(\overline{p_{b_H}} - \overline{p_{b_L}}\right) \\
\iff &\pi_B \leq \frac{(1-\delta_S)\overline{p_{b_L}} - s_L}{(1-\delta_S)\overline{p_{b_H}} - s_L} \\
\iff &\pi_B \leq \frac{(1-\delta_S)b_L - (1+\delta_B)s_L}{(1-\delta_S)b_H - (1+\delta_B)s_L}.
\end{aligned} \tag{A.1}$$

(IC-Int2) is reduced to

$$(1-\pi_S)(1-\pi_B)b_L \geq (1-\pi_S)\pi_B(b_H - b_L) \iff \pi_B \leq \frac{b_L}{b_H}. \tag{A.2}$$

Since $\frac{(1-\delta_S)b_L-(1+\delta_B)s_L}{(1-\delta_S)b_H-(1+\delta_B)s_L} \leq \frac{b_L}{b_H}$ with equality at $s_L = 0$, an ex-post efficient acceptable CE exists if and only if (A.1) holds.

A.4 Proof of Proposition 2

First, I characterize the solution to the linear program (P) when $\pi_B > \overline{\pi_B}$. Rearranging the constraints, (P) can be written as follows:

$$\begin{aligned}
\max_{x=(x_{HH},x_{LH},x_{LL})\in\mathbb{R}^3} \quad & \pi_S\pi_B(b_H - s_H)x_{HH} + (1-\pi_S)\pi_B(b_H - s_L) \\
& + (1-\pi_S)(1-\pi_B)(b_L - s_L)(1-x_{LL}) \\
\text{subject to} \quad & g_1(x) = x_{HH} - x_{LH} \geq 0 \\
& g_2(x) = x_{LL} - x_{LH} \geq 0 \\
& g_3(x) = (1-\pi_B)\left[(1-\delta_S)\overline{p_{b_L}} - s_L\right](1-x_{LL}) \\
& \qquad - \pi_B(1-\delta_S)\left(\overline{p_{b_H}} - \overline{p_{b_L}}\right)(x_{HH} - x_{LH}) \geq 0 \\
& g_4(x) = \pi_S\pi_B b_H x_{HH} + (1-\pi_S)\pi_B(b_H - b_L)x_{LH} - (1-\pi_S)(1-\pi_B)b_L x_{LL} \geq 0 \\
& g_5(x) = -\pi_S\pi_B b_H(1-x_{HH}) - (1-\pi_S)\pi_B(b_H - b_L)(1-x_{LH}) \\
& \qquad + (1-\pi_S)(1-\pi_B)b_L(1-x_{LL}) \geq 0 \\
& g_6(x) = 1 - x_{HH} \geq 0 \\
& g_7(x) = 1 - x_{LH} \geq 0 \\
& g_8(x) = 1 - x_{LL} \geq 0 \\
& g_9(x) = x_{HH} \geq 0 \\
& g_{10}(x) = x_{LH} \geq 0 \\
& g_{11}(x) = x_{LL} \geq 0
\end{aligned}$$

Let η_i denote the Lagrange multiplier associated with the g_i. By the KKT conditions, $(x, \eta) \in \mathbb{R}^3 \times \mathbb{R}^{11}_{+}$ is the solution to the program if and only if it satisfies

Primal feasibility: $g_i(x) \geq 0$ for all i.

Dual feasibility: $\eta_i \geq 0$ for all i, and

$$\pi_S\pi_B(b_H - s_H) + \eta_1 - \pi_B(1-\delta_S)(\overline{p_{b_H}} - \overline{p_{b_L}})\eta_3 + \pi_S\pi_B b_H(\eta_4 + \eta_5) - \eta_6 + \eta_9 = 0 \quad \text{(Dual 1)}$$

$$-\eta_1 - \eta_2 + \pi_B(1-\delta_S)(\overline{p_{b_H}} - \overline{p_{b_L}})\eta_3 + (1-\pi_S)\pi_B(b_H - b_L)(\eta_4 + \eta_5) - \eta_7 + \eta_{10} = 0 \quad \text{(Dual 2)}$$

$$\begin{aligned} &-(1-\pi_S)(1-\pi_B)(b_L - s_L) + \eta_2 - (1-\pi_B)\left[(1-\delta_S)\overline{p_{b_L}} - s_L\right]\eta_3 \\ &\quad -(1-\pi_S)(1-\pi_B)b_L(\eta_4+\eta_5) - \eta_8 + \eta_{11} = 0. \end{aligned} \quad \text{(Dual 3)}$$

Complementary slackness: $\eta_i g_i(x) = 0$ for all i.

For the sake of exposition, I introduce the following notations:

$$\begin{aligned} \overline{x_{HH}} &= \frac{(1-\pi_B)\left[(1-\delta_S)\overline{p_{b_L}} - s_L\right]}{\pi_B(1-\delta_S)\left(\overline{p_{b_H}} - \overline{p_{b_L}}\right)} \\ J &= \pi_S\pi_B b_H + (1-\pi_S)\pi_B(b_H - b_L) - (1-\pi_S)(1-\pi_B)b_L \\ K &= \pi_S\pi_B s_H(1-\overline{x_{HH}}) + (1-\pi_S)\pi_B(b_H - b_L) - (1-\pi_S)(1-\pi_B)s_L. \end{aligned}$$

Note that $\overline{x_{HH}}$ is the value of x_{HH} at which $g_3(x) \geq 0$ binds when $x_{LH} = x_{LL} = 0$ and that $-J$ is the sum of the constant terms in $g_5(x)$. As such, the larger the value of J, the larger the intermediary's incentive to disobey when $\overline{p_{b_L}}$ is recommended.

Case 1: When $J \leq \pi_S \pi_B b_H \overline{x_{HH}} + \min\{K, 0\}$[24]

I show that

$$
\begin{aligned}
x^* &= (\overline{x_{HH}}, 0, 0) \\
\eta_2^* &= \pi_S \pi_B (b_H - s_H) \\
\eta_3^* &= \frac{\pi_S (b_H - s_H)}{(1-\delta_S)(\overline{p_{b_H}} - \overline{p_{b_L}})} \\
\eta_{11}^* &= (1-\pi_S)(1-\pi_B)(b_L - s_L) - \pi_S \pi_B (b_H - s_H)(1 - \overline{x_{HH}}) \\
\eta_i^* &= 0 \ \text{ for all } i \notin \{2, 3, 11\}
\end{aligned}
$$

are the solution to the program.

Primal feasibility: It is easy to see $g_i(x^*) \geq 0$ for all $i \neq \{5, 6\}$. Note that $g_5(x^*) \geq 0$ is reduced to $J \leq \pi_S \pi_B b_H \overline{x_{HH}}$ and that $g_6(x^*) = 1 - \overline{x_{HH}} \geq 0 \iff \pi_B \geq \overline{\pi_B}$.

Dual feasibility: η_2^* and η_3^* are positive, and

$$\eta_{11} \geq 0 \iff J \leq \pi_S \pi_B b_H \overline{x_{HH}} + K.$$

The left-side of (Dual 1) is reduced to

$$\pi_S \pi_B (b_H - s_H) - \pi_S \pi_B (b_H - s_H) = 0.$$

(Dual 2) is satisfied since $\eta_2^* = \pi_B (1-\delta_S)(\overline{p_{b_H}} - \overline{p_{b_L}}) \eta_3^*$. The left-side of (Dual 3) is reduced to

$$\pi_S \pi_B (b_H - s_H) \overline{x_{HH}} - \underbrace{(1-\pi_B)\left[(1-\delta_S)\overline{p_{b_L}} - s_L\right]}_{= \pi_B (1-\delta_S)(\overline{p_{b_H}} - \overline{p_{b_L}}) \overline{x_{HH}}} \frac{\pi_S (b_H - s_H)}{(1-\delta_S)(\overline{p_{b_H}} - \overline{p_{b_L}})} = 0.$$

Complementary slackness: The conditions are satisfied since $\eta_i^* = 0$ for all $i \notin \{2, 3, 11\}$ and $g_2(x^*) = g_3(x^*) = g_{11}(x^*) = 0$.

[24] One can show that this is the only case that can happen if $K < 0$. Hence $K \geq 0$ in the remaining cases.

Case 2: When $\pi_S \pi_B b_H \overline{x_{HH}} < J \leq \pi_S \pi_B b_H \overline{x_{HH}} + K$

I show that

$$\begin{aligned}
x^* &= (1, 1, 1) \\
\eta_2^* &= \pi_S \pi_B (b_H - s_H) + [\pi_S \pi_B b_H + (1 - \pi_S) \pi_B (b_H - b_L)] \, \eta_5^* \\
\eta_3^* &= \frac{\pi_S \pi_B (b_H - s_H) + \pi_S \pi_B b_H \eta_5^*}{\pi_B (1 - \delta_S)(\overline{p_{b_H}} - \overline{p_{b_L}})} \\
\eta_5^* &= \frac{(1 - \pi_S)(1 - \pi_B)(b_L - s_L) - \pi_S \pi_B (b_H - s_H)(1 - \overline{x_{HH}})}{J - \pi_S \pi_B b_H \overline{x_{HH}}} \\
\eta_i^* &= 0 \ \text{ for all } i \notin \{2, 3, 5\}
\end{aligned}$$

are the solution to the program.

Primal feasibility: It is easy to see $g_i(x^*) \geq 0$ for all $i \neq 4$ and $g_4(x^*) = J > 0$ by assumption.

Dual feasibility: Note that η_2^* and η_3^* are positive if η_5^* is nonnegative, which is the case if

$$\pi_S \pi_B b_H \overline{x_{HH}} < J \leq \pi_S \pi_B b_H \overline{x_{HH}} + K.$$

The left-side of (Dual 1) is reduced to

$$\pi_S \pi_B (b_H - s_H) - \pi_S \pi_B (b_H - s_H) - \pi_S \pi_B b_H \eta_5^* + \pi_S \pi_B b_H \eta_5^* = 0.$$

The left-side of (Dual 2) is reduced to

$$\begin{aligned}
&- \pi_S \pi_B (b_H - s_H) - [\pi_S \pi_B b_H + (1 - \pi_S) \pi_B (b_H - b_L)] \eta_5^* \\
&\qquad + \pi_S \pi_B (b_H - s_H) + \pi_S \pi_B b_H \eta_5^* + (1 - \pi_S) \pi_B (b_H - b_L) \eta_5^* = 0.
\end{aligned}$$

The left-side of (Dual 3) is reduced to

$$
\begin{aligned}
& -(1-\pi_S)(1-\pi_B)(b_L-s_L)+\pi_S\pi_B(b_H-s_H)+[\pi_S\pi_B b_H+(1-\pi_S)\pi_B(b_H-b_L)]\,\eta_5^* \\
& \qquad -\underbrace{(1-\pi_B)\left[(1-\delta_S)\overline{p_{b_L}}-s_L\right]}_{=\pi_B(1-\delta_S)(\overline{p_{b_H}}-\overline{p_{b_L}})\overline{x_{HH}}}\frac{\pi_S\pi_B(b_H-s_H)+\pi_S\pi_B b_H\eta_5^*}{\pi_B(1-\delta_S)(\overline{p_{b_H}}-\overline{p_{b_L}})}-(1-\pi_S)(1-\pi_B)b_L\eta_5^* \\
= \; & \pi_S\pi_B(b_H-s_H)(1-\overline{x_{HH}})-(1-\pi_S)(1-\pi_B)(b_L-s_L)+[J-\pi_s\pi_B b_H\overline{x_{HH}}]\eta_5^* \\
= \; & 0.
\end{aligned}
$$

Complementary slackness: The conditions are satisfied since $\eta_i^*=0$ for all $i\notin\{2,3,5\}$ and $g_2(x^*)=g_3(x^*)=g_5(x^*)=0$.

Case 3: When $J>\pi_S\pi_B b_H\overline{x_{HH}}+K$

I show that

$$
\begin{aligned}
& x^*=(1,1,1) \\
& \eta_2^*=\pi_S\pi_B(b_H-s_H) \\
& \eta_3^*=\frac{\pi_S(b_H-s_H)}{(1-\delta_S)(\overline{p_{b_H}}-\overline{p_{b_L}})} \\
& \eta_8^*=\pi_S\pi_B(b_H-s_H)(1-\overline{x_{HH}})-(1-\pi_S)(1-\pi_B)(b_L-s_L) \\
& \eta_i^*=0 \ \text{ for all } i\notin\{2,3,8\}
\end{aligned}
$$

are the solution to the program.

Primal feasibility: It is easy to see $g_i(x^*)\geq 0$ for all $i\neq 4$ and $g_4(x^*)=J>0$ by assumption.

Dual feasibility: Note that η_2^* and η_3^* coincide with those in Case 1 and that $-\eta_8^*$ is equal to "η_{11}^* in Case 1." Thus, the dual feasibility is satisfied if

$$J\geq \pi_S\pi_B b_H\overline{x_{HH}}+K.$$

Complementary slackness: The conditions are satisfied since $\eta_i^*=0$ for all $i\notin\{2,3,8\}$ and $g_2(x^*)=g_3(x^*)=g_8(x^*)=0$.

In summary, the solution to the linear program (P) when $\pi_B > \overline{\pi_B}$ is

- $\left(x^*_{HH}, x^*_{LH}, x^*_{LL}\right) = \left(\overline{x_{HH}}, 0, 0\right)$ if $J \leq \pi_S \pi_B b_H \overline{x_{HH}} + \min\{K, 0\}$;
- $\left(x^*_{HH}, x^*_{LH}, x^*_{LL}\right) = (1, 1, 1)$ otherwise.

The above conditions imply two upper bounds on the ratio $\frac{1+\delta_B}{1-\delta_S}$:

$$J \leq \pi_S \pi_B b_H \overline{x_{HH}} \iff \frac{1+\delta_B}{1-\delta_S} \leq \frac{\pi_S(1-\pi_B) b_H b_L - (b_H - b_L) J}{\pi_S (1-\pi_B) b_H s_L} \tag{A.3}$$

and

$$\begin{aligned} & J \leq \pi_S \pi_B b_H \overline{x_{HH}} + K \\ \iff & \frac{1+\delta_B}{1-\delta_S} \leq \frac{\pi_S(1-\pi_B)(b_H - s_H) b_L - (b_H - b_L)\left[\pi_S \pi_B (b_H - s_H) - (1-\pi_S)(1-\pi_B)(b_L - s_L)\right]}{\pi_S(1-\pi_B)(b_H - s_H) s_L}. \end{aligned} \tag{A.4}$$

The threshold $\overline{h}$ is given by the smaller of the two bounds.

A.5 Proof of Corollary 1

For $h \leq \overline{h}$, the expected social surplus is given by

$$\pi_S \pi_B (b_H - s_H) \overline{x_{HH}} + (1-\pi_S)\pi_B (b_H - s_L) + (1-\pi_S)(1-\pi_B)(b_L - s_L).$$

If $h < \overline{h}$ increases to h' such that $h < h' \leq \overline{h}$, the expected surplus decreases as $\overline{x_{HH}}$ is decreasing in h. If $h \leq \overline{h}$ increases to $h' > \overline{h}$, the expected surplus is now given by

$$\pi_S \pi_B (b_H - s_H) + (1-\pi_S)\pi_B (b_H - s_L).$$

Hence the difference is

$$
\begin{aligned}
&-\pi_S\pi_B(b_H - s_H)(1 - \overline{x_{HH}}) + (1-\pi_S)(1-\pi_B)(b_L - s_L) \\
&= \pi_S\pi_B b_H \overline{x_{HH}} + K - J \\
&\geq 0,
\end{aligned}
$$

where the last inequality follows from the fact that $h \leq \overline{h}$ and hence $J \leq \pi_S\pi_B b_H + K$.

B On CE

In Section B.1, I define SCE and show that any CE is SCE in the mediated bargaining game. In Section B.2, I show the equivalence between ex-ante IC and interim IC.

B.1 Equivalence between CE and SCE

The definition of SCE is involved, but Myerson (1986b, Theorem 2) shows that a mediation plan is an SCE if and only if it is a CE that would never recommend a "codominated action" to any player who has not lied to the mediator. Thus, it suffices to identify the set codominated actions and the set of CE.

To define codominated actions, I introduce some notations. Let C be a correspondence such that

$$
\begin{aligned}
&C(h^0) \subseteq \mathbb{R}_+, \\
&C(\theta_i, p) \subseteq \{Y, N\},\ \forall i \in \{S, B\},\ \forall(\theta_i, p) \in \Theta_i \times \mathbb{R}_+,
\end{aligned}
$$

where h^0 denotes the initial history. For any such C, let $E(C)$ be the set of mediation plans that never recommend actions in $C(\theta_i, p)$ for all $i \in \{S, B\}$ and all $(\theta_i, p) \in \Theta_i \times \mathbb{R}_+$; that is,

$$
E(C) = \{(q, r) \in Q \times R \colon \forall(s, b, p) \in \Theta_S \times \Theta_B \times \mathbb{R}_+,\ r_S(p \mid s, b) \notin C(s, p),\ r_B(p \mid s, b) \notin C(b, p)\}.
$$

For any $p \in \mathbb{R}_+$, let $\phi^1(p)$ be the set of tuples consisting of a mediation plan and the traders' reports such that the mediator recommends p; that is,

$$\phi^1(p) = \{(q, r, s, b) \in Q \times R \times \Theta_S \times \Theta_B \colon q(s, b) = p\}.$$

For each $i \in \{S, B\}$, for any $a_i \in \{Y, N\}$ and any $(\theta_i, p) \in \Theta_i \times \mathbb{R}_+$, let $\phi^2(a_i, \theta_i, p)$ be the set of tuples consisting of a mediation plan, the traders' types, and a price such that i's type is θ_i and the mediator recommends a_i to i as a response to p if the traders do not lie; that is, for $i = S$,

$$\phi^2(a_S, \theta_S, p) = \{(q, r, s, b, p) \in Q \times R \times \Theta_S \times \Theta_B \times \mathbb{R}_+ \colon s = \theta_S,\ r_S(p \mid s, b) = a_S\},$$

and for $i = B$,

$$\phi^2(a_B, \theta_B, p) = \{(q, r, s, b, p) \in Q \times R \times \Theta_S \times \Theta_B \times \mathbb{R}_+ \colon b = \theta_B,\ r_B(p \mid s, b) = a_B\}.$$

Let $\phi^1(C)$ be the union of all sets $\phi^1(p)$ over all $p \in C(h^0)$ and $\phi^2(C)$ be the union of all sets $\phi^2(a_i, \theta_i, p)$ over all $i \in \{S, B\}$, $(\theta_i, p) \in \Theta_i \times \mathbb{R}_+$, and $a_i \in C(\theta_i, p)$; that is,

$$\begin{aligned}
&\phi^1(C) = \{(q, r, s, b) \in Q \times R \times \Theta_S \times \Theta_B \colon q(s, b) \in C(h^\emptyset)\},\\
&\phi^2(C) = \{(q, r, s, b, p) \in Q \times R \times \Theta_S \times \Theta_B \times \mathbb{R}_+ : r_S(p \mid s, b) \in C(s, p) \text{ or } r_B(p \mid s, b) \in C(b, p)\}.
\end{aligned}$$

Now I can define codomination correspondence. C is a *codomination correspondence* if

1. for each $\tau^1 \in \Delta(Q \times R \times \Theta_S \times \Theta_B)$, if $\tau^1(E(C) \times \Theta_S \times \Theta_B) = 1$ and $\tau^1(\phi^1(C)) > 0$, then there exists some $p \in C(h^0)$ and some $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\begin{aligned}
&\sum_{(q, r, s, b) \in \phi^1(p)} \tau^1(q, r, s, b) \cdot v(\alpha(p) \mid s, b) \cdot \mathbf{1}_{\{r(\alpha(p) \mid s, b) = (Y, Y)\}}\\
&> \sum_{(q, r, s, b) \in \phi^1(p)} \tau^1(q, r, s, b) \cdot v(p \mid s, b) \cdot \mathbf{1}_{\{r(p \mid s, b) = (Y, Y)\}}
\end{aligned} \tag{B.1}$$

   and

2. for each $\tau^2 \in \Delta(Q \times R \times \Theta_S \times \Theta_B \times \mathbb{R}_+)$, if $\tau^2(\phi^2(C)) > 0$, then there exists some trader $i \in \{S, B\}$, some $(\theta_i, p) \in \Theta_i \times \mathbb{R}_+$, some $a_i \in C(\theta_i, p)$, and some $\gamma_i : \Theta_i \times \mathbb{R}_+ \to \{Y, N\}$

such that, if $i = S$,

$$\sum_{(q,r,s,b,p)\in\phi^2(a_S,\theta_S,p)} \tau^2(q,r,s,b,p)\cdot(p-s)\cdot \mathbf{1}_{\{\gamma_S(s,p)=Y\cap r_B(p|s,b)=Y\}}$$
$$> \sum_{(q,r,s,b,p)\in\phi^2(a_S,\theta_S,p)} \tau^2(q,r,s,b,p)\cdot(p-s)\cdot \mathbf{1}_{\{a_S=Y\cap r_B(p|s,b)=Y\}} \tag{B.2}$$

or, if $i = B$,

$$\sum_{(q,r,s,b,p)\in\phi^2(a_B,\theta_B,p)} \tau^2(q,r,s,b,p)\cdot(b-p)\cdot \mathbf{1}_{\{\gamma_B(b,p)=Y\cap r_S(p|s,b)=Y\}}$$
$$> \sum_{(q,r,s,b,p)\in\phi^2(a_B,\theta_B,p)} \tau^2(q,r,s,b,p)\cdot(b-p)\cdot \mathbf{1}_{\{a_B=Y\cap r_S(p|s,b)=Y\}}. \tag{B.3}$$

Let D denote the union of all codomination correspondences.[25] The intermediary's offer $p \in \mathbb{R}_+$ is *codominated* if $p \in D(h^0)$. Trader i's response $a_i \in \{Y, N\}$ is *codominated* at the history $(\theta_i, p) \in \Theta_i \times \mathbb{R}_+$ if $a_i \in D(\theta_i, p)$. Intuitively, if codominated actions could be recommended with positive probability, then at least one player could expect to gain by manipulation after being told to use a codominated action.

Lemma 4 shows that no action is codominated in the mediated bargaining game, which implies that every CE is an SCE.

Lemma 4. *No action is codominated in the mediated bargaining game.*

Proof. I show that the only codomination correspondence in the mediated bargaining game is C such that

$$C(h^0) = \emptyset \tag{B.4}$$
$$C(\theta_i, p) = \emptyset,\ \forall i \in \{S, B\},\ \forall(\theta_i, p) \in \Theta_i \times \mathbb{R}_+. \tag{B.5}$$

First, I show (B.5). Note that $C(\theta_i, p) \subseteq \{Y, N\}$ is either $\emptyset$, $\{Y\}$, $\{N\}$, or $\{Y, N\}$ for any $i \in \{S, B\}$ and any $(\theta_i, p) \in \Theta_i \times \mathbb{R}_+$. Fix arbitrary $(s, b, p) \in \Theta_S \times \Theta_B \times \mathbb{R}_+$.

[25] In general, this union may not be well defined if type spaces or action spaces are not finite. Hhowever, there is the unique codomination correspondence in the present model (see Lemma 4). Thus, this concern has no bite.

1. If $C(s,p) = \{Y\}$ and $C(b,p) = \{Y\}$ or $\emptyset$, take $\tau^2 \in \Delta(Q \times R \times \Theta_S \times \Theta_B \times \mathbb{R}_+)$ such that $\tau^2(q,r,s,b,p) = 1$ for $q(s,b) = p$ and $r(p \mid s,b) = (Y,N)$. The seller is recommended an action in C but no manipulation is profitable for him, which violates (B.2).

2. If $C(s,p) = \emptyset$ and $C(b,p) = \{Y\}$, take $\tau^2 \in \Delta(Q \times R \times \Theta_S \times \Theta_B \times \mathbb{R}_+)$ such that $\tau^2(q,r,s,b,p) = 1$ for $q(s,b) = p$ and $r(p \mid s,b) = (N,Y)$. The buyer is recommended an action in C but no manipulation is profitable for him, which violates (B.3).

3. In all the other cases except for $C(s,p) = C(b,p) = \emptyset$, take $\tau^2 \in \Delta(Q \times R \times \Theta_S \times \Theta_B \times \mathbb{R}_+)$ such that $\tau^2(q,r,s,b,p) = 1$ for $q(s,b) = p$ and $r(p \mid s,b) = (N,N)$. At least one trader is recommended an action in C but no manipulation is profitable for him, which violates (B.2) or (B.3).

4. If $C(s,p) = \emptyset$ and $C(b,p) = \emptyset$, then the definition is trivially satisfied.

Since I have chosen (s,b,p) arbitrarily, a codomination correspondence must satisfy $C(s,p) = \emptyset$ and $C(b,p) = \emptyset$ for all $(s,b,p) \in \Theta_S \times \Theta_B \times \mathbb{R}_+$.

Next, I show (B.4). Suppose $C(h^0) \neq \emptyset$. For each $p \in C(h^0)$, take $\tau^1 \in \Delta(Q \times R \times \Theta_S \times \Theta_B)$ such that $\tau^1(q,r,s,b) = 1$ for $q(s,b) = p$ and $r(p' \mid s,b) = (N,N)$ for all $p' \in \mathbb{R}_+$. The intermediary is recommended an action in C but no manipulation is profitable for him, which violates (B.1). Hence, $C(h^0)$ must be empty as well. Therefore, the correspondence C satisfying (B.4) and (B.5) is the only codomination correspondence in the mediated barganing game. $\square$

B.2 Equivalence between ex-ante IC and interim IC

As I only consider pure mediation plans or mixed mediation plans in $\Delta\left(Q \times \{r^A\}\right)$ in the paper, hereafter I focus on mediation plans $\mu \in \Delta\left(Q \times \{r\}\right)$ for some arbitrary $r \in R$. I show that the ex-ante IC constraints (2.1) and (2.2) are satisfied if and only if the interim IC constraints are satisfied.

For trader $i \in \{S,B\}$, let $U_i(q,r \mid \theta_i)$ be type θ_i's interim expected payoff if the mediator uses a pure mediation plan (q, r) and all players are honest and obedient to the mediator:

$$U_S(q, r \mid s) = \sum_{b \in \Theta_B} \Pr(b) \cdot (q(s, b) - s) \cdot \mathbf{1}_{\{r(q(s,b)|s,b)=(Y,Y)\}},$$
$$U_B(q, r \mid b) = \sum_{s \in \Theta_S} \Pr(s) \cdot (b - q(s, b)) \cdot \mathbf{1}_{\{r(q(s,b)|s,b)=(Y,Y)\}},$$

where $\Pr(\theta_i)$ is the prior probability that trader i is of type θ_i. Suppose that the mediator uses a mediation plan $\mu \in \Delta(Q \times \{r\})$. If all players are honest and obedient to the mediator, then trader i of type θ_i expects to get

$$\sum_{q \in Q} \mu(q) U_i(q, r \mid \theta_i).$$

Let $V(\mu \mid p)$ be the intermediary's interim expected payoff if he receives the recommendation p and all players are honest and obedient to the mediator:

$$V(\mu \mid p) = \sum_{(s,b) \in \Theta_S \times \Theta_B} \xi(s, b \mid p) \cdot v(p \mid s, b) \cdot \mathbf{1}_{\{r(p|s,b)=(Y,Y)\}},$$

where $\xi(s, b \mid p)$ is the intermediary's posterior belief that the traders are (s, b) when he receives p, which is given by

$$\xi(s, b \mid p) = \frac{\Pr(s, b) \sum_{q \in \{q' \in Q:\ q'(s,b)=p\}} \mu(q)}{\sum_{(\tilde{s}, \tilde{b}) \in \Theta_S \times \Theta_B} \Pr(\tilde{s}, \tilde{b}) \sum_{q \in \{q' \in Q:\ q'(\tilde{s}, \tilde{b})=p\}} \mu(q)}.$$

For each recommendation p, the intermediary's interim manipulation is represented by the price $\hat{\alpha} \in \mathbb{R}_+$ he actually offers. For trader $i \in \{S, B\}$ of type $\theta_i \in \Theta_i$, his interim manipulation is represented by a pair $(\hat{\beta}_i, \hat{\gamma}_i)$, where $\hat{\beta}_i \in \Theta_i$ is a manipulation in a report and $\hat{\gamma}_i : R_{\mathrm{marg}} \times \mathbb{R}_+ \to \{Y, N\}$ is a manipulation in responses. Let $\hat{\Sigma}_i$ be the set of all such pairs $(\hat{\beta}_i, \hat{\gamma}_i)$. For any $(q, r) \in Q \times R$ and $(\hat{\beta}_i, \hat{\gamma}_i) \in \hat{\Sigma}_i$, the interim expected payoff $U_i((q, r) \circ (\hat{\beta}_i, \hat{\gamma}_i) \mid \theta_i)$ of trader i of type θ_i when he manipulates (q, r) by $(\hat{\beta}_i, \hat{\gamma}_i)$ is naturally defined. If the intermediary receives the recommendation p and manipulates μ by $\hat{\alpha} \in \mathbb{R}_+$, then he expects to get

$$V(\mu \circ \hat{\alpha} \mid p) = \sum_{(s,b) \in \Theta_S \times \Theta_B} \xi(s, b \mid p) \cdot v(\hat{\alpha} \mid s, b) \cdot \mathbf{1}_{\{r(\hat{\alpha}|s,b)=(Y,Y)\}}.$$

The intermediary has no incentive to manipulate $\mu \in \Delta(Q \times \{r\})$ after he receives a recommendation (at Time 5) if, for all $\hat{\alpha} \in \mathbb{R}_+$ and all $p \in \mathbb{R}_+$ such that $p \in \text{Range}(q)$ for some $q \in \text{supp}(\mu)$,

$$V(\mu \mid p) \geq V(\mu \circ \hat{\alpha} \mid p). \tag{B.6}$$

Trader $i \in \{S, B\}$ of type $\theta_i \in \Theta_i$ has no incentive to manipulate $\mu \in \Delta(Q \times \{r\})$ after he learns his type (at Time 3) if, for all $(\hat{\beta}_i, \hat{\gamma}_i) \in \hat{\Sigma}_i$,

$$\sum_{q \in Q} \mu(q) U_i(q, r \mid \theta_i) \geq \sum_{q \in Q} \mu(q) U_i((q, r) \circ (\hat{\beta}_i, \hat{\gamma}_i) \mid \theta_i). \tag{B.7}$$

Lemma 5. *For any mediation plan $\mu \in \Delta(Q \times \{r\})$ for some arbitrary $r \in R$, the ex-ante IC constraints* (2.1) *and* (2.2) *are satisfied if and only if the interim IC constraints* (B.6) *and* (B.7) *are satisfied.*

Proof. Suppose that (B.6) is not satisfied for $\mu \in \Delta(Q \times \{r\})$; that is, there exists some recommendation $\hat{p} \in \mathbb{R}_+$ and the intermediary's manipulation $\hat{\alpha} \in \mathbb{R}_+$ such that

$$\begin{aligned}
&V(\mu \circ \hat{\alpha} \mid \hat{p}) > V(\mu \mid \hat{p}) \\
\iff &\sum_{(s,b) \in \Theta_S \times \Theta_B} \xi(s, b \mid \hat{p}) \cdot v(\hat{\alpha} \mid s, b) \cdot \mathbf{1}_{\{r(\hat{\alpha}|s,b)=(Y,Y)\}} \\
&> \sum_{(s,b) \in \Theta_S \times \Theta_B} \xi(s, b \mid \hat{p}) \cdot v(\hat{p} \mid s, b) \cdot \mathbf{1}_{\{r(\hat{p}|s,b)=(Y,Y)\}}.
\end{aligned} \tag{B.8}$$

Consider an ex-ante manipulation $\alpha \in \Sigma_I$ of the intermediary such that

$$\alpha(p) = \begin{cases} \hat{\alpha} & \text{if } p = \hat{p} \\ p & \text{if } p \neq \hat{p}. \end{cases}$$

In other words, the intermediary disobeys only When he receives $\hat{p}$. Let $Q_\mu(\hat{p})$ and $Q_\mu(s, b, \hat{p})$ be the set of price-recommendations such that $\mu \in \Delta(Q \times \{r\})$ assigns a positive probability and it recommends $\hat{p} \in \mathbb{R}_+$ for some pairs of types and for $(s,b) \in \Theta_S \times \Theta_B$, respectively; that is,

$$Q_\mu(\hat{p}) = \{q \in Q : \mu(q) > 0 \text{ and } q(s,b) = \hat{p} \text{ for some } (s,b) \in \Theta_S \times \Theta_B\},$$
$$Q_\mu(s,b,\hat{p}) = \{q \in Q : \mu(q) > 0 \text{ and } q(s,b) = \hat{p}\}.$$

For any $q \in Q$, let $\Theta_q(\hat{p})$ be the set of pairs of types for which q recommends $\hat{p}$; that is,

$$\Theta_q(\hat{p}) = \{(s,b) \in \Theta_S \times \Theta_B : q(s,b) = \hat{p}\}.$$

Then, I have $\sum_{q \in Q} \mu(q) V(q \circ \alpha, r) = \sum_{q \in Q_\mu(\hat{p})} \mu(q) V(q \circ \alpha, r) + \sum_{q \notin Q_\mu(\hat{p})} \mu(q) V(q, r)$. I can further decompose the first term in the right side as

$$\begin{aligned}
&\sum_{q \in Q_\mu(\hat{p})} \mu(q) V(q \circ \alpha, r) \\
&= \sum_{q \in Q_\mu(\hat{p})} \mu(q) \Bigg[\sum_{(s,b) \in \Theta_q(\hat{p})} \Pr(s,b) \cdot v(\hat{\alpha} \mid s,b) \cdot \mathbf{1}_{\{r(\hat{\alpha} \mid s,b) = (Y,Y)\}} \\
&\qquad + \sum_{(s,b) \notin \Theta_q(\hat{p})} \Pr(s,b) \cdot v(q(s,b) \mid s,b) \cdot \mathbf{1}_{\{r(q(s,b) \mid s,b) = (Y,Y)\}} \Bigg] \\
&= \sum_{(s,b) \in \cup_{q \in Q_\mu(\hat{p})} \Theta_q(\hat{p})} \Pr(s,b) \Bigg[\sum_{q \in Q_\mu(s,b,\hat{p})} \mu(q) \cdot v(\hat{\alpha} \mid s,b) \cdot \mathbf{1}_{\{r(\hat{\alpha} \mid s,b) = (Y,Y)\}} \\
&\qquad + \sum_{q \in Q_\mu(\hat{p}) \setminus Q_\mu(s,b,\hat{p})} \mu(q) \cdot v(q(s,b) \mid s,b) \cdot \mathbf{1}_{\{r(q(s,b) \mid s,b) = (Y,Y)\}} \Bigg] \\
&\quad + \sum_{(s,b) \notin \cup_{q \in Q_\mu(\hat{p})} \Theta_q(\hat{p})} \Pr(s,b) \sum_{q \in Q_\mu(\hat{p})} \mu(q) \cdot v(q(s,b) \mid s,b) \cdot \mathbf{1}_{\{r(q(s,b) \mid s,b) = (Y,Y)\}}.
\end{aligned} \tag{B.9}$$

Since $\xi(s,b \mid \hat{p})$ is given by

$$\xi(s,b \mid \hat{p}) = \frac{\Pr(s,b) \sum_{q \in Q_\mu(s,b,\hat{p})} \mu(q)}{\sum_{(\tilde{s},\tilde{b}) \in \Theta_S \times \Theta_B} \Pr(\tilde{s},\tilde{b}) \sum_{q \in Q_\mu(\tilde{s},\tilde{b},\hat{p})} \mu(q)},$$

$\xi(s, b \mid \hat{p}) = 0$ for all $(s, b) \notin \cup_{q \in Q_\mu(\hat{p})} \Theta_q(\hat{p})$. Thus, (B.8) implies

$$\begin{aligned}
&\sum_{(s,b) \in \cup_{q \in Q_\mu(\hat{p})} \Theta_q(\hat{p})} \Pr(s, b) \sum_{q \in Q_\mu(s,b,\hat{p})} \mu(q) \cdot v(\hat{\alpha} \mid s, b) \cdot \mathbf{1}_{\{r(\hat{\alpha}|s,b)=(Y,Y)\}} \\
&> \sum_{(s,b) \in \cup_{q \in Q_\mu(\hat{p})} \Theta_q(\hat{p})} \Pr(s, b) \sum_{q \in Q_\mu(s,b,\hat{p})} \mu(q) \cdot v(\hat{p} \mid s, b) \cdot \mathbf{1}_{\{r(\hat{p}|s,b)=(Y,Y)\}}.
\end{aligned}$$

Combined with (B.9), I have

$$\begin{aligned}
&\sum_{q \in Q} \mu(q) V(q \circ \alpha, r) \\
&= \sum_{q \in Q_\mu(\hat{p})} \mu(q) V(q \circ \alpha, r) + \sum_{q \notin Q_\mu(\hat{p})} \mu(q) V(q, r) \\
&= \sum_{(s,b) \in \cup_{q \in Q_\mu(\hat{p})} \Theta_q(\hat{p})} \Pr(s, b) \Bigg[\sum_{q \in Q_\mu(s,b,\hat{p})} \mu(q) \cdot v(\hat{\alpha} \mid s, b) \cdot \mathbf{1}_{\{r(\hat{\alpha}|s,b)=(Y,Y)\}} \\
&\quad + \sum_{q \in Q_\mu(\hat{p}) \setminus Q_\mu(s,b,\hat{p})} \mu(q) \cdot v(q(s, b) \mid s, b) \cdot \mathbf{1}_{\{r(q(s,b)|s,b)=(Y,Y)\}} \Bigg] \\
&\quad + \sum_{(s,b) \notin \cup_{q \in q_\mu(\hat{p})} \Theta_q(\hat{p})} \Pr(s, b) \sum_{q \in Q_\mu(\hat{p})} \mu(q) \cdot v(q(s, b) \mid s, b) \cdot \mathbf{1}_{\{r(q(s,b)|s,b)=(Y,Y)\}} \\
&\quad + \sum_{q \notin Q_\mu(\hat{p})} \mu(q) V(q, r) \\
&> \sum_{(s,b) \in \cup_{q \in Q_\mu(\hat{p})} \Theta_q(\hat{p})} \Pr(s, b) \Bigg[\sum_{q \in Q_\mu(s,b,\hat{p})} \mu(q) \cdot v(\hat{p} \mid s, b) \cdot \mathbf{1}_{\{r(\hat{p}|s,b)=(Y,Y)\}} \\
&\quad + \sum_{q \in Q_\mu(\hat{p}) \setminus Q_\mu(s,b,\hat{p})} \mu(q) \cdot v(q(s, b) \mid s, b) \cdot \mathbf{1}_{\{r(q(s,b)|s,b)=(Y,Y)\}} \Bigg] \\
&\quad + \sum_{(s,b) \notin \cup_{q \in Q_\mu(\hat{p})} \Theta_q(\hat{p})} \Pr(s, b) \sum_{q \in Q_\mu(\hat{p})} \mu(q) \cdot v(q(s, b) \mid s, b) \cdot \mathbf{1}_{\{r(q(s,b)|s,b)=(Y,Y)\}} \\
&\quad + \sum_{q \notin Q_\mu(\hat{p})} \mu(q) V(q, r) \\
&= \sum_{q \in Q_\mu(\hat{p})} \mu(q) V(q, r) + \sum_{q \notin Q_\mu(\hat{p})} \mu(q) V(q, r) \\
&= \sum_{q \in Q} \mu(q) V(q, r).
\end{aligned}$$

Hence, (2.1) does not hold.

Next, suppose that (B.7) is not satisfied for $\mu \in \Delta(Q \times \{r\})$; that is, there exists some trader $i \in \{S, B\}$ of type $\hat{\theta}_i \in \Theta_i$ and his manipulation $(\hat{\beta}_i, \hat{\gamma}_i) \in \hat{\Sigma}_i$ such that

$$\sum_{q \in Q} \mu(q) U_i((q, r) \circ (\hat{\beta}_i, \hat{\gamma}_i) \mid \hat{\theta}_i) > \sum_{q \in Q} \mu(q) U_i(q, r \mid \hat{\theta}_i).$$

Consider an ex-ante manipulation $(\beta_i, \gamma_i) \in \Sigma_i$ of trader i such that

$$\beta_i(\theta_i) = \begin{cases} \hat{\beta}_i & \text{if } \theta_i = \hat{\theta}_i \\ \theta_i & \text{if } \theta_i \neq \hat{\theta}_i \end{cases}$$

and for any $r \in R_{\text{marg}}$ and $p \in \mathbb{R}_+$,

$$\gamma_i(\theta_i, r, p) = \begin{cases} \hat{\gamma}_i(r, p) & \text{if } \theta_i = \hat{\theta}_i \\ r_i(p \mid \theta_i, \theta_j) & \text{if } \theta_i \neq \hat{\theta}_i. \end{cases}$$

In other words, trader i disobeys as prescribed in $(\hat{\beta}_i, \hat{\gamma}_i)$ only when his type is $\hat{\theta}_i$. Then, I have

$$\begin{aligned} & \sum_{q \in Q} \mu(q) U_i((q, r) \circ (\beta_i, \gamma_i)) \\ = \;& \Pr(\hat{\theta}_i) \sum_{q \in Q} \mu(q) U_i((q, r) \circ (\hat{\beta}_i, \hat{\gamma}_i) \mid \hat{\theta}_i) + \left(1 - \Pr(\hat{\theta}_i)\right) \sum_{q \in Q} \mu(q) U_i(q, r \mid \theta_i \neq \hat{\theta}_i) \\ > \;& \Pr(\hat{\theta}_i) \sum_{q \in Q} \mu(q) U_i(q, r \mid \hat{\theta}_i) + \left(1 - \Pr(\hat{\theta}_i)\right) \sum_{q \in Q} \mu(q) U_i(q, r \mid \theta_i \neq \hat{\theta}_i) \\ = \;& \sum_{q \in Q} \mu(q) U_i(q, r). \end{aligned}$$

Hence, (2.2) does not hold. This comples the proof of the only-if-part. □

To prove the if-part, suppose that (2.1) is not satisfied for $\mu \in \Delta(Q \times \{r\})$; that is, there exists some manipulation $\alpha \in \Sigma_I$ by the intermediary such that

$$\sum_{q \in Q} \mu(q) V(q \circ \alpha, r) > \sum_{q \in Q} \mu(q) V(q, r). \tag{B.10}$$

For any $\alpha \in \Sigma_I$, let P_α be the set of prices such that α prescribes disobeying; that is,

$$P_\alpha = \{p \in \mathbb{R}_+ : \alpha(p) \neq p\}.$$

Note that (B.10) can be written as

$$\begin{aligned}
&\sum_{q \in Q} \mu(q) V(q \circ \alpha, r) \\
&= \sum_{q \notin \cup_{p \in P_\alpha} Q_\mu(p)} \mu(q) V(q, r) + \sum_{p \in P_\alpha} \sum_{q \in Q_\mu(p)} \mu(q) V(q \circ \alpha, r) \\
&= \sum_{q \notin \cup_{p \in P_\alpha} Q_\mu(p)} \mu(q) V(q, r) + \sum_{p \in P_\alpha} \sum_{q \in Q_\mu(p)} \mu(q) \sum_{(s,b) \in \Theta_q(p)} \Pr(s, b) \cdot v(\alpha(p) \mid s, b) \cdot \mathbf{1}_{\{r(\alpha(p)|s,b)=(Y,Y)\}} \\
&\quad + \sum_{q \in \cup_{p \in P_\alpha} Q_\mu(p)} \mu(q) \sum_{(s,b) \notin \cup_{p \in P_\alpha} \Theta_q(p)} \Pr(s, b) \cdot v(q(s, b) \mid s, b) \cdot \mathbf{1}_{\{r(q(s,b)|s,b)=(Y,Y)\}} \\
&> \sum_{q \notin \cup_{p \in P_\alpha} Q_\mu(p)} \mu(q) V(q, r) + \sum_{p \in P_\alpha} \sum_{q \in Q_\mu(p)} \mu(q) \sum_{(s,b) \in \Theta_q(p)} \Pr(s, b) \cdot v(p \mid s, b) \cdot \mathbf{1}_{\{r(p|s,b)=(Y,Y)\}} \\
&\quad + \sum_{q \in \cup_{p \in P_\alpha} Q_\mu(p)} \mu(q) \sum_{(s,b) \notin \cup_{p \in P_\alpha} \Theta_q(p)} \Pr(s, b) \cdot v(q(s, b) \mid s, b) \cdot \mathbf{1}_{\{r(q(s,b)|s,b)=(Y,Y)\}} \\
&= \sum_{q \notin \cup_{p \in P_\alpha} Q_\mu(p)} \mu(q) V(q, r) + \sum_{p \in P_\alpha} \sum_{q \in Q_\mu(p)} \mu(q) V(q, r) \\
&= \sum_{q \in Q} \mu(q) V(q, r).
\end{aligned}$$

Then, there exists at least one recommendation $\hat{p} \in P_\alpha$ such that

$$
\begin{aligned}
& \sum_{q \in Q_\mu(\hat{p})} \mu(q) \sum_{(s,b) \in \Theta_q(\hat{p})} \Pr(s,b) \cdot v(\alpha(\hat{p}) \mid s,b) \cdot \mathbf{1}_{\{r(\alpha(\hat{p})|s,b)=(Y,Y)\}} \\
& \quad > \sum_{q \in Q_\mu(\hat{p})} \mu(q) \sum_{(s,b) \in \Theta_q(\hat{p})} \Pr(s,b) \cdot v(\hat{p} \mid s,b) \cdot \mathbf{1}_{\{r(\hat{p}|s,b)=(Y,Y)\}} \\
\iff & \sum_{(s,b) \in \cup_{q \in Q_\mu(\hat{p})} \Theta_q(\hat{p})} \Pr(s,b) \sum_{q \in Q_\mu(s,b,\hat{p})} \mu(q) \cdot v(\alpha(\hat{p}) \mid s,b) \cdot \mathbf{1}_{\{r(\alpha(\hat{p})|s,b)=(Y,Y)\}} \\
& \quad > \sum_{(s,b) \in \cup_{q \in Q_\mu(\hat{p})} \Theta_q(\hat{p})} \Pr(s,b) \sum_{q \in Q_\mu(s,b,\hat{p})} \mu(q) \cdot v(\hat{p} \mid s,b) \cdot \mathbf{1}_{\{r(\hat{p}|s,b)=(Y,Y)\}} \\
\iff & \sum_{(s,b) \in \cup_{q \in Q_\mu(\hat{p})} \Theta_q(\hat{p})} \underbrace{\frac{\Pr(s,b) \sum_{q \in Q_\mu(s,b,\hat{p})} \mu(q)}{\sum_{(\tilde{s},\tilde{b}) \in \Theta_S \times \Theta_B} \Pr(\tilde{s},\tilde{b}) \sum_{q \in Q_\mu(\tilde{s},\tilde{b},\hat{p})} \mu(q)}}_{=\xi(s,b|\hat{p})} \cdot v(\alpha(\hat{p}) \mid s,b) \cdot \mathbf{1}_{\{r(\alpha(\hat{p})|s,b)=(Y,Y)\}} \\
& \quad > \sum_{(s,b) \in \cup_{q \in Q_\mu(\hat{p})} \Theta_q(\hat{p})} \underbrace{\frac{\Pr(s,b) \sum_{q \in Q_\mu(s,b,\hat{p})} \mu(q)}{\sum_{(\tilde{s},\tilde{b}) \in \Theta_S \times \Theta_B} \Pr(\tilde{s},\tilde{b}) \sum_{q \in Q_\mu(\tilde{s},\tilde{b},\hat{p})} \mu(q)}}_{=\xi(s,b|\hat{p})} \cdot v(\hat{p} \mid s,b) \cdot \mathbf{1}_{\{r(\hat{p}|s,b)=(Y,Y)\}} \\
\iff & \sum_{(s,b) \in \Theta_S \times \Theta_B} \xi(s,b \mid \hat{p}) \cdot v(\alpha(\hat{p}) \mid s,b) \cdot \mathbf{1}_{\{r(\alpha(\hat{p})|s,b)=(Y,Y)\}} \\
& \quad > \sum_{(s,b) \in \Theta_S \times \Theta_B} \xi(s,b \mid \hat{p}) \cdot v(\hat{p} \mid s,b) \cdot \mathbf{1}_{\{r(\hat{p}|s,b)=(Y,Y)\}} \\
\iff & V(\mu \circ \alpha(\hat{p}) \mid \hat{p}) > V(\mu \mid \hat{p}).
\end{aligned}
$$

Hence, (B.6) does not hold.

Next, suppose that (2.2) is not satisfied for $\mu \in \Delta(Q \times \{r\})$; that is, there exists some trader $i \in \{S, B\}$ and his manipulation $(\beta_i, \gamma_i) \in \Sigma_i$ such that

$$
\sum_{q \in Q} \mu(q) U_i((q,r) \circ (\beta_i, \gamma_i)) > \sum_{q \in Q} \mu(q) U_i(q,r).
$$

Note that, as I have seen in the proof of the only-if-part, trader i's ex-ante expected payoff is a weighted average of his interim expected payoffs. Thus, there exists at least one type $\hat{\theta}_i \in \Theta_i$ who gets a strictly higher interm expected payoff from this manipulation. Let $(\hat{\beta}_i, \hat{\gamma}_i) \in \hat{\Sigma}_i$ denote the interim manipulation by $\hat{\theta}_i$ obtained from (β_i, γ_i). I must have

$$\sum_{q\in Q} \mu(q) U_i((q,r) \circ (\hat{\beta}_i, \hat{\gamma}_i) \mid \hat{\theta}_i) > \sum_{q\in Q} \mu(q) U_i(q, r \mid \hat{\theta}_i).$$

Hence, (B.7) does not hold. This completes the proof.

C The role of the restriction on response-recommendation

In this appendix, I lift the restriction on response-recommendation and provide the necessary and sufficient condition for the existence of CE that implements ex-post efficient outcome.

In what follows, I restrict my attention to the set of pure mediation plans $Q \times R$. For any price-recommendation $q \in Q$, let $r_q \in R$ be such that, for all $(s, b) \in \Theta_S \times \Theta_B \setminus \{(s_H, b_L)\}$,

$$r_q(p \mid s, b) = \begin{cases} (Y, Y) & \text{if } p = q(s, b) \\ (N, N) & \text{otherwise} \end{cases}$$

and

$$r_q(p \mid s_H, b_L) = (N, N) \text{ for all } p \in \mathbb{R}_+.$$

In other words, if there is gain from trade, r_q recommends accepting only the on-path price $q(s, b)$. If there is no gain from trade, it always recommends rejection. Lemma 6 claims that for any price-recommendation q, it suffices to consider (q, r_q) to see if q can be a part of a CE that implements ex-post efficient outcome.

Lemma 6. *For any price-recommendation $q \in Q$, consider a mediation plan $(q, r) \in Q \times R$ that induces an ex-post efficient outcome. If (q, r) is a CE, so is (q, r_q).*

Proof. First, I prove the following lemma, which provides the necessary conditions for CE that implements ex-post efficient outcome.

Lemma 7. *A mediation plan $(q, r) \in Q \times R$ that induces ex-post efficient outcome is a CE only if*

(i) $q(s, b) \in [s, b]$ for all $(s, b) \neq (s_H, b_L)$;

(ii) $r(p \mid s, b) \neq (Y, Y)$ *for all* $p > q(s, b)$ *and all* $(s, b) \neq (s_H, b_L)$*;*

(iii)

$$r(q(s_H, b_L) \mid s_H, b_L) = \begin{cases} (N,Y) \text{ or } (N,N) & \text{if } q(s_H,b_L) \in [0, b_L) \\ (Y,N), (N,Y), \text{ or } (N,N) & \text{if } q(s_H,b_L) \in [b_L, s_H] \\ (Y,N) \text{ or } (N,N) & \text{if } q(s_H,b_L) \in (s_H, +\infty), \end{cases}$$

(iv) $r(p \mid s_H, b_L) \neq (Y, Y)$ *for all* $p > \frac{\lambda s_H - (1-\lambda) b_L}{2\lambda - 1}$.

Proof. I show the contraposition. Consider an arbitrary mediation plan (q, r) that induces ex-post efficient outcome. First, suppose that there exists some $(s, b) \neq (s_H, b_L)$ such that $q(s, b) \notin [s, b]$. Since (q, r) induces an ex-post efficient outcome, it recommends $r(q(s, b) \mid s, b) = (Y, Y)$. However, either the seller or the buyer will get a negative payoff if he accepts $q(s, b)$. Hence, this player has the incentive to disobey and reject $q(s, b)$.

Second, suppose that there exists some $(s, b) \neq (s_H, b_L)$ and $p > q(s, b)$ such that $r(p \mid s, b) = (Y, Y)$. If the intermediary is obedient and offers $q(s, b)$, then it will be accepted and he will get $v(q(s, b) \mid s, b)$. Since $v(\cdot \mid s, b)$ is strictly increasing, he has the incentive to disobey and offer p, which gives him $v(p \mid s, b) > v(q(s, b) \mid s, b)$.

Third, suppose that (iii) does not hold. Note that $r(q(s_H, b_L) \mid s_H, b_L) \neq (Y, Y)$ since (q, r) induces an ex-post efficient outcome. If $q(s_H, b_L) \in [0, b_L)$ and $r(q(s_H, b_L) \mid s_H, b_L) = (Y, N)$, then the low-type buyer b_L has the incentive to disobey and accept $q(s_H, b_L)$, which gives him $b_L - q(s_H, b_L) > 0$. If $q(s_H, b_L) \in (s_H, +\infty)$ and $r(q(s_H, b_L) \mid s_H, b_L) = (N, Y)$, then the high-type seller s_H has the incentive to disobey and accept $q(s_H, b_L)$, which gives him $q(s_H, b_L) - s_H > 0$.

Finally, suppose that there exists some $p > \frac{\lambda s_H - (1-\lambda) b_L}{2\lambda - 1}$ such that $r(p \mid s_H, b_L) = (Y, Y)$. If the intermediary is obedient and offers $q(s_H, b_L)$, then it will be rejected and he will get 0. Since $v(p \mid s, b) > 0 \Leftrightarrow p > \frac{\lambda s_H - (1-\lambda) b_L}{2\lambda - 1}$, he has the incentive to disobey and offer p.

Therefore, if (q, r) fails to satisfy any of (i)–(iv), then some player has the incentive to disobey, which implies that (q, r) is not a CE. □

Note that if q satisfies condition (i) in Lemma 7, then no manipulation by the intermediary is profitable in (q, r_q); for $(s, b) \neq (s_H, b_L)$, any off-path offer $p \neq q(s, b)$ will be rejected and give him the payoff 0, while the on-path offer $q(s, b)$ will be accepted and give him a positive payoff $v(q(s, b) \mid s, b)$. For (s_H, b_L), no offer will be accepted and he will always get 0.

Let $r \in R$ be an arbitrary response-recommendation such that (q, r) induces an ex-post efficient outcome and satisfies the necessary conditions in Lemma 7. Since (q, r) induces an ex-post efficient outcome, r must also satisfy $r(q(s, b) \mid s, b) = (Y, Y)$ for all $(s, b) \neq (s_H, b_L)$ and $r(q(s_H, b_L) \mid s_H, b_L) \neq (Y, Y)$. Hence, r can be different from r_q in; first, $r(p \mid s, b) \neq (N, N)$ for some $(s, b) \neq (s_H, b_L)$ and some $p \neq q(s, b)$; second, $r(q(s_H, b_L) \mid s_H, b_L) \neq (N, N)$; and third, $r(p \mid s_H, b_L) \neq (N, N)$ for some $p \neq q(s_H, b_L)$.

Case 1: Suppose that there exists some $(s, b) \neq (s_H, b_L)$ and $p \neq q(s, b)$ such that $r(p \mid s, b) \neq (N, N)$. The traders' IC constraints in (q, r) are the same as those in (q, r_q) because the off-path price p never appears in the traders' constraints. Note that condition (ii) in Lemma 7 ensures that p can be accepted only when $p \leq q(s, b)$. Hence, the intermediary can get at most $v(p \mid s, b) \leq v(q(s, b) \mid s, b)$ by manipulation, which implies that no manipulation is profitable as in (q, r_q).

Case 2: Suppose $r(q(s_H, b_L) \mid s_H, b_L) \neq (N, N)$. Note that the IC constraints for the high-type seller s_H and the low-type buyer b_L in (q, r) are the same as those in (q, r_q). This is because condition (iii) in Lemma 7 ensures that $r(q(s_H, b_L) \mid s_H, b_L) \neq (N, N)$ does not give rise to additional manipulations by them that can be profitable. Note also that no manipulation by the intermediary is profitable as in (q, r_q).

However, the IC constraints for the low-type seller s_L and the high-type buyer b_H can be more stringent than those in (q, r_q). If $q(s_H, b_L) \in (s_L, b_L)$ and $r(q(s_H, b_L) \mid s_H, b_L) = (N, Y)$, then s_L can have the incentive to misreport his type and accept $q(s_H, b_L)$, which gives him $q(s_H, b_L) - s_L > 0$. If $q(s_H, b_L) \in (s_H, b_H)$ and $r(q(s_H, b_L) \mid s_H, b_L) = (Y, N)$, then b_H can have the incentive to misreport his type and accept $q(s_H, b_L)$, which gives him $b_H - q(s_H, b_L) > 0$. Since there are now additional manipulations that must be deterred, the IC constraints for them in (q, r) are at least as stringent as those in (q, r_q).

Case 3: Suppose that there exists some $p \neq q(s_H, b_L)$ such that $r(p \mid s_H, b_L) \neq (N, N)$. The traders' IC constraints in (q, r) are the same as those in (q, r_q) because the off-path price p never appears in the traders' constraints. Note that condition (iv) in Lemma 7 ensures that p can be accepted only when $p \leq \frac{\lambda s_H - (1-\lambda) b_L}{2\lambda - 1} \Leftrightarrow v(p \mid s, b) \leq 0$. Hence, the intermediary can get at most $v(p \mid s, b) \leq 0$ by manipulation, which implies that no manipulation is profitable as in (q, r_q).

In conclusion, the players' IC constraints in (q, r) are at least as stringent as those in (q, r_q). Therefore, if (q, r) is a CE, so is (q, r_q). □

Proposition 3 provides a necessary and sufficient condition for the existence of pure CE that implements ex-post efficient outcome.

Proposition 3. *If the intermediary is seller-biased ($\lambda > 1/2$), then a pure CE that implements ex-post efficient outcome exists if and only if*

$$\pi_S \pi_B b_H + (1 - \pi_S) b_L \geq \pi_B s_H + (1 - \pi_S)(1 - \pi_B) s_L. \tag{C.1}$$

Proof. As I will discuss in Section **??**, the condition (C.1) is the same as that in the case of an unbiased intermediary with commitment and enforcement power. Thus, the only-if part of the proof is trivial; if ex-post efficiency cannot be hoped even with commitment and enforcement power, it cannot be hoped either without them.

In what follows, I show that a such CE exists if and only if there exists some $p \in (b_L, s_H)$ such that $\pi_S \geq \frac{p - b_L}{b_H - s_H + p - b_L}$ and $\pi_B \leq \frac{b_L - s_L}{s_H - p + b_L - s_L}$. By solving these inequalities with respect to p, I obtain

$$p \in \left[s_H - \frac{1 - \pi_B}{\pi_B}(b_L - s_L),\ b_L + \frac{\pi_S}{1 - \pi_S}(b_H - s_H) \right].$$

Since $s_H - \frac{1-\pi_B}{\pi_B}(b_L - s_L) < s_H$ and $b_L < b_L + \frac{\pi_S}{1-\pi_S}(b_H - s_H)$, such $p \in (b_L, s_H)$ exists if and only if

$$s_H - \frac{1 - \pi_B}{\pi_B}(b_L - s_L) \leq b_L + \frac{\pi_S}{1 - \pi_S}(b_H - s_H)$$
$$\Longleftrightarrow \pi_S \pi_B b_H + (1 - \pi_S) b_L \geq \pi_B s_H + (1 - \pi_S)(1 - \pi_B) s_L.$$

Table C.1: The price-recommendation q.

	b_H	b_L
s_H	s_H	$q(s_H, b_L) \in (b_H, +\infty)$
s_L	$p \in (b_L, s_H)$	b_L

Therefore, the condition is equivalent to (C.1).

Proof of the if-part: Let $q \in Q$ be such that $q(s_H, b_H) = s_H$, $q(s_{H,} b_L) \in (b_H, +\infty)$, $q(s_L, b_H) = p$ for some $p \in (b_L, s_H)$, and $q(s_{L,} b_L) = b_L$ (see Table C.1).

I show that a mediation plan (q, r_q) that induces an ex-post efficient outcome is a CE if and only if $\pi_S \geq \frac{p-b_L}{b_H-s_H+p-b_L}$ and $\pi_B \leq \frac{b_L-s_L}{s_H-p+b_L-s_L}$, where $p \in (b_L, s_H)$. It is easy to see that no manipulation by the high-type seller s_H and the low-type buyer b_L is profitable. If the low-type seller s_L is honest and obedient, then he will get $\pi_B(p - s_L) + (1 - \pi_B)(b_L - s_L)$. If he misreports his type, then he can get at most $\pi_B(s_H - s_L)$. Hence, he has no profitable manipulation if and only if

$$\pi_B(p - s_L) + (1 - \pi_B)(b_L - s_L) \geq \pi_B(s_H - s_L) \iff \pi_B \leq \frac{b_L - s_L}{s_H - p + b_L - s_L}.$$

Similarly, if the high-type buyer b_H is honest and obedient, he will get $\pi_S(b_H - s_H) + (1 - \pi_S)(b_H - p)$. If he misreports his type, then he can get at most $(1 - \pi_S)(b_H - b_L)$. Hence, he has no profitable manipulation if and only if

$$\pi_S(b_H - s_H) + (1 - \pi_S)(b_H - p) \geq (1 - \pi_S)(b_H - b_L) \iff \pi_S \geq \frac{p - b_L}{b_H - s_H + p - b_L}.$$

By the same argument as in the proof of Lemma 6, the intermediary has no profitable manipulation in (q, r_q), as $q(s, b) \in [s, b]$ for all $(s, b) \neq (s_H, b_L)$. Thus, (q, r_q) is a CE if and only if $\pi_S \geq \frac{p-b_L}{b_H-s_H+p-b_L}$ and $\pi_B \leq \frac{b_L-s_L}{s_H-p+b_L-s_L}$, where $p \in (b_L, s_H)$. □

D Alternative Assumptions on the Intermediary

D.1 Unbiased intermediary

D.1.1 Without commitment and enforcement power

If $\lambda = 1/2$, the intermediary's payoff no longer depends on price; he gets $(b - s)/2$ if trade occurs between (s, b) and 0 if no trade occurs. This gives the intermediary much less incentive to disobey, as he is now indifferent between prices that give the same expected trade probability. This effect is so strong that the refinement does not work at all; a pure acceptable CE that induces ex-post efficient outcome exists under the same condition as in Proposition 3. Proposition 4 summarizes this observation.

Proposition 4. *If the intermediary is unbiased ($\lambda = 1/2$), then a pure acceptable CE that induces ex-post efficient outcome exists if and only if*

$$\pi_S \pi_B b_H + (1 - \pi_S) b_L \geq \pi_B s_H + (1 - \pi_S)(1 - \pi_B) s_L.$$

Proof. As in the proof of Proposition 3, I show that such pure acceptable CE exists if and only if there exists some $p \in (b_L, s_H)$ such that $\pi_S \geq \frac{p - b_L}{b_H - s_H + p - b_L}$ and $\pi_B \leq \frac{b_L - s_L}{s_H - p + b_L - s_L}$, which is equivalent to $\pi_S \pi_B b_H + (1 - \pi_S) b_L \geq \pi_B s_H + (1 - \pi_S)(1 - \pi_B) s_L$.

Let $q \in Q$ be such that $q(s_H, b_H) = s_H$, $q(s_H, b_L) \in (b_H, +\infty)$, $q(s_L, b_H) = p$ for some $p \in (b_L, s_H)$, and $q(s_L, b_L) = b_L$. I show that the mediation plan (q, r^A) that induces ex-post efficient outcome is an acceptable CE if and only if $\pi_S \geq \frac{p - b_L}{b_H - s_H + p - b_L}$ and $\pi_B \leq \frac{b_L - s_L}{s_H - p + b_L - s_L}$, where $p \in (b_L, s_H)$. As shown in the proof of Proposition 3, the traders' interim IC constraints are satisfied if and only if $\pi_S \geq \frac{p - b_L}{b_H - s_H + p - b_L}$ and $\pi_B \leq \frac{b_L - s_L}{s_H - p + b_L - s_L}$.[26]

Next, consider the intermediary's incentive for manipulation. When he receives $q(s, b) \neq q(s_H, b_L)$, his expected payoff is $(b - s)/2 > 0$ if he offers $p \in [s, b]$ and 0 otherwise. Hence, following the recommendation is optimal for him. When he receives $q(s_H, b_L)$, he believes that the traders are (s_H, b_L), who accept no offer. Hence, following the recommendation is optimal for him.

[26]Note that I have the same interim IC constaints in (q, r_q) and (q, r^A).

Thus, (q, r^A) is an acceptable CE if and only if $\pi_S \geq \frac{p-b_L}{b_H-s_H+p-b_L}$ and $\pi_B \leq \frac{b_L-s_L}{s_H-p+b_L-s_L}$. This completes the proof of the if-part.

To prove the only-if-part, suppose $\pi_S < \frac{p-b_L}{b_H-s_H+p-b_L}$ or $\pi_B > \frac{b_L-s_L}{s_H-p+b_L-s_L}$ for all $p \in (b_L, s_H)$. By Lemma **??**, this implies $\pi_S < \frac{s_H-b_L}{b_H-b_L}$ and $\pi_B > \frac{b_L-s_L}{s_H-s_L}$. I show that no mediation plan that induces an ex-post efficient outcome is an acceptable CE. Note that a price-recommendation q of such mediation plan (q, r^A) must fall into one of the following three cases.

Case 1: Recommend two different prices for $(s, b) \neq (s_H, b_L)$; that is, $p_H = q(s_H, b_H) = q(s_L, b_H)$ and $p_L = q(s_L, b_L)$;

Case 2: Recommend two different prices for $(s, b) \neq (s_H, b_L)$; that is, $p_H = q(s_H, b_H)$ and $p_L = q(s_L, b_H) = q(s_L, b_L)$;

Case 3: Recommend three different prices for $(s, b) \neq (s_H, b_L)$; that is, $p_{HH} = q(s_H, b_H)$, $p_{LH} = q(s_L, b_H)$, and $p_{LL} = q(s_L, b_L)$.

As I focus on pure acceptable CE that induces an ex-post efficient outcome, I restrict my attention to (q, r^A) in each case.

Case 1: Let $q \in Q$ be such that $q(s_H, b_H) = q(s_L, b_H) = p_H$ for some $p_H \in [s_H, b_H]$, $q(s_H, b_L) \in \mathbb{R}_+$, and $q(s_L, b_L) = p_L$ for some $p_L \in [s_L, b_L]$. As shown in the proof of Proposition 3, the traders' interim IC constraints are satisfied if and only if $\pi_S \geq \frac{p_H-p_L}{b_H-p_L}$.

Next, consider the intermediary's incentive for manipulation. When he receives p_H, if he offers p, then his expected payoff is

$$\begin{cases} \frac{\pi_S\pi_B}{\pi_S\pi_B+(1-\pi_S)\pi_B} \cdot \frac{b_H-s_H}{2} + \frac{(1-\pi_S)\pi_B}{\pi_S\pi_B+(1-\pi_S)\pi_B} \cdot \frac{b_H-s_L}{2} & \text{if } p \in [s_H, b_H] \\ \frac{(1-\pi_S)\pi_B}{\pi_S\pi_B+(1-\pi_S)\pi_B} \cdot \frac{b_H-s_L}{2} & \text{if } p \in [s_L, s_H) \\ 0 & \text{otherwise.} \end{cases}$$

Hence, following the recommendation is optimal for him. When he receives p_L, his expected payoff is $(b_L - s_L)/2$ if he offers $p \in [s_L, b_L]$ and 0 otherwise. Hence, following the recommendation is optimal for him. When he receives $q(s_H, b_L)$, he believes that the traders are (s_H, b_L), who accept no offer. Hence, following the recommendation is optimal for him.

Thus, (q, r^A) is an acceptable CE if and only if $\pi_S \geq \frac{p_H - p_L}{b_H - p_L}$, which is least stringent at $(p_H, p_L) = (s_H, b_L)$ and reduces to $\pi_S \geq \frac{s_H - b_L}{b_H - b_L}$. Since $\pi_S < \frac{s_H - b_L}{b_H - b_L}$, the mediation plan (q, r^A) cannot be an acceptable CE.

Case 2: Let $q \in Q$ be such that $q(s_H, b_H) = p_H$ for some $p_H \in [s_H, b_H]$, $q(s_H, b_L) \in \mathbb{R}_+$, and $q(s_L, b_H) = q(s_L, b_L) = p_L$ for some $p_L \in [s_L, b_L]$. As shown in the proof of Proposition 3, the traders' interim IC constraints are satisfied if and only if $\pi_B \leq \frac{p_L - s_L}{p_H - s_L}$.

Next, consider the intermediary's incentive for manipulation. When he receives p_H, his expected payoff is $(b_H - s_H)/2$ if he offers $p \in [s_H, b_H]$ and 0 otherwise. Hence, following the recommendation is optimal for him. When he receives p_L, if he offers p, then his expected payoff is

$$\begin{cases} \frac{(1-\pi_S)\pi_B}{(1-\pi_S)\pi_B + (1-\pi_S)(1-\pi_B)} \cdot \frac{b_H - s_L}{2} & \text{if } p \in (b_L, b_H] \\ \frac{(1-\pi_S)\pi_B}{(1-\pi_S)\pi_B + (1-\pi_S)(1-\pi_B)} \cdot \frac{b_H - s_L}{2} + \frac{(1-\pi_S)(1-\pi_B)}{(1-\pi_S)\pi_B + (1-\pi_S)(1-\pi_B)} \cdot \frac{b_L - s_L}{2} & \text{if } p \in [s_L, b_L] \\ 0 & \text{otherwise.} \end{cases}$$

Hence, following the recommendation is optimal for him. When he receives $q(s_H, b_L)$, by the same argument as in Case 1, following the recommendation is optimal for him.

Thus, (q, r^A) is an acceptable CE if and only if $\pi_B \leq \frac{p_L - s_L}{p_H - s_L}$, which is least stringent at $(p_H, p_L) = (s_H, b_L)$ and reduces to $\pi_B \leq \frac{b_L - s_L}{s_H - s_L}$. Since $\pi_B > \frac{b_L - s_L}{s_H - s_L}$, the mediation plan (q, r^A) cannot be an acceptable CE.

Case 3: Let $q \in Q$ be such that $q(s_H, b_H) = p_{HH}$ for some $p_{HH} \in [s_H, b_H]$, $q(s_H, b_L) \in \mathbb{R}_+$, $q(s_L, b_H) = p_{LH}$ for some $p_{LH} \in [s_L, b_H]$, and $q(s_L, b_L) = p_{LL}$ for some $p_{LL} \in [s_L, b_L]$, where $p_{HH} \neq p_{LH} \neq p_{LL}$. As shown in the proof of Proposition 3, the traders' interim IC constraints are satisfied if and only if $p_{HH} > p_{LH} > p_{LL}$, $\pi_S \geq \frac{p_{LH} - p_{LL}}{b_H - p_{HH} + p_{LH} - p_{LL}}$, and $\pi_B \leq \frac{p_{LL} - s_L}{p_{HH} - p_{LH} + p_{LL} - s_L}$.

Next, consider the intermediary's incentive for manipulation. When he receives $q(s, b) \neq q(s_H, b_L)$, his expected payoff is $(b - s)/2 > 0$ if he offers $p \in [s, b]$ and 0 otherwise. Hence, following the recommendation is optimal for him. When he receives $q(s_H, b_L)$, by the same argument as in Case 1 and 2, following the recommendation is optimal for him. Thus, (q, r^A) is an acceptable CE if and only if $p_{HH} > p_{LH} > p_{LL}$, $\pi_S \geq \frac{p_{LH}-p_{LL}}{b_H-p_{HH}+p_{LH}-p_{LL}}$, and $\pi_B \leq \frac{p_{LL}-s_L}{p_{HH}-p_{LH}+p_{LL}-s_L}$, which are least stringent at $(p_{HH}, p_{LL}) = (s_H, b_L)$ and reduce to $\pi_S \geq \frac{p_{LH}-b_L}{b_H-s_H+p_{LH}-b_L}$ and $\pi_B \leq \frac{b_L-s_L}{s_H-p_{LH}+b_L-s_L}$ for some $p_{LH} \in (b_L, s_H)$. Since $\pi_S < \frac{p-b_L}{b_H-s_H+p-b_L}$ or $\pi_B > \frac{b_L-s_L}{s_H-p+b_L-s_L}$ for all $p \in (b_L, s_H)$, the mediation plan (q, r^A) cannot be an acceptable CE.

In conclusion, if $\pi_S < \frac{p-b_L}{b_H-s_H+p-b_L}$ or $\pi_B > \frac{b_L-s_L}{s_H-p+b_L-s_L}$ for all $p \in (b_L, s_H)$, then, for each possible q above, a mediation plan (q, r^A) that induces an ex-post efficient outcome is not an acceptable CE. □

When I construct an acceptable CE, I can deter the intermediary from disobeying by letting the traders reject all off-path prices (namely, use r_q for each price-recommendation q). Intuitively, being unbiased has the same effect on the intermediary's incentive as the commitment. Although the traders (s, b) will accept any price $p \in [s, b]$, the intermediary is indifferent between these prices.

D.1.2 With commitment and enforcement power

Matsuo (1989) studies the mechanism design problem for bilateral trade à la Myerson and Satterthwaite (1983) while putting the same assumptions on the traders' valuations as in this paper. Namely, the traders' valuations are binary and independently distributed and satisfy $s_L < b_L < s_H < b_H$. The principal in his model can be viewed as an unbiased intermediary who can commit to and enforce the prices he offers. He shows that a Bayesian incentive compatible, individually rational, and ex-post efficient mechanism exists if and only if $\pi_S\pi_B b_H + (1-\pi_S)b_L \geq \pi_B s_H + (1-\pi_S)(1-\pi_B)s_L$, the same condition as in Proposition 3 and 4. This implies that the lack of commitment and enforcement power has no bite on efficiency if the intermediary is unbiased.

D.2 Without intermediary

So far, I have studied the mediated bargaining game in which the intermediary makes an offer instead of the traders. To clarify the effect of intermediary, I consider another mediated bargaining game in which the seller makes an offer. Let $\hat{r}\colon \Theta_S \times \Theta_B \times \mathbb{R}_+ \longrightarrow \{Y, N\}$ be a response-recommendation to the buyer and $\hat{R}$ be the set of all such recommendations. A mediation plan is a probability distribution over $Q \times \hat{R}$. I consider the mediated seller-offer bargaining game that proceeds as follows:

1. The mediator publicly commits to a mediation plan $\mu \in \Delta(Q \times \hat{R})$ that she will use;
2. The seller and the buyer privately observe the realizations of their types;
3. The traders confidentially report their types $s \in \Theta_S$ and $b \in \Theta_B$ to the mediator;
4. The mediator privately picks up a pure mediation plan $(q, \hat{r}) \in Q \times \hat{R}$ with probability $\mu(q, \hat{r})$. She confidentially recommends a price $q(s, b) \in \mathbb{R}_+$ to the seller and a response-rule $\hat{r}(\cdot \mid s, b)$ to the buyer;
5. The seller offers a price $p \in \mathbb{R}_+$;
6. The buyer responds to p, by either acceptance (Y) or rejection (N). If he accepts, trade occurs at the price p, and the payoffs are realized; otherwise, no trade occurs, and all players get 0.

It is straightforward to show that no acition is codominated in this game. Thus, any CE is an acceptable CE as in the mediated bargaining game. Let $\hat{r}^A \in \hat{R}$ be such that for all $(s, b) \in \Theta_S \times \Theta_B$,

$$\hat{r}^A(p \mid s, b) = \begin{cases} Y & \text{if } p \in [0, b] \\ N & \text{if } p \in (b, +\infty) \end{cases}.$$

A CE $\mu \in \Delta\left(Q \times \{\hat{r}^A\}\right)$ is *acceptable* if it $\mu(Q \times \{\hat{r}^A\}) = 1$. With this definition of acceptable CE, I can show analogs of Lemma 3 and 2.

Lemma 8. *A mediation plan $\mu \in \Delta\left(Q \times \{\hat{r}^A\}\right)$ that induces an ex-post efficient outcome is an acceptable CE only if, for all $q \in \mathrm{supp}(\mu)$, $q(s_H, b_H) = b_H$, $q(s_L, b_H) = q(s_L, b_L) = b_L$, and $q(s_H, b_L) \notin [0, b_H)$.*

Proof. Consider an arbitrary mediation plan $\mu \in \Delta\left(Q \times \{\hat{r}^A\}\right)$. Suppose that there exist some $q \in \operatorname{supp}(\mu)$ such that $q(s_H, b_H) \neq b_H$. If the high-type seller s_H is honest to the mediator and recommended $q(s_H, b_H)$, then his expected payoff from offering p is

$$\begin{cases} \xi(b_H \mid q(s_H, b_H))(p - s_H) & \text{if } p \in (b_L, b_H] \\ p - s_H & \text{if } p \in [0, b_L] \\ 0 & \text{otherwise} \end{cases},$$

where $\xi(b \mid p)$ is his posterior belief that the buyer is of type $b \in \Theta_B$ When he receives $p \in \mathbb{R}_+$. Since $p - s_H < 0$ for all $p \in [0, s_H)$ and $\xi(b_H \mid q(s_H, b_H)) > 0$, he has the incentive to disobey to offer b_H. Next, suppose that there exist some $q \in \operatorname{supp}(\mu)$ and $b \in \Theta_B$ such that $q(s_L, b) \notin \{b_H, b_L\}$. If the seller of type s_L is honest to the mediator and recommended $q(s_L, b)$, then his expected payoff from offering p is

$$\begin{cases} \xi(b_H \mid q(s_L, b))(p - s_L) & \text{if } p \in (b_L, b_H] \\ p - s_L & \text{if } p \in [0, b_L] \\ 0 & \text{otherwise} \end{cases}.$$

Thus, he has the incentive to disobey to offer either b_H or b_L. This shows that $\mu \in \Delta\left(Q \times \{\hat{r}^A\}\right)$ is an acceptable CE only if, for all $q \in \operatorname{supp}(\mu)$, $q(s_H, b_H) = b_H$ and $q(s_L, b) \in \{b_H, b_L\}$ for all $b \in \Theta_B$. In addition, any mediation plan $\mu \in \Delta\left(Q \times \{\hat{r}^A\}\right)$ that induces an ex-post efficient outcome must also satisfy $q(s, b) \in [s, b]$ for all $(s, b) \neq (s_H, b_L)$ and all $q \in \operatorname{supp}(\mu)$, which implies $q(s_L, b_L) = b_L$.

If the high-type buyer b_H is honest and obedient, then he can get $b_H - b_L$ when the seller is of low-type and the price b_L is offered. If he misreports his type, then he can get $b_H - b_L$ when the seller is of low-type, and $b_H - p > 0$ when the seller is of high-type and the price $p \in [0, b_H)$

Table D.1: The price-recommendation $\hat{q}^*$.

	b_H	b_L
s_H	b_H	b_H
s_L	b_L	b_L

is offered. Thus, he has no profitable manipulation if

$$\begin{aligned}
&\sum_{q\in\{q'\in Q:\, q'(s_L,b_H)=b_L\}} \mu(q)(1-\pi_S)(b_H-b_L) \\
&\quad\geq \sum_{p\in[0,b_H)} \sum_{q\in\{q'\in Q:\, q'(s_H,b_L)=p\}} \mu(q)\pi_S(b_H-p)+(1-\pi_S)(b_H-b_L) \\
\iff &\sum_{p\in[0,b_H)} \sum_{q\in\{q'\in Q:\, q'(s_H,b_L)=p\}} \mu(q)\pi_S(b_H-p) \\
&\quad+\left(1-\sum_{q\in\{q'\in Q:\, q'(s_L,b_H)=b_L\}} \mu(q)\right)(1-\pi_S)(b_H-b_L)\leq 0.
\end{aligned}$$

This implies that μ is not an acceptable CE if there exists some $q \in \operatorname{supp}(\mu)$ such that $q(s_L, b_H) = b_H$ or $q(s_H, b_L) \in [0, b_H)$. This completes the proof. □

Lemma 9. *There exists a price-recommendation $\hat{q}^* \in Q$ such that if a mediation plan $\mu \in \Delta(Q \times \hat{r}^A)$ that induces an ex-post efficient outcome is an acceptable CE, so is $(\hat{q}^*, \hat{r}^A)$.*

Proof. Let $\hat{q}^* \in Q$ be such that $\hat{q}^*(s_H, b_H) = \hat{q}^*(s_H, b_L) = b_H$ and $\hat{q}^*(s_L, b_H) = \hat{q}^*(s_L, b_L) = b_L$ (see Table D.1). In other words, the price-recommendation only depends on the seller's report. As such, no manipulation by the buyer is profitable. Note that the seller's posterior belief about the buyer's type is the same as the prior regardless of the recommendation he receives. Thus, if the seller of type s offers p, then his expected payoff is

$$\begin{cases} \pi_B(p-s) & \text{if } p \in (b_L, b_H] \\ p-s & \text{if } p \in [0, b_L] \\ 0 & \text{otherwise.} \end{cases}$$

Since $p - s_H < 0$ for all $p \in [0, b_L]$, the high-type seller s_H has no profitable manipulation. The low-type seller s_L does not also have a profitable manipulation if

$$b_L - s_L \geq \pi_B(b_H - s_L) \iff \pi_B \leq \frac{b_L - s_L}{b_H - s_L}. \tag{D.1}$$

Consider an arbitrary mediation plan $\mu \in \Delta(Q \times \hat{r}^A)$ that induces an ex-post efficient outcome and satisfies the necessary conditions in Lemma 8 (otherwise, μ is not an acceptable CE and hence the statement of the lemma holds). Suppose that there exists some $q \in \text{supp}(\mu)$ such that $q(s_H, b_L) \neq b_H$. Then, the seller learns the buyer's type When he receives $q(s_H, b_L)$. Thus, if the low-type seller s_L misreports his type and receives $q(s_H, b_L)$, then he will offer b_L. When he misreports his type, he receives b_H with probability $\pi_B + (1-\pi_B)(1-\sum_{q \in \{q' \in Q:\ q'(s_H, b_L) \neq b_H\}} \mu(q))$. In this case, he believes that the buyer is of high-type with probability $\pi_B/(\pi_B + (1-\pi_B)(1 - \sum_{q \in \{q' \in Q:\ q'(s_H, b_L) \neq b_H\}} \mu(q)))$. Therefore, he has no profitable manipulation if

$$b_L - s_L \geq \pi_B(b_H - s_L) + \sum_{q \in \{q' \in Q:\ q'(s_H, b_L) \neq b_H\}} \mu(q)(1-\pi_B)(b_L - s_L),$$

which is more stringent than (D.1). Note that even if such q exists, the interim IC constraints for the buyer and the high-type seller s_H are the same as those in $(\hat{q}^*, \hat{r}^A)$.

In conclusion, the players' IC constraints in μ are at least stringent as those in $(\hat{q}^*, \hat{r}^A)$. Therefore, if μ is an acceptable CE, so is $(\hat{q}^*, \hat{r}^A)$. □

Focusing on $(q^{**}, \hat{r}^A)$, I can establish the necessary and sufficient condition for the existence of pure acceptable CE that implements ex-post efficient outcome in the mediated seller-offer bargaining game.

Proposition 5. *In the mediated seller-offer bargaining game, a pure acceptable CE that implements ex-post efficient outcome exists if and only if* $\pi_B \leq \frac{b_L - s_L}{b_H - s_L}$.

Proof. As shown in the proof of Lemma 9, the mediation plan $(q^{**}, \hat{r}^A)$ is an acceptable CE if and only if $\pi_B \leq \frac{b_L - s_L}{b_H - s_L}$. By Lemma 9, no FB mediation plan $\mu \in \Delta(Q \times \hat{r}^A)$ is an acceptable CE if $\pi_B > \frac{b_L - s_L}{b_H - s_L}$. □

Proposition 5 implies that achieving ex-post efficiency is possible with a seller-biased intermediary (if and) only if it is possible without an intermediary. Intuitively, a seller-biased intermediary has the same incentive as that the seller has in the mediated seller-offer bargaining game. Thus, a seller-biased intermediary is of no help in achieving ex-post efficiency, no matter how small the bias is.

E Implementation without Mediator

The acceptable CE characterized in Section 4 are the equilibria of the game with the fictitious mediator. Therefore, the analysis so far tells nothing about under which communication structure the corresponding equilibrium outcomes can be achieved in the underlying game; that is, the game without the mediator. In this appendix, I come back to the underlying game to answer this question. I show that the noiseless direct communication between the traders and the intermediary is sufficient to achieve ex-post efficiency. However, to achieve the SB outcomes, I need to introduce artificial noise into the communication.

I consider the game between the seller, the buyer, and the intermediary. Their preferences are the same as those defined in Section 2. The timing of the game is as follows:

1. The seller and the buyer privately observe realizations of their types;
2. The traders confidentially report their types $s \in \Theta_S$ and $b \in \Theta_B$ to the intermediary;
3. The intermediary offers a price $p \in \mathbb{R}_+$;
4. The traders simultaneously respond to p, by either acceptance (Y) or rejection (N). If both accept, trade occurs at the price p, and the payoffs are realized; otherwise, no trade occurs, and all players get 0.

The communication between the traders and the intermediary at Time 2 is *noiseless* if the intermediary always receives what the traders have reported; otherwise, it is *noisy*.

E.1 Implementation of the FB outcome

The ex-post efficient outcome characterized in Section 4 is summarized as follows:

- (s_H, b_H) trade at the price b_H;

Table E.1: The intermediary's offer in σ^* as a function of the reports

	b_H	b_L
s_H	b_H	b_H
s_L	b_L	b_L

- (s_H, b_L) do not trade;
- (s_L, b_H) trade at the price b_L;
- (s_L, b_L) trade at the price b_L.

Consider the following strategy profile and the intermediary's system of beliefs about the traders' types:

Strategy σ^*

- Both types of seller report truthfully;
- Both types of buyer report b_L;
- The intermediary offers b_H if he receives (s_H, b_H) or (s_H, b_L) and b_L if he receives (s_L, b_H) or (s_L, b_L) (see Table E.1);
- The seller (resp. the buyer) accepts an offer if and only if it is greater (resp. smaller) than or equal to his type.

The intermediary's system of beliefs ξ^*

- When he receives (s_H, b_H), he assigns probability 1 to (s_H, b_H);
- When he receives (s_H, b_L), he assigns probability π_B to (s_H, b_H) and $1 - \pi_B$ to (s_H, b_L);
- When he receives (s_L, b_H), he assigns probability 1 to (s_L, b_L);
- When he receives (s_L, b_L), he assigns probability π_B to (s_L, b_H) and $1 - \pi_B$ to (s_L, b_L).

Note that following σ^* results in the ex-post efficient outcome described above. Proposition 6 shows that (σ^*, ξ^*) is a perfect Bayesian equilibrium (PBE) of the game if the ex-post efficient outcome can be implemented by acceptable CE in the mediated bargaining game.

Proposition 6. *Assume that the communication between the traders and the intermediary is noiseless. Then, (σ^*, ξ^*) is a PBE of the game if the ex-post efficient outcome can be implemented by acceptable CE in the mediated bargaining game.*

Proof. Note that the system of beliefs ξ^* are consistent with σ^*. The high-type seller s_H has no profitable deviation because misreporting his type leads to the offer b_L, which is unacceptable to him. The low-type seller has no profitable deviation if $b_L - s_L \geq \pi_B(b_H - s_L) \iff \pi_B \leq \frac{b_L - s_L}{b_H - s_L}$. Both types of buyer have no profitable deviation because the intermediary's offer does not depend on the buyer's report.

Given ξ^*, the intermediary's offers are clearly optimal except for when he receives (s_L, b_L). In this case, it is optimal for him to offer b_L if $b_L - s_L \geq \pi_B \lambda(b_H - s_L) \iff \pi_B \leq \frac{b_L - s_L}{\lambda(b_H - s_L)}$. Thus, no player has a profitable deviation and hence (σ^*, ξ^*) is a PBE if $\pi_B \leq \frac{b_L - s_L}{b_H - s_L}$; that is, when the ex-post efficient outcome can be implemented by acceptable CE in the mediated bargaining game. □

This proposition implies that the noiseless direct communication is sufficient to achieve ex-post efficiency.

E.2 Implementation of the SB outcomes

As shown in Proposition 2, there are two possible SB acceptable CE, $x^{**1} = (\overline{x_{HH}}, 0, 0)$ and $x^{**2} \equiv (1, 1, 1)$, where $\overline{x_{HH}} = \frac{(1-\pi_B)(b_L - s_L)}{\pi_B(b_H - b_L)}$.

The SB outcome associated with x^{**1} is summarized as follows:

- (s_H, b_H) trade at the price b_H with probability $\overline{x_{HH}}$ and do not trade with probability $1 - \overline{x_{HH}}$;
- (s_H, b_L) do not trade;
- (s_L, b_H) trade at the price b_L;
- (s_L, b_L) trade at the price b_L.

Consider the following strategy profile, the intermediary's system of beliefs about the traders' types, and the noisy communication:

Table E.2: The intermediary's offer as a function of the reports in σ^{**1}

	b_H	b_L
s_H	b_H	$p \in (b_H, +, \infty)$
s_L	b_L	b_L

Strategy σ^{**1}

- Both types of seller report truthfully;
- Both types of buyer report b_H;
- The intermediary offers b_H if he receives (s_H, b_H), $p \in (b_H, +, \infty)$ if he receives (s_H, b_L), and b_L if he receives (s_L, b_H) or (s_L, b_L) (see Table E.2);
- The seller (resp. the buyer) accepts an offer if and only if it is greater (resp. smaller) than or equal to his type.

The intermediary's system of beliefs ξ^{**1}

- When he receives (s_H, b_H), he assigns probability π_B to (s_H, b_H) and $1 - \pi_B$ to (s_H, b_L);
- When he receives (s_H, b_L), he assigns probability 1 to (s_H, b_L);
- When he receives (s_L, b_H), he assigns probability $\frac{\pi_S \pi_B (1-\overline{x_{HH}})}{\pi_S(1-\overline{x_{HH}})+(1-\pi_S)}$ to (s_H, b_H), $\frac{\pi_S(1-\pi_B)(1-\overline{x_{HH}})}{\pi_S(1-\overline{x_{HH}})+(1-\pi_S)}$ to (s_H, b_L), $\frac{(1-\pi_S)\pi_B}{\pi_S(1-\overline{x_{HH}})+(1-\pi_S)}$ to (s_L, b_H), and $\frac{(1-\pi_S)(1-\pi_B)}{\pi_S(1-\overline{x_{HH}})+(1-\pi_S)}$ to (s_L, b_L);
- When he receives (s_L, b_L), he assigns probability 1 to (s_L, b_L).

Noisy communication

- If the seller reports s_H, then the intermediary will receive s_H with probability $\overline{x_{HH}}$ and s_L with probability $(1 - \overline{x_{HH}})$;
- In the other cases, the intermediary always receives what the traders have reported.

Note that, combined with the above noisy communication, following σ^{**1} results in the SB outcome associated with x^{**1}. Proposition 7 shows that $(\sigma^{**1}, \xi^{**1})$ is a PBE of the game if the SB acceptable CE is x^{**1}.

Proposition 7. *Assume that the communication between the traders and the intermediary is noisy as described above. Then,* $(\sigma^{**1}, \xi^{**1})$ *is a PBE of the game if the SB acceptable CE is* x^{**1}*.*

Proof. Note that given the noisy communication described above, the system of beliefs ξ^{**1} are consistent with σ^{**1}. It is easy to see that the high-type seller s_H and both types of buyer have no profitable deviation. If the low-type seller s_L follows σ^{**1}, then his expected payoff is $b_L - s_L$. If he misreports his type, then his expected payoff is at most

$$\pi_B \overline{x_{HH}}(b_H - s_L) + \pi_B(1 - \overline{x_{HH}})(b_L - s_L) = \pi_B(b_H - s_L) - \pi_B(1 - \overline{x_{HH}})(b_H - b_L) = b_L - s_L.$$

Thus, he has no profitable deviation.

Given ξ^{**1}, the intermediary's offers are clearly optimal except for when he receives (s_L, b_H). In this case, if he offers p, then his expected payoff is

$$\begin{cases} \xi^{**1}(s_H, b_H \mid s_L, b_H)v(p \mid s_H, b_H) + \xi^{**1}(s_L, b_H \mid s_L, b_H)v(p \mid s_L, b_H) & \text{if } p \in [s_H, b_H] \\ \xi^{**1}(s_L, b_H \mid s_L, b_H)v(p \mid s_L, b_H) & \text{if } p \in (b_L, s_H) \\ \xi^{**1}(s_L, b_H \mid s_L, b_H)v(p \mid s_L, b_H) + \xi^{**1}(s_L, b_L \mid s_L, b_H)v(p \mid s_L, b_L) & \text{if } p \in [s_L, b_L] \\ 0 & \text{otherwise,} \end{cases}$$

where $\xi^{**1}(\tilde{s}, \tilde{b} \mid s, b)$ is the intermediary's belief that the traders are $(\tilde{s}, \tilde{b}) \in \Theta_S \times \Theta_B$ when he receives $(s, b) \in \Theta_S \times \Theta_B$. Thus, it is optimal for him to offer b_L if

$$\begin{aligned} & \xi^{**1}(s_L, b_H \mid s_L, b_H)v(b_L \mid s_L, b_H) + \xi^{**1}(s_L, b_L \mid s_L, b_H)v(b_L \mid s_L, b_L) \\ & \geq \xi^{**1}(s_H, b_H \mid s_L, b_H)v(b_H \mid s_H, b_H) + \xi^{**1}(s_L, b_H \mid s_L, b_H)v(b_H \mid s_L, b_H) \\ \iff & (1 - \pi_S)\pi_B[\lambda(b_L - s_L) + (1 - \lambda)(b_H - b_L)] + (1 - \pi_S)(1 - \pi_B)\lambda(b_L - s_L) \\ & \geq \pi_S \pi_B(1 - \overline{x_{HH}})\lambda(b_H - s_H) + (1 - \pi_S)\pi_B\lambda(b_H - s_L), \end{aligned}$$

which is equivalent to (IC-Int2). Thus, no player has a profitable deviation and hence $(\sigma^{**1}, \xi^{**1})$ is a PBE if the SB-ACE is characterized by x^{**1}. □

Next, the SB outcome associated with x^{**2} is summarized as follows:

Table E.3: The intermediary's offer as a function of the reports in σ^{**2}

	b_H	b_L
s_H	b_H	$p \in (b_H, +, \infty)$
s_L	b_H	b_H

- (s_H, b_H) trade at the price b_H;
- (s_H, b_L) do not trade;
- (s_L, b_H) trade at the price b_H;
- (s_L, b_L) do not trade.

Consider the following strategy profile and the intermediary's system of beliefs about the traders' types:

Strategy σ^{**2}

- Both types of seller report truthfully;
- Both types of buyer report b_H;
- The intermediary offers b_H if he receives (s_H, b_H), (s_L, b_H), or (s_L, b_L) and $p \in (b_H, +, \infty)$ if he receives (s_H, b_L) (see Table E.3);
- The seller (resp. the buyer) accepts an offer if and only if it is greater (resp. smaller) than or equal to his type.

The intermediary's system of beliefs ξ^{**2}

- When he receives (s_H, b_H), he assigns probability π_B to (s_H, b_H) and $1 - \pi_B$ to (s_H, b_L);
- When he receives (s_H, b_L), he assigns probability 1 to (s_H, b_L);
- When he receives (s_L, b_H), he assigns probability $\frac{\pi_S \pi_B (1-\overline{x}_{HH})}{\pi_S(1-\overline{x}_{HH})+(1-\pi_S)}$ to (s_H, b_H), $\frac{\pi_S(1-\pi_B)(1-\overline{x}_{HH})}{\pi_S(1-\overline{x}_{HH})+(1-\pi_S)}$ to (s_H, b_L), $\frac{(1-\pi_S)\pi_B}{\pi_S(1-\overline{x}_{HH})+(1-\pi_S)}$ to (s_L, b_H), and $\frac{(1-\pi_S)(1-\pi_B)}{\pi_S(1-\overline{x}_{HH})+(1-\pi_S)}$ to (s_L, b_L);
- When he receives (s_L, b_L), he assigns probability 1 to (s_L, b_H).

Two strategy profiles σ^{**1} and σ^{**2} differ only in the intermediary's offers when he receives (s_L, b_H) or (s_L, b_L). Also, two system of beliefs ξ^{**1} and ξ^{**2} differ only in the off-path belief when the intermediary receives (s_L, b_L). Note that, combined with the same noisy communication as before, following σ^{**2} results in the SB outcome associated with x^{**2}. Proposition 8 shows that $(\sigma^{**2}, \xi^{**2})$ is a PBE of the game if the SB acceptable CE is x^{**2}.

Proposition 8. *Assume that the communication between the traders and the intermediary is noisy as described above. Then,* $(\sigma^{**2}, \xi^{**2})$ *is a PBE of the game if the SB acceptable CE is* x^{**2}.

Proof. Note that given the noisy communication described above, the system of beliefs ξ^{**2} are consistent with σ^{**2}. It is easy to see that the traders have no profitable deviation.

By the proof of Proposition 7, the intermediary's offers are optimal given ξ^{**2} if

$$\begin{aligned}&(1-\pi_S)\pi_B[\lambda(b_L - s_L) + (1-\lambda)(b_H - b_L)] + (1-\pi_S)(1-\pi_B)\lambda(b_L - s_L)\\ &\leq \pi_S\pi_B(1-\overline{x_{HH}})\lambda(b_H - s_H) + (1-\pi_S)\pi_B\lambda(b_H - s_L).\end{aligned}$$

Thus, no player has a profitable deviation and hence $(\sigma^{**2}, \xi^{**2})$ is a PBE if the SB acceptable CE is x^{**2}. □

Proposition 7 and 8 imply that, in contrast to the implementation of the ex-post efficient outcome, some artificial noise is necessary to implement the SB outcomes if I restrict my attention to the one-round direct communication.

References

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

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

Bester, Helmut and Roland Strausz (2007), "Contracting with imperfect commitment and noisy communication." *Journal of Economic Theory*, 136(1), 236–259.

Čopič, Jernej and Clara Ponsatí (2008), "Robust bilateral trade and mediated bargaining." *Journal of the European Economic Association*, 6(2-3), 570–580.

Doval, Laura and Vasiliki Skreta (2022), "Mechanism design with limited commitment." *Econometrica*, 90(4), 1463–1500.

Eilat, Ran and Ady Pauzner (2021), "Bilateral trade with a benevolent intermediary." *Theoretical Economics*, 16(4), 1655–1714.

Fanning, Jack (2021), "Mediation in reputational bargaining." *American Economic Review*, 111(8), 2444–72.

Fanning, Jack (2023), "Optimal dynamic mediation." *Journal of Political Economy*, 131(11), 2969–3002.

Forges, Francoise (1986), "An approach to communication equilibria." *Econometrica*, 54(6), 1375–1385.

Ganguly, Chirantan and Indrajit Ray (2012), "*Simple* mediation in a cheap-talk game." Working Paper, University of Birmingham.

Goltsman, Maria, Johannes Hörner, Gregory Pavlov, and Francesco Squintani (2009), "Mediation, arbitration and negotiation." *Journal of Economic Theory*, 144(4), 1397–1420.

Gottardi, Piero and Claudio Mezzetti (2024), "Shuttle diplomacy." *Journal of Economic Theory*, 216, 105794.

Hörner, Johannes, Massimo Morelli, and Francesco Squintani (2015), "Mediation and peace." *The Review of Economic Studies*, 82(4), 1483–1501.

Jarque, Xavier, Clara Ponsatí, and József Sákovics (2003), "Mediation: incomplete information bargaining with filtered communication." *Journal of Mathematical Economics*, 39(7), 803–830.

Kydd, Andrew H. (2003), "Which side are you on? bias, credibility, and mediation." *American Journal of Political Science*, 47(4), 597–611.

Kydd, Andrew H. (2006), “When can mediators build trust?” *American Political Science Review*, 100(3), 449–462.

Loertscher, Simon and Leslie M. Marx (2022), “Incomplete information bargaining with applications to mergers, investment, and vertical integration.” *American Economic Review*, 112(2), 616—49.

Lomys, Niccolò and Takuro Yamashita (2021), “A mediator approach to mechanism design with limited commitment.” Working Paper.

Matsuo, Toshihide (1989), “On incentive compatible, individually rational, and ex post efficient mechanisms for bilateral trading.” *Journal of Economic Theory*, 49(1), 189–194.

Myerson, Roger B. (1982), “Optimal coordination mechanisms in generalized principal-agent problems.” *Journal of Mathematical Economics*, 10(1), 67–81.

Myerson, Roger B. (1986a), “Acceptable and predominant correlated equilibria.” *International Journal of Game Theory*, 15(3), 133–154.

Myerson, Roger B. (1986b), “Multistage games with communication.” *Econometrica*, 54(2), 323–358.

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

Sugaya, Takuo and Alexander Wolitzky (2021), “The revelation principle in multistage games.” *The Review of Economic Studies*, 88(3), 1503–1540.