7-2024

Efficient bilateral trade with interdependent values: The use of two-stage mechanisms

Takashi KUNIMOTO
Singapore Management University, tkunimoto@smu.edu.sg

Cuiling ZHANG
Singapore Management University, clzhang.2015@phdecons.smu.edu.sg

Efficient Bilateral Trade with Interdependent Values: The Use of Two-Stage Mechanisms*

Takashi Kunimoto† Cuiling Zhang‡

First Version: June 2020
This Version: July 2024

Abstract

Efficient, voluntary bilateral trades are generally not implementable in an interdependent values environment where agents' information is ex ante symmetric (i.e., both parties have private information and each party's valuation depends on the other's information in the same way). Thus, we seek more positive results by employing *two-stage* mechanisms in which (i) the outcome (e.g., allocation of the goods) is determined first; (ii) the agents partially learn the state via their own outcome-decision payoffs; and (iii) transfers are finally made. We propose the *generalized shoot-the-liar mechanism* as a two-stage mechanism and a mild condition under which there is no loss of generality in focusing on the generalized shoot-the-liar mechanism to implement efficient voluntary trades. We then identify a necessary and sufficient condition for the generalized shoot-the-liar mechanism to implement efficient, voluntary trades. We argue by example that the identified condition implies that the degree of interdependence of the buyer's preferences is not too high relative to the seller's counterpart. The identified condition becomes a vacuous constraint if we settle for virtual (or, approximate) implementation.

JEL Classification: C72, D78, D82.

*We are indebted to Claudio Mezzetti for a number of his insightful comments on the first draft of the paper. We thank Shurojit Chatterji, Yi-Chun Chen, Daisuke Hirata, Atsushi Kajii, Jiangtao Li, Ming Li, Nozomu Muto, Daisuke Oyama, Jingyi Xue, Takuro Yamashita, Seung Han Yoo, and the conference/seminar audiences at Hitotsubashi University, Korea University, Nanjing Audit University, and SMU for useful comments.

†School of Economics, Singapore Management University; `tkunimoto@smu.edu.sg`

‡The Institute for Advanced Economic Research, Dongbei University of Finance and Economics; `cuilingzhang01@outlook.com`

Keywords: bilateral trade, interdependent values, two-stage mechanisms.

1 Introduction

This paper investigates when efficient, voluntary bilateral trades are possible in an interdependent values environment with two-sided asymmetric information, where both parties have private information and each party's valuation depends on the other's information in the same way. By "bilateral trade" we mean a simple trading problem in which two individuals, one of whom (a seller) has a single indivisible object to sell to the other (a buyer), attempt to agree to exchange the object for money. Thus, in this setup, the seller has the full property rights over the object to be sold, while she is not fully informed about the quality of the good at the outset. For example, it is quite natural to imagine that the owner of an apartment does not necessarily know the full value of the apartment while thinking about selling it to a potential buyer. If we instead assume that the seller is fully informed about the quality of the good from the beginning, we refer the reader to Kunimoto and Zhang (2022) for the analysis in such a case. Efficiency adopted in this paper requires, conditional on the realization of both agents types, that (i) the good be traded if and only if the buyer's expected valuation of the good is strictly higher as the seller's counterpart (*decision efficiency*) and (ii) the expected payment made by the buyer be always exactly the expected payment received by the seller (*budget balance*). It is worth noting that our formulation allows an additional uncertainty about both agents' valuation for the good, even conditional on their type profile. This additional generality plays an important role when we deal with two-stage mechanisms later. Voluntary trade means that each agent of every type weakly prefers to participate in the mechanism (*interim individual rationality*).

By the well-known revelation principle, we say that efficient, voluntary trades are implementable if there exists a *direct revelation* mechanism in which each agent is asked to report his own type and telling the true type profile constitutes a Bayesian Nash equilibrium (i.e., Bayesian incentive compatibility (BIC)) such that the mechanism also satisfies decision efficiency (EFF), interim individual rationality (IIR), and budget balance (BB). In our bilateral trade environment, however, Fieseler, Kittsteiner, and Moldovanu (henceforth, FKM, 2003) and Gresik (1991) show that efficient, voluntary bilateral trade is generally not implementable.

The implicit assumption underlying FKM (2003), Gresik (1991), and many other papers in the literature is that agents cannot observe their outcome-decision payoffs until the mechanism has played out. We now drop this assumption, and seek more positive results by incorporating *two-stage* mechanisms (Mezzetti (2004)) in which (i) the outcome (e.g., the allocation of the goods) is determined first; (ii) the agents then observe their own outcome-decision payoffs; and (iii) finally, transfers are made. In his Proposition 1, Mezzetti (2003) establishes the generalized revelation principle, which states that any perfect Bayesian equilibrium outcome of any two-stage mechanism can be implemented as a perfect Bayesian equilibrium of a two-stage generalized revelation mechanism in which reporting his true allocation payoff in the second stage and reporting his true type in the first stage is an equilibrium strategy for each player. By this generalized revelation principle, a two-stage generalized revelation mechanism is simply called a two-stage mechanism in this paper. Following the generalized revelation principle, we thus modify the notion of Bayesian incentive compatibility: a two-stage mechanism satisfies BIC if there exists a perfect Bayesian equilibrium of that two-stage mechanism in which all agents tell the truth in both stages. The main question of our paper is thus reformulated as: "when does there exist a two-stage mechanism satisfying BIC, IIR, EFF, and BB?" We consider this as a primarily normative question, drawing the line between the existence and non-existence of such two-stage mechanisms. Thus, we closely follow Myerson and Satterthwaite (1983) in terms of the main research question we address.

The use of two-stage mechanisms can be justified. Imagine that two parties in a bilateral trade setup invite a trusted mediator (a third party) to their contractual relationship: the mediator asks both agents to deposit a large amount of money in the mediator's account, and the mediator returns the remaining deposit to each agent after the two-stage mechanism has played out. For example, in the context of a labor market, an employer (the buyer) learns the quality of the worker (the seller) after hiring him. If the worker was not hired by the employer, i.e., no trade takes place, we assume that the worker will eventually be hired by another employer. When this happens, the worker learns the value of himself through his signed constract. The power of two-stage mechanisms relies crucially on the fact that the extra information the agents learn via their experienced payoff can be used in its second stage to detect the lies which might have occurred in its first stage. What is not standard here is that the contract spans the stage where the worker who was not hired by this employer, but was eventually hired by someone else, must report his experienced payoff to the original two-stage mechanism. We admit that the power of two-stage mechanism may well be compromised, since this contractual apparatus requires a strong commitment on the part of agents over a long time horizon. Therefore, we are primarily concerned with the theoretical possibility of two-stage mechanisms. Nevertheless, we may be able to overcome this formidable contractual condition through a *smart contract* based on the blockchain technology as a commitment device that prevents agents from reneging on the contract terms (see, for example, Matsushima and Noda (2023)). Even without a trusted mediator or smart contracts, a two-stage mechanism can sometimes be realized through a long-term relationship. See Mezzetti (2004) for such an argument.

We then describe the main analysis of the paper. In Section 3, we propose the *generalized shoot-the-liar mechanism (henceforth, GS mechanism)* as a generalization of Mezzetti's (2007) shoot-the-liar mechanism used in an auction setting. We show in Proposition 1 that the GS mechanism always satisfies EFF, BB, and the buyer's IIR.

In Section 4, we introduce Assumption 1, and show in our Proposition 2 that the GS mechanism satisfies BIC if and only if Assumption 1 is satisfied. We next show that when Assumption 1 holds, the GS mechanism is canonical in the sense that if there exists a two-stage mechanism satisfying BIC, EFF, BB, and IIR, the GS mechanism has the same desired properties. Finally, we discuss when Assumption 1 holds and when it does not.

In Section 5, we introduce Assumption 2 and show in Proposition 4 that when Assumption 1 holds, the GS mechanism satisfies the seller's IIR constraint if and only if Assumption 2 is satisfied. Therefore, we conclude that, when Assumption 1 holds, Assumption 2 is a necessary and sufficient condition for the existence of two-stage mechanisms satisfying BIC, EFF, BB, and IIR.

Section 6 scrutinizes the implications of Assumption 2. We apply our main results from Section 5 to a stylized model in which each agent's type is chosen from the uniform distribution over $[0, 1]$ and the state of the residual uncertainty $\omega = (\omega_1, \omega_2)$ is a two-dimensional vector where, conditional on the type profile, ω_1 and ω_2 are independently drawn from the uniform distribution in $[-0.1, 0.1]$. The valuation of each agent i for the object is determined by the type profile as well as the state of the residual uncertainty, which is $\tilde{u}_i(\theta_i, \theta_{-i}; \omega) = (1 + \omega_i)(\theta_i + \gamma_i\theta_{-i})$, where γ_i denotes the degree of interdependence of preferences for agent i. In this context, Assumption 2 is satisfied as long as the degree of interdependence of the buyer's preferences (γ_2) is not too high relative to the seller's counterpart (γ_1). By providing a set of simulation results in the example, we also conclude that Assumption 2 certainly has a bite but is satisfied for many of the cases.

Section 7 consists of three subsections. In Section 7.1, we show that Assumption 2 can be completely dispensed with if we settle for virtual (i.e., approximate) implementation. We also show that Assumption 1 is replaced with Assumption

3 if we settle for virtual implementation. Section 7.2 compares our results with Theorem 4 of Galavotti, Muto, and Oyama (henceforth GMO, 2011), who consider the problem of efficient partnership dissolution in an interdependent values environment.[1] GMO (2011, Theorem 4) show that when GMO's Assumption 5.1 is satisfied, Mezzetti's (2007) shoot-the-liar mechanism satisfies BIC, IIR, EFF, and BB for all ownership structures (including our bilateral trade case) in the deterministic model.[2] To make our comparison meaningful, we focus on bilateral trade in the deterministic model, i.e., there are only two agents, the seller has the full property rights over the good, and the state of the residual uncertainty is a deterministic function of the type profile. We then argue that GMO's Assumption 5.1 is stronger than our Assumption 2 in the deterministic model. Section 7.3 discusses the relation between our Theorem 2 and Theorem 3.1 of Makowski and Mezzetti (1994), who consider the general mechanism design problem with private values.

The rest of the paper is organized as follows. In Section 2, we introduce the general notation and basic concepts of the paper. In Section 3, we introduce the GS mechanism. In Section 4, we show that the GS mechanism satisfies BIC if and only if Assumption 1 holds. We also discuss when Assumption 1 is satisfied and when it is violated. In Section 5, if Assumption 1 holds, we show that there exists a two-stage mechanism satisfying BIC, EFF, BB, and IIR if and only if Assumption 2 is satisfied. Section 6 specializes in a highly stylized but well studied model of bilateral trade with interdependent values and explores the implications of our main results in Section 5. In Section 7, we clarify that if we settle for virtual implementation, Assumption 2 can be dropped and Assumption 1 is replaced with Assumption 3. We also compare our paper with GMO (2011) and Makowski and Mezzetti (1994). Finally, Section 8 concludes the paper. The Appendix contains all the proofs of the results omitted from the main text of the paper.

2 Preliminaries

2.1 The Basic Setup

We consider a bilateral trade environment in which a *seller (agent 1)* initially owns an indivisible object which one potential *buyer (agent 2)* intends to obtain. Each

[1] Efficient dissolution of a partnership consists in allocating the partnership's asset (e.g., the developed product/technology or the company itself) to the partner with the highest valuation, in exchange for monetary compensations.

[2] To be precise, their result is stronger than this because GMO (2011) strengthen IIR into ex post individual rationality (EPIR).

agent $i \in \{1,2\}$ has private information, which is summarized as his own type θ_i. For each $i \in \{1,2\}$, we let $\Theta_i = [\underline{\theta}, \overline{\theta}]$ be the common set of agent i's types such that $\underline{\theta} < \overline{\theta}$. Throughout the paper, we use the notation convention that $\Theta = [\underline{\theta}, \overline{\theta}]^2$; and $\Theta_{-i} = \Theta_j$ where $j \neq i$ with a generic element θ_{-i}. Types are independently drawn from an identical distribution where $f : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}_+$ and $F : [\underline{\theta}, \overline{\theta}] \to [0,1]$ denote the probability density function and cumulative distribution function, respectively. We further assume that $f(\theta_i) > 0$ for all $\theta_i \in (\underline{\theta}, \overline{\theta})$ and $i \in \{1,2\}$.

We introduce what Mezzetti (2007) calls the "random model." Let ω denote a residual uncertainty and Ω be the set of residual uncertainties, which is assumed to be a nonempty subset of a Euclidean space. Let $q \in Q = [0,1]$ be the probability that the good is transferred from the seller (agent 1) to the buyer (agent 2), or trading probability for short. The residual uncertainty ω is a random function of the type profile θ drawn from a cumulative distribution function conditional on the agents' types, denoted by $H : \Omega \times \Theta \to [0,1]$. Preferences of each agent $i \in \{1,2\}$ are given by $U_i : Q \times \Theta \times \Omega \times \mathbb{R} \to \mathbb{R}$, which depends on the trading probability q, the type profile θ, the residual uncertainty ω, and their monetary transfer p_i:

$$\begin{aligned}
U_1(q,\theta,\omega,p_1) &= u_1(q,\theta,\omega) + p_1 = (1-q)\tilde{u}_1(\theta;\omega) + p_1; \\
U_2(q,\theta,\omega,p_2) &= u_2(q,\theta,\omega) + p_2 = q\tilde{u}_2(\theta;\omega) + p_2,
\end{aligned}$$

where $u_i(q,\theta,\omega)$ is agent i's allocation payoff and $\tilde{u}_i(\theta;\omega)$ is agent i's valuation for the object when the true type profile is θ and the residual uncertainty is $\omega \in \Omega$. Hence, we assume expected, quasilinear utility. Agent i's "expected" valuation for the object, conditional on type profile θ, is

$$\tilde{u}_i^e(\theta) \equiv \int_\Omega \tilde{u}_i(\theta;\omega)dH(\omega|\theta).$$

It is standard in the literature that one uses a reduced-form model in which the residual uncertainty ω is suppressed and an agent's outcome-decision payoff depends directly on the types of all the agents. For instance, we can reconcile our formulation with the reduced-form model by interpreting θ as a vector of type profile θ and $(\tilde{u}_i^e(\theta))_{i\in\{1,2\}}$ as the profile of agents' expected valuations in type profile θ. These two formulations turn out to be equivalent if an agent cannot observe his own outcome-decision payoff. However, when an agent can observe his own outcome-decision payoff before final transfers are made, this equivalence breaks down. It is worth mentioning that the agent's ability to observe his own outcome-decision payoff becomes irrelevant if we consider private values environments. Hence, we need genuinely *interdependent* values environments for our discussion. Furthermore, our formulation follows Mezzetti (2004, 2007) so that it allows residual uncertainties in the agent's payoff even after observing his own outcome-decision payoff. The example below illustrates our formulation.

Example 1. Suppose that each agent i's type θ_i is drawn from the uniform distribution on the unit interval $[0,1]$. Conditional on the type profile θ, ω_1 and ω_2 are independently drawn from the uniform distribution on the closed interval $[-0.1, 0.1]$. Then, we define $\Omega = \{(\omega_1, \omega_2)|\ \omega_1, \omega_2 \in [-0.1, 0.1]\}$ as the set of residual uncertainties. Each agent's valuation for the object when the true type profile is (θ_1, θ_2) and the residual uncertainty is $\omega = (\omega_1, \omega_2) \in \Omega$ is given as follows:

$$\begin{aligned}\tilde{u}_1(\theta_1, \theta_2; \omega) &= \theta_1 + \gamma_1\theta_2 + \omega_1(\theta_1 + \gamma_1\theta_2) = (1+\omega_1)(\theta_1 + \gamma_1\theta_2);\\ \tilde{u}_2(\theta_1, \theta_2; \omega) &= \theta_2 + \gamma_2\theta_1 + \omega_2(\theta_2 + \gamma_2\theta_1) = (1+\omega_2)(\theta_2 + \gamma_2\theta_1),\end{aligned}$$

where γ_i denotes the degree of interdependence of preferences for agent i. We assume $\gamma_i \in (0,1)$ for all $i \in \{1,2\}$. In this paper, we assume that the agents are risk-neutral so that we could restrict attention to the agents' expected utilities. We thus set the support of ω_i to be small relative to that of θ_i so that our result holds if agents are slightly risk averse.

Throughout the paper, we revisit Example 1 multiple times to illustrate the implications of our analysis. Recall $\tilde{u}_i^e(\theta_i, \theta_{-i}) = \int_\Omega \tilde{u}_i(\theta_i, \theta_{-i}; \omega) dH\left(\omega|(\theta_i, \theta_{-i})\right)$. We assume that $\tilde{u}_i^e$ is bounded and that $H(\cdot|\theta)$ is differentiable with respect to ω. So, we write $h(\omega|\theta) = dH(\omega|\theta)/d\omega$. Using Leibniz's formula, we obtain

$$\begin{aligned}&\frac{\partial}{\partial\theta_i}\int_\Omega \tilde{u}_i(\theta_i, \theta_{-i}; \omega) h\left(\omega|(\theta_i, \theta_{-i})\right) d\omega\\ =\ &\int_\Omega \left[\frac{\partial \tilde{u}_i(\theta_i, \theta_{-i}; \omega)}{\partial\theta_i} h\left(\omega|(\theta_i, \theta_{-i})\right) + \tilde{u}_i(\theta_i, \theta_{-i}; \omega)\frac{\partial h\left(\omega|(\theta_i, \theta_{-i})\right)}{\partial\theta_i}\right] d\omega.\end{aligned}$$

We assume that $\tilde{u}_i(\theta_i, \theta_{-i}; \omega)$ and $h\left(\omega|(\theta_i, \theta_{-i})\right)$ are also differentiable in both θ_i and θ_{-i}. As a result, $\tilde{u}_i^e(\cdot)$ is differentiable in both θ_i and θ_{-i}. Moreover, we assume each agent's expected valuation function is strictly increasing in his own type, i.e., $\tilde{u}_{i,i}^e \equiv \partial\tilde{u}_i^e(\theta_i, \theta_{-i})/\partial\theta_i > 0$. We see this feature in Example 1: $\partial\tilde{u}_i(\theta_i, \theta_{-i}; \omega)/\partial\theta_i = 1 + \omega_i > 0$ and $\partial h\left(\omega|(\theta_i, \theta_{-i})\right)/\partial\theta_i = 0$ because $\omega = (\omega_i, \omega_{-i})$ and conditional on the type profile, ω_i and ω_{-i} are independently drawn from the uniform distribution in $[-0.1, 0.1]$. As a result, we obtain $\frac{\partial}{\partial\theta_i}\int_\Omega \tilde{u}_i(\theta_i, \theta_{-i}; \omega) h\left(\omega|(\theta_i, \theta_{-i})\right) d\omega > 0$ so that $\tilde{u}_i^e : (\theta_i, \theta_{-i}) \mapsto \mathbb{R}$ is strictly increasing in θ_i.

2.2 One-Stage Mechanisms

We first introduce the notion of (one-stage) direct revelation mechanism. A one-stage direct revelation mechanism is defined as a triple (Θ, x, t) in which each agent announces his type and thereafter, the allocation decision is determined by the rule $x : \Theta \to [0,1]$ and the monetary transfer is determined by $t : \Theta \to \mathbb{R}^2$ "simultaneously" based on all agents' type announcements. By the standard revelation principle, we lose nothing to focus on direct revelation mechanisms in which truth-telling of each agent's type constitutes a Bayesian Nash equilibrium, which is known as Bayesian incentive compatibility (BIC). From now on, we simply call a one-stage direct mechanism a one-stage mechanism.

Consider a special case of Example 1 in which $\omega_i = 0$ for each agent $i \in \{1,2\}$. Then, the residual uncertainty $\omega = (\omega_1, \omega_2)$ is reduced to $(0,0)$, which further implies that $\tilde{u}_i(\theta_i, \theta_{-i}; \omega) = (1+\omega_i)(\theta_i + \gamma_i \theta_{-i})$ is reduced to $\theta_i + \gamma_i \theta_{-i}$. Making use of Theorem 5 of FKM (2003), Kunimoto and Zhang (2022) show that there are no one-stage mechanisms satisfying BIC, IIR, EFF, and BB in this example. This is the main reason why we introduce two-stage mechanisms below.

2.3 Two-Stage Mechanisms

To justify the use of two-stage mechanisms, which will be introduced below, we assume that once agent i receives the object, he observes his realized allocation payoff $\tilde{u}_i(\theta_i, \theta_{-i}; \omega)$. We follow Mezzetti (2004) to define a *two-stage* mechanism as a quadruple (M^1, M^2, δ, τ) such that

- M_i^1 is agent i's message space in the first stage and M_i^2 is agent i's message space in the second stage, respectively;
- $\delta : M^1 \to [0,1]$ is the decision rule specifying the trading probability; and
- $\tau : M^1 \times M^2 \to \mathbb{R}^2$ is the transfer rule specifying the monetary transfer for both agents.

In words, in the first stage, after observing his own type, agent i sends a message from M_i^1 simultaneously and then the good is allocated according to the decision rule δ; in the second stage, after the decision is implemented and agent i observes his realized allocation payoff, he is asked to send a message from M_i^2; and finally, the monetary transfers are finalized based on the reports of both stages. We denote by $\widehat{Q} = \{0,1\}$ the final decision outcome after randomization and $\Pi_i = \{\hat{u}_i \in \mathbb{R} :$ there exist $q \in \hat{Q}, \theta \in \Theta,$ and $\omega \in \Omega$ such that $u_i(q, \theta, \omega) = \hat{u}_i\}$ the range of agent i's allocation payoffs. Note that whoever does not receive the good receives zero allocation payoff. We further denote by $r_i = (r_i^1, r_i^2)$ agent i's strategy such that $r_i^1 : \Theta_i \to M_i^1$ is his strategy in the first stage and $r_i^2 : \widehat{Q} \times \Theta_i \times \Pi_i \to M_i^2$ is his strategy in the second stage. The reader is referred to the fourth paragraph of the introduction for the justification of two-stage mechanisms.

In particular, if we set $M_i^1 = \Theta_i$ and $M_i^2 = \Pi_i$, i.e., the agents are asked to report their types in the first stage and realized allocation payoffs in the second stage, then we can construct the corresponding *generalized revelation mechanism* (Θ, Π, x, t) as follows: the decision rule $x : \Theta \to [0,1]$ is given by the composite function $x(\theta) = \delta(r^1(\theta))$ and the transfer rule $t : \Theta \times \Pi \to \mathbb{R}^2$ is given by the composite function $t_i(\theta, \pi) = \tau_i(r^1(\theta), r^2(\delta(r^1(\theta)), \theta, \pi))$. Since each agent i's allocation payoff $u_i(x(\theta_i, \theta_{-i}), \theta_i, \theta_{-i}, \omega)$ depends on the true type profile and the residual uncertainty, then the second-stage reports in the generalized revelation mechanism indeed provide extra information about the true type profile and the residual uncertainty, while there is a loss of generality in assuming that the designer only uses one-stage mechanisms.

There are some scenarios where the agent experiencing the good learns its value before the transfers are finalized. Imagine that two parties in a bilateral trade setup invite a trusted mediator (a third party) to their contractual relationship: the mediator asks both agents to deposit a large amount of money in the mediator's account, and the mediator returns the remaining deposit to each agent after the two-stage mechanism has played out. For example, we consider an apartment owner (the seller) who partially knows the value of her property and is interested in selling it to a potential buyer (the buyer). When the buyer acquires this property, he receives noisy but additional information about its quality. If the seller could not sell her apartment to the original buyer, we assume that both the seller and the original buyer, who did not get the property, are still bound by the original two-stage mechanism until the seller eventually sells her property to another buyer, receives noisy but additional information about the value of her property, and reports her experienced value to the mechanism. We recognize that this is not a standard assumption, since this contractual apparatus requires a strong commitment on the part of agents over a long time horizon. Thus, we are primarily concerned with the theoretical possibility of two-stage mechanisms so that we postpone our discussion of how to implement such a theoretical possibility. Nevertheless, we may be able to implement this formidable contractual apparatus through a *smart contract* based on the blockchain technology as a commitment device that prevents agents from reneging on the contract terms (see, for example, Matsushima and Noda (2023)).

Following Mezzetti (2003), we adopt *perfect Bayesian equilibrium* as a solution concept in extensive games with incomplete information and appeal to the following generalized revelation principle, i.e., the counterpart of revelation principle in one-stage mechanisms.[3]

Lemma 1 (The Generalized Revelation Principle in Mezzetti (2003))**.** Any perfect Bayesian equilibrium outcome of any two-stage mechanism can be implemented as a perfect Bayesian equilibrium of a generalized revelation mechanism in which reporting their true allocation payoff in the second stage and reporting their true type in the first stage is an equilibrium strategy for each player.

From now on, by the generalized revelation principle, we call a generalized revelation mechanism simply a two-stage mechanism. We denote by (θ_1^r, θ_2^r) the first-stage report and by (u_1^r, u_2^r) the second-stage report in a two-stage mechanism, respectively. We next introduce the main properties we want our two-stage mechanisms to satisfy. We adopt the BIC constraint proposed by Mezzetti (2004).

Definition 1. A two-stage mechanism (Θ, Π, x, t) satisfies *Bayesian incentive compatibility* (BIC) if truth-telling in both stages constitutes an equilibrium strategy of each agent in a perfect Bayesian equilibrium; that is, for each agent $i \in \{1, 2\}$, each pair of type profiles $(\theta_i, \theta_{-i}), (\theta_i^r, \theta_{-i}^r) \in \Theta$, and each realized state of the world ω, the equilibrium second-stage report is $u_i^r = u_i(x(\theta_i^r, \theta_{-i}^r), \theta_i, \theta_{-i}, \omega)$ and the equilibrium first-stage report is $\theta_i^r = \theta_i$.

BIC implies that, given the first-stage report, each agent reports their realized allocation payoff truthfully in the second stage. BIC further implies that, on the equilibrium path, each agent reports their type truthfully in the first stage and for any type profile $(\theta_i, \theta_{-i}) \in \Theta$, $u_i(x(\theta_i, \theta_{-i}), \theta_i, \theta_{-i}, \omega)$ is agent i's true allocation payoff.

We assume that each agent has the option of not participating in the two-stage mechanism (Θ, Π, x, t) and let $U_i^O(\theta_i)$ be the expected utility of agent i with type θ_i from non-participation. To be specific,

$$U_1^O(\theta_1) = \int_{\Theta_2} \int_{\Omega} \tilde{u}_1(\theta_1, \theta_2; \omega) dH\left(\omega | (\theta_1, \theta_2)\right) dF(\theta_2) \text{ for all } \theta_1 \in \Theta_1$$

and

$$U_2^O(\theta_2) = 0 \text{ for all } \theta_2 \in \Theta_2.$$

Note that when contemplating over whether to participate in the mechanism, each agent is uninformed of the other agent's type or the residual uncertainty. So, when

[3] For perfect Bayesian equilibrium, for example, the reader is referred to Osborne and Rubinstein (1994, pp.232-233).

we compute an agent's expected utility, we should take expectations with respect to both the other agent's type and the residual uncertainty. We introduce the following individual rationality constraint:

Definition 2. A two-stage mechanism (Θ, Π, x, t) satisfies *interim individual rationality* (IIR) if, for each agent $i \in \{1, 2\}$ and each $\theta_i \in \Theta_i$,

$$\int_{\Theta_{-i}} \int_{\Omega} [u_i(x(\theta_i, \theta_{-i}), \theta_i, \theta_{-i}, \omega) + t_i(\theta_i, \theta_{-i}; u_i, u_{-i})] \, dH\left(\omega | (\theta_i, \theta_{-i})\right) dF(\theta_{-i}) \geq U_i^O(\theta_i),$$

where $u_i = u_i(x(\theta_i, \theta_{-i}), \theta_i, \theta_{-i}, \omega)$ and $u_{-i} = u_{-i}(x(\theta_i, \theta_{-i}), \theta_i, \theta_{-i}, \omega)$.

We further impose the decision efficiency constraint. In two-stage mechanisms, the efficient decision rule must be based on the "expected" valuations conditional on the realized type profile.[4]

Definition 3. A two-stage mechanism (Θ, Π, x, t) satisfies *decision efficiency* (EFF) if, for all $(\theta_1, \theta_2) \in \Theta$,

$$x(\theta_1, \theta_2) = \begin{cases} 1 & \text{if } \tilde{u}_2^e(\theta_1, \theta_2) \geq \tilde{u}_1^e(\theta_1, \theta_2); \\ 0 & \text{otherwise,} \end{cases}$$

where $\tilde{u}_i^e(\theta_i, \theta_{-i}) = \int_{\Omega} \tilde{u}_i(\theta_i, \theta_{-i}; \omega) dH\left(\omega | (\theta_i, \theta_{-i})\right)$ is the expected valuation of agent i conditional on the type profile being (θ_i, θ_{-i}).

In what follows, we denote by x^* the efficient decision rule. In the definition of budget balance below, we require that the total "expected" transfer conditional on the realized type profile equal zero.[5]

Definition 4. A two-stage mechanism (Θ, Π, x, t) satisfies *budget balance* (BB) if, for all $(\theta_1, \theta_2) \in \Theta$,

$$\sum_{i=1}^{2} t_i^e(\theta_1, \theta_2; u_1, u_2) = 0,$$

where, for each agent $i \in \{1, 2\}$,

$$t_i^e(\theta_1, \theta_2; u_1, u_2) = \int_{\Omega} t_i\left(\theta_1, \theta_2; u_1(x(\theta_1, \theta_2), \theta_1, \theta_2, \omega), u_2(x(\theta_1, \theta_2), \theta_1, \theta_2, \omega)\right) dH\left(\omega | (\theta_1, \theta_2)\right)$$

is the expected monetary transfer of agent i conditional on the type profile (θ_1, θ_2) when both agents report their type truthfully in both stages.[6]

[4]Mezzetti (2004) also introduces the same decision efficiency constraint.

[5]Mezzetti (2004) uses the same budget balance constraint and says "When the state variable ω is a random function of the state profile, the budget will not be balanced for all realizations of ω, but it will be balanced (on average) for all type profiles."

[6]In fact, this budget balance can be strengthened to be achieved for all realizations of the residual uncertainty ω.

We maintain the following assumption imposed on two-stage mechanisms throughout the paper:

Assumption 0. Each agent's transfer is determined independently of their own payoff report in the second stage.[7]

As a result, the agents are indifferent between reporting their payoffs truthfully and lying in the second stage so that they have no incentive to deviate from the truth-telling in their payoff report in the second stage. Although Assumption 0 sounds a substantial restriction, it is necessary for our BIC constraint to make sense.[8] To see this, we suppose, on the contrary, that agent i's transfer depends on his own payoff report in the second stage. Fix the first-stage report $(\theta_i^r, \theta_{-i}^r) \in \Theta$ and the other agent's second-stage report $u_{-i}^r \in \Pi_{-i}(\theta_i^r, \theta_{-i}^r)$, and let u_i and u_i' be two distinct payoff reports of agent i in the second stage. Assume $t_i(\theta_i^r, \theta_{-i}^r; u_i, u_{-i}^r) > t_i(\theta_i^r, \theta_{-i}^r; u_i', u_{-i}^r)$ without loss of generality. Then, when agent i's true allocation payoff is u_i', he can increase his monetary transfer, thereby increase his utility, by deviating to u_i in the second stage, violating BIC.[9]

3 The Generalized Shoot-the-Liar Mechanism

In this paper, we explore the existence of two-stage mechanisms satisfying BIC, IIR, EFF and BB in bilateral trade environments. Even though Mezzetti (2004) establishes that the generalized two-stage Groves mechanism always satisfies BIC, EFF and BB in the general mechanism design models, Kunimoto and Zhang (2022) show by a stylized model of bilateral trade that the generalized two-stage Groves mechanism violates IIR. We thus seek more positive results by considering the *generalized shoot-the-liar* (henceforth, GS) mechanism.

We generalize the shoot-the-liar mechanism proposed in Mezzetti (2007) and apply it to our bilateral trade model. As the GS mechanism builds on the effi-

[7]The reader is referred to Jehiel and Moldovanu (2006), who critically examine this assumption and cast doubt on the real-world applicability of two-stage mechanisms.

[8]One may conjecture that Assumption 0 can be relaxed if we weakened our notion of BIC. This is indeed true in Nath and Zoeter (2013), who propose a two-stage mechanism in which (i) truth-telling in both stages is an ex post Nash equilibrium of the entire game, leading to a weaker BIC constraint; (ii) reporting the true payoff in the second stage is a strict best-response for each agent, implying that our Assumption 0 is not satisfied. However, since Assumption 0 is necessary for our BIC constraint, then truth-telling in both stages in their mechanism does not constitute part of a perfect Bayesian equilibrium, implying that our BIC constraint is violated. The reader is referred to Nath and Zoeter (2013) for the discussion of their mechanism.

[9]See the same argument in Mezzetti (2003, p18).

cient decision rule, we recall that the efficient decision rule dictates that, for each $(\theta_1, \theta_2) \in \Theta$,

$$x^*(\theta_1, \theta_2) = \begin{cases} 1 & \text{if } \tilde{u}_2^e(\theta_1, \theta_2) \geq \tilde{u}_1^e(\theta_1, \theta_2); \\ 0 & \text{otherwise,} \end{cases}$$

where $\tilde{u}_i^e(\theta_i, \theta_{-i}) = \int_\Omega \tilde{u}_i(\omega) dH\left(\omega | (\theta_i, \theta_{-i})\right)$ is the expected valuation of agent i conditional on the type profile being (θ_i, θ_{-i}).

Furthermore, for each agent $i \in \{1, 2\}$ and each type profile $\theta \in \Theta$, we denote by $\widetilde{\Pi}_i(\theta)$ the set of feasible allocation payoffs of agent i given the type profile θ, that is,

$$\widetilde{\Pi}_i(\theta) = \{\hat{u}_i \in \mathbb{R} : \exists \omega \in \Omega \text{ such that } h(\omega|\theta) > 0 \text{ and } \tilde{u}_i(\theta; \omega) = \hat{u}_i\}.$$

To define the generalized shoot-the-liar mechanism below, we introduce the following notation: $\Theta_2^{**} \equiv \{\theta_2 \in \Theta_2 : x^*(\theta_1, \theta_2) = 1 \text{ for all } \theta_1 \in \Theta_1\}$ and

$$\theta_2^* = \begin{cases} \inf \Theta_2^{**} & \text{if } \Theta_2^{**} \neq \emptyset; \\ \overline{\theta} & \text{if } \Theta_2^{**} = \emptyset. \end{cases}$$

Definition 5. A two-stage mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ is called the *generalized shoot-the-liar mechanism* if, for each type report $(\theta_1^r, \theta_2^r) \in \Theta$ and each payoff report $(u_1^r, u_2^r) \in \Pi_1 \times \Pi_2$,

$$t_1^{GS}(\theta_1^r, \theta_2^r; u_1^r, u_2^r) = \begin{cases} \tilde{u}_2^e(\theta_1^r, \theta_2^r) & \text{if } \theta_2^r \notin \Theta_2^{**},\ x^*(\theta_1^r, \theta_2^r) = 1, \text{ and } u_2^r \in \widetilde{\Pi}_2(\theta_1^r, \theta_2^r); \\ g(\theta_1^r) & \text{if } \theta_2^r \in \Theta_2^{**},\ x^*(\theta_1^r, \theta_2^r) = 1, \text{ and } u_2^r \in \widetilde{\Pi}_2(\theta_1^r, \theta_2^r); \\ -\psi & \text{if } x^*(\theta_1^r, \theta_2^r) = 1 \text{ and } u_2^r \notin \widetilde{\Pi}_2(\theta_1^r, \theta_2^r); \\ 0 & \text{if } x^*(\theta_1^r, \theta_2^r) = 0, \end{cases}$$

and

$$t_2^{GS}(\theta_1^r, \theta_2^r; u_1^r, u_2^r) = \begin{cases} -\tilde{u}_2^e(\theta_1^r, \theta_2^r) & \text{if } \theta_2^r \notin \Theta_2^{**} \text{ and } x^*(\theta_1^r, \theta_2^r) = 1; \\ -g(\theta_1^r) & \text{if } \theta_2^r \in \Theta_2^{**} \text{ and } x^*(\theta_1^r, \theta_2^r) = 1; \\ 0 & \text{if } x^*(\theta_1^r, \theta_2^r) = 0 \text{ and } u_1^r \in \widetilde{\Pi}_1(\theta_1^r, \theta_2^r); \\ -\psi & \text{if } x^*(\theta_1^r, \theta_2^r) = 0 \text{ and } u_1^r \notin \widetilde{\Pi}_1(\theta_1^r, \theta_2^r), \end{cases}$$

where ψ is a nonnegative constant (which is determined later), and

$$g(\theta_1^r) = \begin{cases} \tilde{u}_2^e(\theta_1^r, \theta_2^*) & \text{if } \int_{\Theta_2^{**}} dF(\theta_2) = 0; \\ G(\theta_1^r) / \int_{\Theta_2^{**}} dF(\theta_2) & \text{if } \int_{\Theta_2^{**}} dF(\theta_2) > 0, \end{cases}$$

with

$$
\begin{aligned}
G(\theta_1^r) \;=\; & -\int_{\Theta_2^*(\theta_1^r)\setminus\Theta_2^{**}} \tilde{u}_2^e(\theta_1^r,\theta_2)dF(\theta_2) + \int_{\Theta_2^*(\theta_1^r)} \tilde{u}_1^e(\theta_1^r,\theta_2)dF(\theta_2) \\
& + \int_{\Theta_1}\int_{\Theta_2^*(\theta_1)} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_1^e(\theta_1,\theta_2)\right) dF(\theta_2)dF(\theta_1) \\
& - \int_{\Theta_1}\int_{\Theta_2^{**}} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2^*)\right) dF(\theta_2)dF(\theta_1), \qquad (1)
\end{aligned}
$$

where for each $\theta_1^r \in \Theta_1$, $\Theta_2^*(\theta_1^r) = \{\theta_2 \in \Theta_2 : x^*(\theta_1^r,\theta_2) = 1\}$.

In the GS mechanism $(\Theta,\Pi,x^*,t^{GS},\psi)$, if each agent i reports their true type θ_i and true allocation payoff $u_i = u_i(x^*(\theta_i,\theta_{-i}),\omega)$, then the following three properties are satisfied.

1. $x^*(\theta_1,\theta_2) = 0$ implies $t_1^{GS}(\theta_1,\theta_2;u_1,u_2) = t_2^{GS}(\theta_1,\theta_2;u_1,u_2) = 0$, i.e., when no trade occurs, there are no monetary transfers;
2. $x^*(\theta_1,\theta_2) = 1$ and $\theta_2 \notin \Theta_2^{**}$ implies $t_1^{GS}(\theta_1,\theta_2;u_1,u_2) = -t_2^{GS}(\theta_1,\theta_2;u_1,u_2) = \tilde{u}_2^e(\theta_1,\theta_2)$, i.e., the seller extracts the full surplus; and
3. $x^*(\theta_1,\theta_2) = 1$ and $\theta_2 \in \Theta_2^{**}$ implies $t_1^{GS}(\theta_1,\theta_2;u_1,u_2) = -t_2^{GS}(\theta_1,\theta_2;u_1,u_2) = g(\theta_1)$ which is independent of the buyer's type. Since trade occurs with probability one whenever the buyer's reported type is in Θ_2^{**}, the buyer of type $\theta_2 \in \Theta_2^{**}$ has no incentive to deviate to type $\theta_2^r \in \Theta_2^{**}$ with $\theta_2^r \neq \theta_2$. Moreover, if $\int_{\Theta_2^{**}} dF(\theta_2) = 0$, then $g(\theta_1) = \tilde{u}_2^e(\theta_1,\theta_2^*)$. By the transfer rule under $\Theta_2\setminus\Theta_2^{**}$, this implies that the seller extracts the full surplus. As we will show in the proof of Claim 1, if $\int_{\Theta_2^{**}} dF(\theta_2) > 0$, the $g(\cdot)$ function is designed in such a way that the buyer of type $\theta_2 \in \Theta_2^{**}$ is left with positive expected surplus.[10]

The GS mechanism satisfies EFF and BB by its construction. Indeed, since each agent's monetary transfer is independent of the residual uncertainty ω, then budget balance is achieved for each $\omega \in \Omega$, or equivalently, budget balance is achieved at the ex post stage. On the other hand, the GS mechanism allows us to

[10]In contrast, the seller always extracts the full surplus in the shoot-the-liar mechanism proposed by Mezzetti (2007). Moreover, Mezzetti (2007) considers the auction setup and in the shoot-the-liar mechanism he proposes, the seller plays the role of an outsider whose valuation is normalized to zero and the seller makes no monetary transfer other than collecting payments from the buyers. In the GS mechanism we propose, however, the seller has private information which should be elicited within the mechanism and she is asked to make monetary transfers based on the reports.

break the budget off the equilibrium path in the following manner: if trade does not occur and the seller's payoff report is not consistent with the first-stage type reports, then the buyer is punished with the penalty ψ. Similarly, if trade occurs and the buyer's payoff report is not consistent with the first-stage type reports, then the seller is punished with the penalty ψ.

We know from the definition of the GS mechanism that there are two cases: $\int_{\Theta_2^{**}} dF(\theta_2) = 0$ and $\int_{\Theta_2^{**}} dF(\theta_2) > 0$. We illustrate the working of the $g(\cdot)$ function in the GS mechanism in the following claim.

Claim 1. When $\int_{\Theta_2^{**}} dF(\theta_2) > 0$, the $g(\cdot)$ function works as follows:

1. The seller's IIR constraint is satisfied if, for each $\theta_1 \in \Theta_1$, there exist a nonnegative constant C such that

$$\int_{\Theta_2^{**}} g(\theta_1) dF(\theta_2) \quad = \quad -\int_{\Theta_2^*(\theta_1)\setminus\Theta_2^{**}} \tilde{u}_2^e(\theta_1, \theta_2) dF(\theta_2) + \int_{\Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2) + C, \tag{2}$$

   where $\tilde{u}_i^e(\theta_i, \theta_{-i}) = \int_\Omega \tilde{u}_i(\theta_i, \theta_{-i}; \omega) dH(\omega|(\theta_i, \theta_{-i}))$; $\Theta_2^*(\theta_1) = \{\theta_2 \in \Theta_2 : x^*(\theta_1, \theta_2) = 1\}$; and $\Theta_2^{**} = \{\theta_2 \in \Theta_2 : x^*(\theta_1, \theta_2) = 1 \text{ for all } \theta_1 \in \Theta_1\}$.

2. The IIR constraint for the buyer of type $\theta_2 \in \Theta_2^{**}$ is satisfied if the constant C equals the following:

$$\begin{aligned} C \quad = \quad & \int_{\Theta_1} \int_{\Theta_2^*(\theta_1)} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_1^e(\theta_1, \theta_2)\right) dF(\theta_2) dF(\theta_1) \\ & - \int_{\Theta_1} \int_{\Theta_2^{**}} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^*)\right) dF(\theta_2) dF(\theta_1), \end{aligned} \tag{3}$$

   where $\theta_2^* = \inf \Theta_2^{**}$ in this case. Finally, combining (2) and (3), we obtain the expression of $g(\cdot)$ in the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$. Assumption 2, which will be introduced in Section 5.1, is equivalent to $C \geq 0$.

Proof. The proof is in the Appendix. □

We further show that the GS mechanism satisfies the following properties.

Proposition 1. We obtain the following result:

1. When $\int_{\Theta_2^{**}} dF(\theta_2) = 0$, the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ always satisfies EFF, BB, and IIR.[11]

2. When $\int_{\Theta_2^{**}} dF(\theta_2) > 0$, the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ always satisfies EFF, BB, and IIR only for the buyer.

Proof. Given that both EFF and BB are already built in the GS mechanism, it remains to verify that IIR is satisfied when $\int_{\Theta_2^{**}} dF(\theta_2) = 0$ and that buyer's IIR constraint is satisfied when $\int_{\Theta_2^{**}} dF(\theta_2) > 0$. The rest of the proof is in the Appendix. $\square$

4 When the GS Mechanism satisfies BIC

Whether BIC is satisfied in the GS mechanism depends on whether all the possible lies in the first stage can be detected and appropriately detered. This exhibits a contrast with the generalized two-stage Groves mechanism, which satisfies BIC automatically. In the GS mechanism, the designer detects a lie if, whenever agent i deviates in the first stage, the good is allocated to agent $j \neq i$ and agent j's reported allocation payoff is not consistent with the first-stage type reports. To guarantee this detectability, we introduce Assumption 1, which turns out to be necessary and sufficient for the GS mechanism to satisfy BIC.

4.1 Assumption 1

Consider the GS mechanism with $\psi = 0$ where the designer imposes zero penalty off the equilibrium path. We consider such a mechanism with the following reasons: (i) within the class of the GS mechanisms with $\psi \geq 0$, the one with $\psi = 0$ is most fragile in the sense that it has the maximum number of profitable deviations; (ii) if the GS mechanism with $\psi > 0$ is able to detect all the profitable deviations in the mechanism with $\psi = 0$, then it satisfies BIC; and (iii) if the GS mechanism with $\psi > 0$ satisfies BIC, it is able to detect all the profitable deviations in the mechanism with $\psi = 0$.

To formally define Assumption 1, we develop some useful notation. Since each agent's transfer is independent of their own second-stage report, each agent has

[11] Suppose that in the deterministic model, each agent's ex post valuation, which is a deterministic function of the type profile, varies as the type profile varies. Then, the GS mechanism satisfies BIC. Together with Proposition 1, we conclude that the GS mechanism satisfies BIC, IIR, EFF and BB when $\int_{\Theta_2^{**}} dF(\theta_2) = 0$ of the deterministic model. Kunimoto and Zhang (2022) established the same result in a stylized example.

no incentive to deviate in the second stage. For each agent $i \in \{1,2\}$ and each $\theta_i, \theta_i^r \in \Theta_i$, we denote by $U_i^{\psi_0}(\theta_i;\theta_i^r)$ the expected utility of agent i of type θ_i when agent i reports θ_i^r and agent $j \neq i$ reports their type truthfully in the first stage in the GS mechanism with $\psi = 0$. Recall that $\widetilde{\Pi}_j(\theta_i^r, \theta_j)$ denotes the set of feasible allocation payoffs of agent $j \neq i$ given the reported type profile (θ_i^r, θ_j), that is,

$$\widetilde{\Pi}_j(\theta_i^r, \theta_j) = \{\hat{u}_j \in \mathbb{R} : \text{ there exists } \omega \in \Omega \text{ s.t. } h\left(\omega|(\theta_i^r, \theta_j)\right) > 0 \text{ and } \tilde{u}_j(\theta_i^r, \theta_j; \omega) = \hat{u}_j\}.$$

We also denote by $\Omega^{\psi}(\theta_i, \theta_j; \theta_i^r)$ the set of residual uncertainties in which agent j's allocation payoff is consistent with the true type profile (θ_i, θ_j), but it is inconsistent with the reported type profile (θ_i^r, θ_j):

$$\Omega^{\psi}(\theta_i, \theta_j; \theta_i^r) = \left\{\omega \in \Omega : \tilde{u}_j(\theta_i, \theta_j; \omega) \in \widetilde{\Pi}_j(\theta_i, \theta_j) \text{ but } \tilde{u}_j(\theta_i, \theta_j; \omega) \notin \widetilde{\Pi}_j(\theta_i^r, \theta_j)\right\}.$$

We denote by $x_i^*(\theta)$ the probability that the good is allocated to agent i when θ is the reported type profile.

Assumption 1. For each agent $i \in \{1,2\}$ and each pair of types $\theta_i, \theta_i^r \in \Theta_i$, if $U_i^{\psi_0}(\theta_i;\theta_i^r) > U_i^{\psi_0}(\theta_i;\theta_i)$, then there exists a positive-measure set $\Theta_j^+ \subseteq \Theta_j$ where $j \neq i$ such that (i) $\int_{\Omega^{\psi}(\theta_i,\theta_j;\theta_i^r)} dH\left(\omega|(\theta_i,\theta_j)\right) > 0$ for all $\theta_j \in \Theta_j^+$, and (ii) $\int_{\Theta_j^+} x_j^*(\theta_i^r, \theta_j) dF(\theta_j) > 0$.

Our Assumption 1 adapts Mezzetti's (2007) Assumption 1 from the model of auctions to that of bilateral trade and further extends it from the deterministic model to the random model.[12]

Condition (i) in our Assumption 1 implies that $\widetilde{\Pi}_j(\theta_i, \theta_j) \neq \widetilde{\Pi}_j(\theta_i^r, \theta_j)$. In particular, if $\widetilde{\Pi}_j(\theta_i, \theta_j) \cap \widetilde{\Pi}_j(\theta_i^r, \theta_j) = \emptyset$, then we have $\Omega^{\psi}(\theta_i, \theta_j; \theta_i^r) = \Omega$. This further implies $\int_{\Omega^{\psi}(\theta_i,\theta_j;\theta_i^r)} dH\left(\omega|(\theta_i,\theta_j)\right) = 1$. Hence, Condition (i) is automatically satisfied. On the other hand, if $\widetilde{\Pi}_j(\theta_i, \theta_j) \cap \widetilde{\Pi}_j(\theta_i^r, \theta_j) \neq \emptyset$, then Condition (i) requires that the set of residual uncertainties leading to $\widetilde{\Pi}_j(\theta_i, \theta_j) \backslash \widetilde{\Pi}_j(\theta_i^r, \theta_j)$ must have nonzero measure in Ω.

To show the lemma below, we introduce the following categories. By the *trivial* case, we mean that it is always efficient to trade. We call any other case a *nontrivial* case.

[12] Mezzetti's Assumption 1 requires that if an agent i of type θ_i has an incentive to deviate to θ_i^r in the shoot-the-liar mechanism with zero penalty, then there exists a positive-measure set $\Theta_{-i}^+ \subseteq \Theta_{-i}$ and an agent $j \neq i$ such that (i) $\tilde{u}_j^e(\theta_i, \theta_{-i}) \neq \tilde{u}_j^e(\theta_i^r, \theta_{-i})$ for all $\theta_{-i} \in \Theta_{-i}^+$, and (ii) $\int_{\Theta_{-i}^+} x_j^*(\theta_i^r, \theta_{-i}) dF_{-i}(\theta_{-i}) > 0$.

Lemma 2. In all nontrivial cases, if Assumption 1 is satisfied and $\Theta_2^{**} \neq \emptyset$, then $\theta_2^* = \inf \Theta_2^{**}$ is the unique cutoff point such that

$$\int_{\Theta_1} x^*(\theta_1, \theta_2) dF(\theta_1) \begin{cases} < 1 & \text{if } \theta_2 < \theta_2^*; \\ = 1 & \text{if } \theta_2 \geq \theta_2^*. \end{cases}$$

Proof. Consider the buyer of type $\theta_2 \notin \Theta_2^{**}$. The buyer's expected utility under truth-telling is

$$\int_{\Theta_1^*(\theta_2)} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2)\right) dF(\theta_1) = 0,$$

where $\Theta_1^*(\theta_2) = \{\theta_1 \in \Theta_1 : x^*(\theta_1, \theta_2) = 1\}$. Note that the expected utility of the buyer under truth-telling is identical to that in the GS mechanism with $\psi = 0$, which is $U_2^{\psi_0}(\theta_2; \theta_2)$. On the other hand, if he deviates to $\theta_2^r \in \Theta_2^{**}$, his expected utility becomes

$$\int_{\Theta_1} \left(\tilde{u}_2^e(\theta_1, \theta_2) - g(\theta_1)\right) dF(\theta_1) = \int_{\Theta_1} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^*)\right) dF(\theta_1),$$

where the equality follows from the proof of Claim 1 that $\int_{\Theta_1} g(\theta_1) dF(\theta_1) = \int_{\Theta_1} \tilde{u}_2^e(\theta_1, \theta_2^*) dF(\theta_1)$. Note that the expected utility of the buyer after deviation is identical to that in the GS mechanism with $\psi = 0$, which is $U_2^{\psi_0}(\theta_2; \theta_2^r)$. Moreover, note that after the buyer deviates to $\theta_2^r \in \Theta_2^{**}$, the seller is never allocated the good. Then, according to Assumption 1, $U_2^{\psi_0}(\theta_2; \theta_2) \geq U_2^{\psi_0}(\theta_2; \theta_2^r)$ must hold, which is equivalent to

$$0 \geq \int_{\Theta_1} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^*)\right) dF(\theta_1).$$

Since the buyer's expected valuation $\tilde{u}_2^e(\cdot)$ is strictly increasing in his own type, $\theta_2 \leq \theta_2^*$ must be satisfied. Moreover, since θ_2 is an arbitrary type within the set $\Theta_2 \backslash \Theta_2^{**}$ and $\theta_2^* = \inf \Theta_2^{**}$, we obtain that θ_2^* is the unique cutoff such that $\int_{\Theta_1} x^*(\theta_1, \theta_2) dF(\theta_1) = 1$ whenever $\theta_2 > \theta_2^*$ and $\int_{\Theta_1} x^*(\theta_1, \theta_2) dF(\theta_1) < 1$ when $\theta_2 < \theta_2^*$.

It remains to verify that Θ_2^{**} is a closed set, which implies $\theta_2^* \in \Theta_2^{**}$. If $\Theta_2^{**} \neq \emptyset$ but $\int_{\Theta_2^{**}} dF(\theta_2) = 0$, then Θ_2^{**} is a singleton so that it is trivially closed. So, we assume $\int_{\Theta_2^{**}} dF(\theta_2) > 0$. Choose a sequence $(\theta_2^n)_{n \in \mathbb{N}} \subseteq \Theta_2$ such that $\theta_2^n \in \Theta_2^{**}$, i.e., $\theta_2^n \geq \theta_2^*$ for each $n \in \mathbb{N}$ and $\theta_2^n \to \theta_2^*$ as $n \to \infty$. Choose $\theta_1 \in \Theta_1$ arbitrarily. Since $\theta_2^n \in \Theta_2^{**}$ for each $n \in \mathbb{N}$, we have

$$\tilde{u}_2^e(\theta_1, \theta_2^n) \geq \tilde{u}_1^e(\theta_1, \theta_2^n), \ \forall n.$$

Since we assume that each $\tilde{u}_i^e(\cdot)$ is differentiable in both θ_i and θ_{-i}, which implies that they are also continuous in both θ_i and θ_{-i}, we have

$$\tilde{u}_2^e(\theta_1, \theta_2^*) \geq \tilde{u}_1^e(\theta_1, \theta_2^*).$$

Since θ_1 is chosen arbitrarily, this implies that $x^*(\theta_1, \theta_2^*) = 1$ for all $\theta_1 \in \Theta_1$. Thus, $\theta_2^* \in \Theta_2^{**}$. This completes the proof. $\square$

4.2 The Role of Assumption 1 in the GS Mechanism

We shall show that Assumption 1 is a necessary and sufficient condition for the GS mechanism to satisfy BIC.

Proposition 2. There exists a threshold $\psi^* > 0$ such that the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ with $\psi \geq \psi^*$ satisfies BIC if and only if Assumption 1 is satisfied.

Proof. We first prove the necessity. Suppose on the contrary that Assumption 1 is violated. Then, there exists an agent i and types θ_i and θ_i^r such that (i) $U_i^{\psi_0}(\theta_i; \theta_i^r) > U_i^{\psi_0}(\theta_i; \theta_i)$, and (ii) for all subsets $\Theta_j^+ \subseteq \Theta_j$ with positive measure where $j \neq i$, either $\int_{\Omega^{\psi}(\theta_i, \theta_j; \theta_i^r)} dH\left(\omega | (\theta_i, \theta_j)\right) = 0$ for all $\theta_j \in \Theta_j^+$, or $\int_{\Theta_j^+} x_j^*(\theta_i^r, \theta_j) dF(\theta_j) = 0$. If $\int_{\Omega^{\psi}(\theta_i, \theta_j; \theta_i^r)} dH\left(\omega | (\theta_i, \theta_j)\right) = 0$ for all $\theta_j \in \Theta_j^+$, then all allocation payoffs which are feasible given the true type profile (θ_i, θ_j) are also feasible given the deviation (θ_i^r, θ_j). According to the transfer rule in the GS mechanism, agent i is not punished. Moreover, if $\int_{\Theta_j^+} x_j^*(\theta_i^r, \theta_j) dF(\theta_j) = 0$, then the good is not allocated to agent $j \neq i$ and agent i is not punished, either. As a result, agent i's expected utility after deviating to θ_i^r is the same as $U_i^{\psi_0}(\theta_i; \theta_i^r)$, which is higher than that under truth-telling. We conclude that if Assumption 1 is violated, the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ violates BIC for any $\psi \geq 0$. This completes the proof of the necessity part.

We next prove the sufficiency of Assumption 1. If Assumption 1 is satisfied, according to Lemma 2, θ_2^* is the unique cutoff type of the buyer where the expected trading probabilities differ for types above and below it. Recall that we set $\theta_2^* = \overline{\theta}$ if $\int_{\Theta_1} x^*(\theta_1, \theta_2) dF(\theta_1) < 1$ for all $\theta_2 \in (\underline{\theta}, \overline{\theta})$. Therefore, if Assumption 1 is satisfied, we can divide the analysis into two cases: Case 1: $\theta_2^* = \overline{\theta}$ and Case 2: $\theta_2^* \in (\underline{\theta}, \overline{\theta})$. To have a better understanding of the two cases, we also provide two figures illustrating the allocation decision at different type profiles. The shaded region represents $\Theta_* = \{(\theta_1, \theta_2) \in \Theta : x^*(\theta_1, \theta_2) = 1\}$, which describes the set of possible type profiles for which it is efficient to trade. In Figure 1, we have $\int_{\Theta_1} x^*(\theta_1, \theta_2) dF(\theta_1) < 1$ for all $\theta_2 < \overline{\theta}$. In Figure 2, it is always efficient to trade when θ_2 is greater than the cutoff type θ_2^*. It remains to show that every possible deviation at the first stage can be detected and punished. We show this by considering Cases 1 and 2, respectively. The rest of the sufficiency proof is in the Appendix.

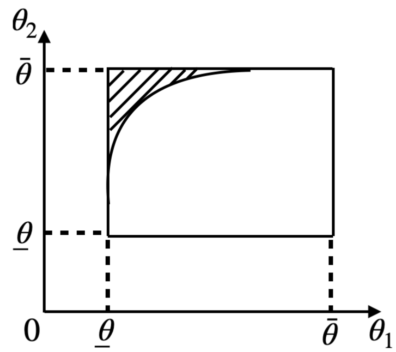


Figure 1

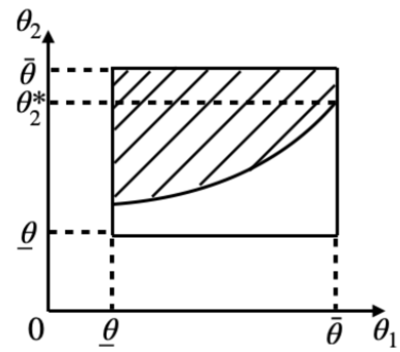


Figure 2

$\square$

4.3 The Canonicality of the GS Mechanism

We first show that if the GS mechanism satisfies BIC, it maximizes the seller's ex ante expected utility within the class of all two-stage mechanisms satisfying BIC, EFF, BB as well as buyer's IIR. Recall that $t_i^e(\theta_i, \theta_{-i}; u_i, u_{-i})$ is the expected monetary transfer of agent i conditional on the type profile being (θ_i, θ_{-i}) when both agents report truthfully in both stages, which is given by

$$\int_\Omega t_i\left(\theta_i, \theta_{-i}; u_i(x^*(\theta_i, \theta_{-i}), \theta_i, \theta_{-i}, \omega), u_{-i}(x^*(\theta_i, \theta_{-i}), \theta_i, \theta_{-i}, \omega)\right) dH\left(\omega | (\theta_i, \theta_{-i})\right).$$

Proposition 3. Suppose that there exists a threshold $\psi^* > 0$ such that the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ with $\psi \geq \psi^*$ satisfies BIC. Then, it maximizes the seller's ex ante expected utility among all two-stage mechanisms satisfying BIC, EFF, BB as well as the buyer's IIR constraint.

Proof. If $\Theta_2^{**} \neq \emptyset$, then it follows from Proposition 2 and Lemma 2 that $\theta_2^* = \inf \Theta_2^{**}$ is the unique cutoff point such that

$$\int_{\Theta_1} x^*(\theta_1, \theta_2) dF(\theta_1) \begin{cases} < 1 & \text{if } \theta_2 < \theta_2^*; \\ = 1 & \text{if } \theta_2 \geq \theta_2^*. \end{cases}$$

When we compute the seller's ex ante expected utility, we divide the buyer's types into two categories: $\theta_2 < \theta_2^*$ and $\theta_2 \geq \theta_2^*$. The proof is completed by the following four steps.

Step I: In any two-stage mechanism $(\Theta, \Pi, x^*, t, \psi)$ satisfying buyer's IIR, EFF and BB, the following inequality holds for the buyer of type $\theta_2 < \theta_2^*$: for all $\theta_2 < \theta_2^*$,

$$\int_{\Theta_1} t_1^e(\theta_1, \theta_2; u_1, u_2) dF(\theta_1) \quad \leq \quad \int_{\Theta_1^*(\theta_2)} \tilde{u}_2^e(\theta_1, \theta_2) dF(\theta_1),$$

where $\tilde{u}_2^e(\theta_1, \theta_2) = \int_\Omega \tilde{u}_2(\theta_1, \theta_2; \omega) dH(\omega|(\theta_1, \theta_2))$ and $\Theta_1^*(\theta_2) = \{\theta_1 \in \Theta_1 : x^*(\theta_1, \theta_2) = 1\}$.

Proof. Consider the buyer with type $\theta_2 < \theta_2^*$ in a two-stage mechanism $(\Theta, \Pi, x^*, t, \psi)$. Suppose that both agents report truthfully in both stages. Then, the buyer receives the following expected utility after participating in the two-stage mechanism $(\Theta, \Pi, x^*, t, \psi)$:

$$\int_{\Theta_1^*(\theta_2)} \tilde{u}_2^e(\theta_1, \theta_2) dF(\theta_1) + \int_{\Theta_1} t_2^e(\theta_1, \theta_2; u_1, u_2) dF(\theta_1).$$

Recall that the buyer's outside option utility is always zero. Then, the buyer's IIR constraint is as follows:

$$\begin{aligned} & \int_{\Theta_1^*(\theta_2)} \tilde{u}_2^e(\theta_1, \theta_2) dF(\theta_1) + \int_{\Theta_1} t_2^e(\theta_1, \theta_2; u_1, u_2) dF(\theta_1) \geq 0 \\ \Rightarrow \quad & \int_{\Theta_1} t_2^e(\theta_1, \theta_2; u_1, u_2) dF(\theta_1) \geq - \int_{\Theta_1^*(\theta_2)} \tilde{u}_2^e(\theta_1, \theta_2) dF(\theta_1). \end{aligned}$$

Since the two-stage mechanism $(\Theta, \Pi, x^*, t, \psi)$ satisfies BB, for each $(\theta_1, \theta_2) \in \Theta$, we have

$$t_2^e(\theta_1, \theta_2; u_1, u_2) \quad = \quad -t_1^e(\theta_1, \theta_2; u_1, u_2).$$

Therefore, the buyer's IIR constraint can be further rewritten as follows:

$$\int_{\Theta_1} t_1^e(\theta_1, \theta_2; u_1, u_2) dF(\theta_1) \leq \int_{\Theta_1^*(\theta_2)} \tilde{u}_2^e(\theta_1, \theta_2) dF(\theta_1).$$

This completes the proof of Step I. □

Step II: In any two-stage mechanism $(\Theta, \Pi, x^*, t, \psi)$ satisfying BIC, EFF, BB and the buyer's IIR, there exists $\hat{g} : \Theta_1 \to \mathbb{R}$ such that for all $\theta_1 \in \Theta_1$ and all $\theta_2 \geq \theta_2^*$,

$$t_1^e(\theta_1, \theta_2; u_1, u_2) = -t_2^e(\theta_1, \theta_2; u_1, u_2) = \hat{g}(\theta_1),$$

and

$$\int_{\Theta_1} \hat{g}(\theta_1) dF(\theta_1) \leq \int_{\Theta_1} \tilde{u}_2^e(\theta_1, \theta_2^*) dF(\theta_1).$$

Proof. The proof is in the Appendix. $\square$

Step III: In any two-stage mechanism satisfying BIC, EFF, BB as well as the buyer's IIR, the seller's ex ante expected utility is at most as high as the following:

$$\int_{\Theta_1}\int_{\Theta_2} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2)dF(\theta_1) + \int_{\Theta_1}\int_{\Theta_2^*(\theta_1)} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_1^e(\theta_1,\theta_2)\right)dF(\theta_2)dF(\theta_1)$$
$$- \int_{\Theta_1}\int_{\Theta_2^{**}} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2^*)\right)dF(\theta_2)dF(\theta_1), \tag{4}$$

where $\tilde{u}_i^e(\theta_i,\theta_{-i}) = \int_\Omega \tilde{u}_i(\theta_i,\theta_{-i};\omega)dH\left(\omega|(\theta_i,\theta_{-i})\right)$ for each agent $i \in \{1,2\}$, $\Theta_2^*(\theta_1) = \{\theta_2 \in \Theta_2 : x^*(\theta_1,\theta_2) = 1\}$ for each $\theta_1 \in \Theta_1$, and $\Theta_2^{**} = [\theta_2^*, \overline{\theta}]$.

Proof. The proof is in the Appendix. $\square$

Step IV: The seller's ex ante expected utility in the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ is identical to the upper bound in (4).

Proof. The proof is in the Appendix. $\square$

Therefore, we conclude that the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ maximizes the seller's ex ante expected utility among all two-stage mechanisms satisfying BIC, EFF, BB as well as the buyer's IIR constraint. This completes the proof of Proposition 3. $\square$

The result below shows that, under Assumption 1, there is no loss of generality in focusing on the GS mechanism for finding two-stage mechanisms satisfying BIC, IIR, EFF, and BB.

Theorem 1. Suppose that Assumption 1 holds. If there exists a two-stage mechanism that satisfies BIC, EFF, BB, and IIR, then the GS mechanism satisfies the same properties.

Proof. We know from Proposition 1 that the GS mechanism always satisfies EFF, BB as well as the buyer's IIR. It then suffices to verify that the GS mechanism satisfies BIC and seller's IIR constraints. Since Assumption 1 holds, it follows from Proposition 2 that there exists a threshold $\psi^* > 0$ such that the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ with $\psi \geq \psi^*$ satisfies BIC. Moreover, by our hypothesis that there exists a two-stage mechanism satisfies IIR, that mechanism must satisfy the ex ante individual rationality as well. This together with Proposition 3 further implies that the GS mechanism must also satisfy the seller's ex ante individual rationality.

If both agents report truthfully in both stages, the proof of Step IV of Proposition 3 allows us to obtain the following ex ante expected utility for the seller after participating in the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$:

$$\int_{\Theta_1}\int_{\Theta_2} U_1^O(\theta_1)dF(\theta_2)dF(\theta_1) + \int_{\Theta_1}\int_{\Theta_2^*(\theta_1)} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_1^e(\theta_1,\theta_2)\right) dF(\theta_2)dF(\theta_1)$$
$$-\int_{\Theta_1}\int_{\Theta_2^{**}} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2^*)\right) dF(\theta_2)dF(\theta_1),$$

where $U_1^O(\theta_1)$ denotes the interim outside option utility for the seller of type θ_1. Ex ante individual rationality requires that the following expression be nonnegative:

$$\int_{\Theta_1}\int_{\Theta_2^*(\theta_1)} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_1^e(\theta_1,\theta_2)\right) dF(\theta_2)dF(\theta_1) - \int_{\Theta_1}\int_{\Theta_2^{**}} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2^*)\right) dF(\theta_2)dF(\theta_1). \quad (5)$$

Finally, recall in the proof of Claim 1 that the interim expected utility of the seller of type θ_1 after participating in the GS mechanism is

$$\begin{aligned} U_1^{GS}(\theta_1) &= U_1^O(\theta_1) + \int_{\Theta_1}\int_{\Theta_2^*(\theta_1)} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_1^e(\theta_1,\theta_2)\right) dF(\theta_2)dF(\theta_1) \\ &\quad - \int_{\Theta_1}\int_{\Theta_2^{**}} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2^*)\right) dF(\theta_2)dF(\theta_1). \end{aligned}$$

Since expression (5) is nonnegative, we conclude that the seller's interim expected utility after participating in the GS mechanism is at least as high as that from her outside option, implying that the seller's IIR constraint is satisfied in the GS mechanism. This completes the proof of the theorem. □

4.4 Assumption 1 is Satisfied in Example 1

Recall in Example 1 that the residual uncertainty ω is expressed as (ω_1, ω_2) where conditional on the type profile, ω_1 and ω_2 are independently drawn from a uniform distribution in $[-0.1, 0.1]$, and that each agent i's valuation function is $\tilde{u}_i(\theta_i, \theta_{-i}; \omega) = (1+\omega_i)(\theta_i + \gamma_i\theta_{-i})$. Then, each agent i's expected valuation conditional on the type profile being (θ_i, θ_{-i}) is given as follows:

$$\tilde{u}_i^e(\theta_i,\theta_{-i}) = \int_\Omega \tilde{u}_i(\theta_i,\theta_{-i};\omega)dH\left(\omega|(\theta_i,\theta_{-i})\right) = \int_{-0.1}^{0.1} (1+\omega_i)(\theta_i+\gamma_i\theta_{-i})dH_i\left(\omega_i|(\theta_i,\theta_{-i})\right) = \theta_i+\gamma_i\theta_{-i},$$

where $H_i\left(\omega_i|(\theta_i,\theta_{-i})\right)$ denotes the conditional cumulative distribution function of ω_i, which is the uniform distribution. Hence, for each type profile $(\theta_1, \theta_2) \in \Theta$, we have

$$\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_1^e(\theta_1,\theta_2) = (\theta_2 + \gamma_2\theta_1) - (\theta_1 + \gamma_1\theta_2) = (1-\gamma_1)\theta_2 - (1-\gamma_2)\theta_1,$$

implying that the efficient decision rule depends on the values of γ_1 and γ_2. Since $\gamma_1 < 1$ and $\gamma_2 < 1$, we are left with two cases to consider: (i) $0 < \gamma_2 \leq \gamma_1 < 1$ and (ii) $0 < \gamma_1 < \gamma_2 < 1$. The following two figures (Figures 3 and 4) illustrate the decision at different type profile in the two cases, respectively; in particular, the shaded region in each figure represents $\Theta_* = \{(\theta_1, \theta_2) \in [0,1]^2 : x^*(\theta_1, \theta_2) = 1\}$ in each case.

Figure 3: when $0 < \gamma_2 \leq \gamma_1 < 1$

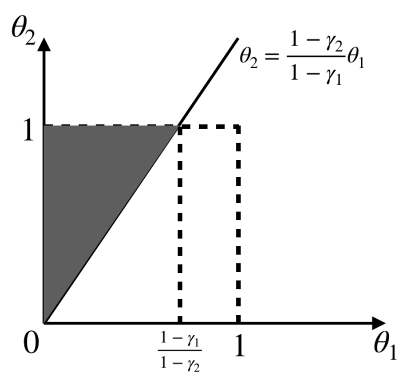
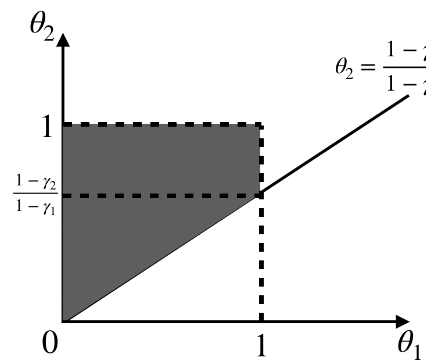


Figure 4: when $0 < \gamma_1 < \gamma_2 < 1$



It can easily be seen from Figures 3 that $\int_{\Theta_1} x^*(\theta_1, \theta_2) dF(\theta_1) < 1$ for all $\theta_2 < \overline{\theta}$. In this case, we define $\theta_2^* = \overline{\theta}$. In Figure 4, we have $\int_{\Theta_1} x^*(\theta_1, \theta_2) dF(\theta_1) < 1$ for all $\theta_2 < (1-\gamma_2)/(1-\gamma_1)$ and $\int_{\Theta_1} x^*(\theta_1, \theta_2) dF(\theta_1) = 1$ otherwise. Hence, $\theta_2^* = \inf \Theta_2^{**} = (1-\gamma_2)/(1-\gamma_1) \in (\underline{\theta}, \overline{\theta})$ in Figure 4. We confirm the following.

Claim 2. Assumption 1 is satisfied in Example 1.

Proof. The proof is in the Appendix. □

4.5 When Assumption 1 is Violated

There are some scenarios in which Assumption 1 may well be violated. We provide two examples. In the first example, there is a large degree of additional uncertainty in which (i) agent $i \in \{1,2\}$ of type θ_i deviates to θ_i^r in the first stage, resulting in a profitable deviation in the GS mechanism with $\psi = 0$ (i.e., $U_i^{\psi_0}(\theta_i; \theta_i^r) > U_i^{\psi_0}(\theta_i; \theta_i)$), and (ii) every feasible allocation payoff of the other agent $j \neq i$ given the true type profile (θ_i, θ_j) is also feasible given the deviation (θ_i^r, θ_j) (i.e., $\widetilde{\Pi}_j(\theta_i, \theta_j) = \widetilde{\Pi}_j(\theta_i^r, \theta_j)$). In other words, Condition (i) in Assumption 1 is violated. We proceed to the details.

Example 2. Suppose that each agents' type θ_i is drawn from the uniform distribution on the unit interval $[0,1]$. The residual uncertainty ω is also drawn from the unit interval $[0,1]$. Let $h(\cdot|\theta)$ be the probability density function defined over $[0,1]$ conditional on the type profile θ: for each $\theta \in \Theta$, (i) if $\theta_2 < 0.5$, then $h(\omega|\theta) = 1$ for all $\omega \in [0,1]$; (ii) if $\theta_2 \geq 0.5$, for any $\omega \in [0,1]$,

$$h(\omega|\theta) = \begin{cases} 0.9 & \text{if } \omega < 0.5, \\ 1.1 & \text{if } \omega \geq 0.5. \end{cases}$$

Then, we define $\Omega = [0,1]$ as the set of residual uncertainties. Each agent's valuation for the object in the type profile θ and residual uncertainty ω is given as follows:

$$\begin{aligned} \tilde{u}_1(\theta;\omega) &= \omega; \\ \tilde{u}_2(\theta;\omega) &= \theta_2 - \theta_1. \end{aligned}$$

The buyer's expected valuation is the same across different realizations of ω, while the seller's expected valuation varies: for any $\theta \in \Theta$,

$$\tilde{u}_1^e(\theta) = \begin{cases} \int_0^1 \omega d\omega = 0.5 & \text{if } \theta_2 < 0.5; \\ \\ 0.9\int_0^{0.5} \omega d\omega + 1.1\int_{0.5}^1 \omega d\omega = 0.525 & \text{if } \theta_2 \geq 0.5, \end{cases}$$

where we consider

$$0.9\int_0^{0.5} \omega d\omega + 1.1\int_{0.5}^1 \omega d\omega = \int_0^1 \omega d\omega + 0.1\left(\int_{0.5}^1 \omega d\omega - \int_0^{0.5} \omega d\omega\right) = 0.525.$$

It is efficient to trade if and only if $\tilde{u}_2^e(\theta_1,\theta_2) \geq \tilde{u}_1^e(\theta_1,\theta_2)$, which is equivalent to $\theta_2 \geq \theta_1 + 0.525$ in this example. Hence, we have that $\theta_2^* = \overline{\theta}$ in this example so that it is always efficient not to trade if $\theta_2 \leq 0.525$.

Claim 3. Condition (i) of Assumption 1 is violated in Example 2.

Proof. We will show that the buyer of type $\theta_2 > 0.525$ has an incentive to deviate to $\theta_2^r \in (0.525, \theta_2)$, but condition (i) of Assumption 1 is violated.

We take two types of the buyer $\theta_2', \theta_2'' \in \Theta_2$ such that $\theta_2'' > \theta_2' > 0.525$. Observe that for each $\theta_1 \in \Theta_1$, $\tilde{u}_2^e(\theta_1,\theta_2'') - \tilde{u}_2^e(\theta_1,\theta_2') = (\theta_2'' - \theta_1) - (\theta_2' - \theta_1) = \theta_2'' - \theta_2' > 0$, implying that the buyer's expected valuation is strictly increasing in θ_2. Moreover, if the buyer's true type is θ_2'' and both agents report truthfully in both stages, the buyer receives the following expected utility in the GS mechanism with $\psi = 0$:

$$\int_{\Theta_1^*(\theta_2'')} \left(\tilde{u}_2^e(\theta_1,\theta_2'') - \tilde{u}_2^e(\theta_1,\theta_2'')\right) dF(\theta_1) = 0,$$

where $\Theta_1^*(\theta_2'') = \{\theta_1 \in \Theta_1 : x^*(\theta_1, \theta_2'') = 1\}$. Note that the buyer's expected utility under truth-telling is identical to that in the GS mechanism with $\psi = 0$, which is $U_2^{\psi_0}(\theta_2''; \theta_2'')$. However, if the buyer of type θ_2'' deviates to report θ_2' in the first stage, his expected utility in the GS mechanism with $\psi = 0$ becomes

$$\int_{\Theta_1^*(\theta_2')} (\tilde{u}_2^e(\theta_1, \theta_2'') - \tilde{u}_2^e(\theta_1, \theta_2'))\, dF(\theta_1) > 0.$$

The above inequality follows because $\tilde{u}_2^e(\theta_1, \theta_2'') > \tilde{u}_2^e(\theta_1, \theta_2')$ for all $\theta_1 \in \Theta_1$ and $\Theta_1^*(\theta_2') \neq \emptyset$. Note that the expected utility of the buyer after deviation is also identical to that in the GS mechanism with $\psi = 0$, which is $U_2^{\psi_0}(\theta_2''; \theta_2')$. Then, we have $U_2^{\psi_0}(\theta_2''; \theta_2') > U_2^{\psi_0}(\theta_2''; \theta_2'')$, implying that the buyer of type θ_2'' has an incentive to deviate to θ_2' in the GS mechanism with $\psi = 0$.

However, in this example, ω can take any value in $[0, 1]$ regardless of the agents' types, implying that all possible allocation payoffs of the seller are feasible regardless of the agents' types. This exhibits a large degree of payoff uncertainty beyond the level determined by the type profile. Hence, Condition (i) of Assumption 1 is violated. This completes the proof. □

We next provide another example where Condition (ii) of Assumption 1 is rather violated.

Example 3. Suppose that each agent i's type θ_i is drawn from the uniform distribution on the unit interval $[0, 1]$. Conditional on the type profile θ, ω_1 and ω_2 are independently drawn from the uniform distribution on the closed interval $[-0.1, 0.1]$. Then, we define $\Omega = \{(\omega_1, \omega_2) |\ \omega_1, \omega_2 \in [-0.1, 0.1]\}$ as the set of residual uncertainties. Each agent's valuation for the object when the true type profile is (θ_1, θ_2) and the residual uncertainty is $\omega = (\omega_1, \omega_2) \in \Omega$ is given as follows:

$$\begin{aligned} \tilde{u}_1(\theta_1, \theta_2; \omega) &= \theta_1 + 2\theta_2 + \omega_1; \\ \tilde{u}_2(\theta_1, \theta_2; \omega) &= 0.5 + 2\theta_1 + \theta_2 + \omega_2. \end{aligned}$$

For each type profile (θ_1, θ_2), we have

$$\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_1^e(\theta_1, \theta_2) = 0.5 + 2\theta_1 + \theta_2 - (\theta_1 + 2\theta_2) = 0.5 + \theta_1 - \theta_2,$$

implying that it is efficient to trade if and only if $\theta_2 < 0.5 + \theta_1$. In particular, if $\theta_2 \leq 0.5$, then $\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_1^e(\theta_1, \theta_2) = 0.5 + \theta_1 - \theta_2 \geq 0$ regardless of the seller's type, implying that it is always efficient to trade. In the example, $\Theta_2^{**} = [0, 0.5]$ and $\theta_2^* = \inf \Theta_2^{**} = 0$.

Claim 4. Condition (ii) of Assumption 1 is violated in Example 3.

Proof. We will show that the buyer of type $\theta_2 > 0.5$ has an incentive to deviate to $\theta_2^r \leq 0.5$ in the GS mechanism with $\psi = 0$, but condition (ii) of Assumption 1 is violated.

Suppose that the buyer's true type is $\theta_2 > 0.5$. If everyone reports truthfully in both stages, according to the transfer rule in the GS mechanism, the buyer is left with zero expected utility. Note that the expected utility of the buyer under truth-telling is identical to that in the GS mechanism with $\psi = 0$, which is $U_2^{\psi_0}(\theta_2; \theta_2)$. If the buyer deviates to $\theta_2^r \leq 0.5$, however, his expected utility becomes

$$\int_{\Theta_1} [\tilde{u}_2^e(\theta_1, \theta_2) - g(\theta_1)]\, dF(\theta_1) = \int_{\Theta_1} [\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^*)]\, dF(\theta_1) > 0.$$

The equality follows because by the proof of Claim 1, we have $\int_{\Theta_1} g(\theta_1) dF(\theta_1) = \int_{\Theta_1} \tilde{u}_2^e(\theta_1, \theta_2^*) dF(\theta_1)$. The inequality follows because the buyer's expected valuation is strictly increasing in his own type and $\theta_2 > \theta_2^* = 0$. Note that the expected utility of the buyer after deviation is also identical to that in the GS mechanism with $\psi = 0$, which is $U_2^{\psi_0}(\theta_2; \theta_2^r)$. Then, we have $U_2^{\psi_0}(\theta_2; \theta_2^r) > U_2^{\psi_0}(\theta_2; \theta_2)$, implying that the buyer of type θ_2 has an incentive to deviate to $\theta_2^r \leq 0.5$ in the GS mechanism with $\psi = 0$. However, the good is never allocated to the seller after the buyer deviates, implying that condition (ii) of Assumption 1 is violated. This completes the proof. $\square$

5 When the GS Mechanism Satisfies IIR

This section is organized as follows. In Section 5.1, we introduce Assumption 2. In Section 5.2, we show in Theorem 2 that when Assumption 1 holds, there exists a two-stage mechanism satisfying BIC, IIR, EFF, and BB if and only if Assumption 2 is satisfied.

5.1 A Key Assumption

To have an intuitive account for Assumption 2, we introduce the following terminologies. The ex ante gains from trade over the entire type space Θ equals

$$\int_{\Theta_1} \int_{\Theta_2^*(\theta_1)} (\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_1^e(\theta_1, \theta_2))\, dF(\theta_2) dF(\theta_1),$$

where $\Theta_2^*(\theta_1) = \{\theta_2 \in \Theta_2 : x^*(\theta_1, \theta_2) = 1\}$ for each $\theta_1 \in \Theta_1$. Recall in the proof of Proposition 2 that if agents' reports in both stages are truthful and the buyer has type $\theta_2 < \theta_2^*$, then the buyer's expected utility in the GS mechanism is

$$\int_{\Theta_1^*(\theta_2)} (\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2))\, dF(\theta_1) = 0,$$

where $\Theta_1^*(\theta_2) = \{\theta_1 \in \Theta_1 : x^*(\theta_1,\theta_2) = 1\}$; on the other hand, if the buyer has type $\theta_2 \geq \theta_2^*$, his expected utility becomes

$$\int_{\Theta_1} (\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2^*))\, dF(\theta_1),$$

where $\theta_2^* \in (\underline{\theta}, \overline{\theta}]$ is the cutoff point identified in Lemma 2. As a result, the buyer's ex ante expected utility in the GS mechanism equals

$$\int_{\Theta_1} \int_{\Theta_2^{**}} (\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2^*))\, dF(\theta_2) dF(\theta_1),$$

where $\Theta_2^{**} = [\theta_2^*, \overline{\theta}]$. We are ready to introduce Assumption 2.

Assumption 2.

$$\begin{aligned} &\int_{\Theta_1} \int_{\Theta_2^*(\theta_1)} (\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_1^e(\theta_1,\theta_2))\, dF(\theta_2) dF(\theta_1) \\ - &\int_{\Theta_1} \int_{\Theta_2^{**}} (\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2^*))\, dF(\theta_2) dF(\theta_1) \geq 0. \end{aligned} \tag{6}$$

Remark: (1) Since the ex ante gains from trade is always nonnegative and the buyer receives zero expected utility in the GS mechanism if $\theta_2^* = \overline{\theta}$, Assumption 2 is automatically satisfied in Case 1: $\theta_2^* = \overline{\theta}$.
(2) Assumption 2 guarantees that the ex ante gains from trade are split between the buyer and seller; otherwise, the buyer receives too much gains from trade such that the seller is worse off after participating in the mechanism, implying that the seller's IIR constraint is violated.

5.2 Existence of Desired Two-Stage Mechanisms

First, we show that the GS mechanism satisfies the seller's IIR constraints if and only if Assumption 2 is satisfied.

Proposition 4. Suppose that Assumption 1 holds. The GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ satisfies the seller's IIR constraints if and only if Assumption 2 is satisfied.

Proof. The proof is in the Appendix. □

The following is the main result of this section.

Theorem 2. Suppose that Assumption 1 holds. Then, there exists a two-stage mechanism satisfying BIC, IIR, EFF and BB if and only if Assumption 2 is satisfied.

Proof. It follows from Proposition 2 that there exists a threshold $\psi^* > 0$ such that the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ with $\psi \geq \psi^*$ satisfies BIC if and only if Assumption 1 is satisfied. The sufficiency part is straightforward, as we know from Proposition 4 that the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ satisfies IIR if Assumption 2 is satisfied.

To prove the necessity part, we suppose, on the contrary, that Assumption 2 is violated. We have shown in the proof of Proposition 4 that if Assumption 2 is violated, the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ violates the seller's ex ante individual rationality constraint, which is weaker than the seller's IIR constraint. We know from Proposition 3 that the GS mechanism maximizes the seller's ex ante expected utility among all two-stage mechanisms satisfying BIC, EFF, BB, and the buyer's IIR. We thus conclude that there exist no two-stage mechanisms satisfying BIC, IIR, EFF and BB. This completes the proof of the necessity part. □

6 Implications of Assumption 2

To assess the permissiveness of Assumption 2, we provide a set of simulation results based on Example 1. Recall that we can divide the analysis of Example 1 into two cases: (i) $0 < \gamma_2 \leq \gamma_1 < 1$ and (ii) $0 < \gamma_1 < \gamma_2 < 1$. It follows from Figures 3 and 4 that we have $\theta_2^* = \overline{\theta}$ in Case (i) and $\theta_2^* \in (\underline{\theta}, \overline{\theta})$ in Case (ii), respectively. Recall that Assumption 1 is satisfied in Example 1. Then, there always exists a $\psi^* > 0$ such that the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ with $\psi \geq \psi^*$ satisfies BIC in Example 1. Moreover, since Assumption 2 is always satisfied in Case (i): $\theta_2^* = \overline{\theta}$, we obtain by Proposition 4 that the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ with $\psi \geq \psi^*$ satisfies BIC, IIR, EFF and BB in Case (i). Thus, what remains to investigate is the extent to which there exists a two-stage mechanism satisfying all the desired properties in Case (ii) $0 < \gamma_1 < \gamma_2 < 1$, or equivalently, the extent to which Assumption 2 is satisfied. In the simulation, we select finitely many values of γ_1 and γ_2 satisfying $0 < \gamma_1 < \gamma_2 < 1$, that is, $\gamma_1 \in \{0.01, 0.02, \cdots, 0.98\}$ and $\gamma_2 \in \{\gamma_1 + 0.01, \gamma_1 + 0.02, \cdots, 0.99\}$ for each γ_1. To check whether Assumption 2 is satisfied in Case (ii), we use the following result.

Lemma 3. In Case (ii): $\theta_2^* \in (\underline{\theta}, \overline{\theta})$ of Example 1, our Assumption 2 is equivalent to

$$\frac{1}{6}\frac{(1-\gamma_2)^2}{1-\gamma_1} + \frac{1-\gamma_2}{1-\gamma_1} - \frac{1}{2}\left(\frac{1-\gamma_2}{1-\gamma_1}\right)^2 + \frac{1}{2}(\gamma_2 - \gamma_1 - 1) \geq 0. \tag{7}$$

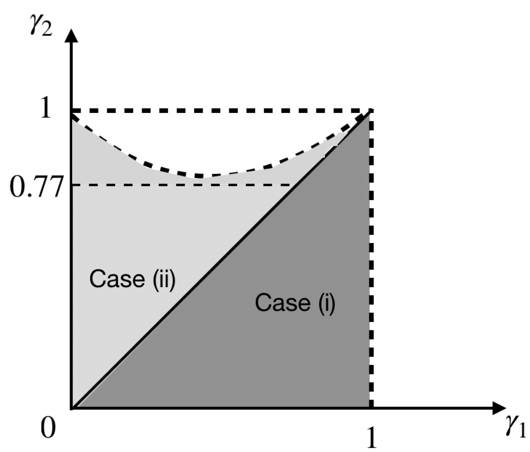
Proof. The proof is in the Appendix. □

This lemma shows that Assumption 2 is reduced to an inequality in terms of γ_1 and γ_2. In Figure 5, we track all possible pairs of $(\gamma_1, \gamma_2) \in (0,1)^2$ satisfying Assumption 2. In particular, the upper triangle in $[0,1]^2$, i.e., the region where $\gamma_2 > \gamma_1$, corresponds to Case (ii). The lightly shaded region describes all pairs of (γ_1, γ_2) within this upper triangle for which our Assumption 2 is satisfied. Note that Assumption 2 is always satisfied for all (γ_1, γ_2) satisfying $0 < \gamma_1 < \gamma_2 \leq 0.77$.

On the other hand, the lower triangle in the unit square, i.e., the region where $\gamma_2 < \gamma_1$ corresponds to Case (i). Since Assumption 2 is always satisfied within this region, the heavily shaded region describes all pairs of (γ_1, γ_2) within the lower triangle for which our Assumption 2 is satisfied.

Therefore, the lightly and heavily shaded regions together characterize the set of (γ_1, γ_2) for which our Assumption 2 is satisfied. Since the whole shaded (regardless of whether lightly or heavily) region spans quite a large part of the unit square, we conclude that our Assumption 2 can be satisfied in many cases in the example.

Figure 5: Summary of Simulation



7 Discussion

In Subsection 7.1, we discuss how our Assumption 2 can be dropped if we settle for *virtual* implementation rather than exact implementation which we pursue in this paper. In contrast, we show that Assumption 1 is replaced with Assumption 3 if we move from exact to virtual implementation. Subsection 7.2 discusses the relation with Galavotti, Muto, and Oyama (2011) who consider the problem of partnership dissolution in a deterministic model with interdependent values. In

Subsection 7.3, we discuss the relation with Makowski and Mezzetti (1994) who study the existence of BIC, IIR, EFF and BB "one-stage" mechanisms in general mechanism design problems with private values.

7.1 Virtual Implementation

We introduce the following metric over allocation rules: for any allocation rules $x, x^{'} : \Theta \to [0, 1]$, define

$$d(x, x^{'}) = \max_{\theta \in \Theta} |x(\theta) - x^{'}(\theta)|.$$

Let $\varepsilon \in (0, 1)$. We say that the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ that satisfies BIC, IIR, EFF, and BB is *ε-implementable* if there exists an allocation rule x^ε with $d(x^\varepsilon, x^*) = \varepsilon$ such that the modified GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{GS}, \psi)$ satisfies BIC, IIR and BB. We further say that the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ is *virtually implementable* if, for any $\varepsilon \in (0, 1)$, it is ε-implementable.

For each $\theta \in \Theta$, we define $x^\varepsilon(\theta) = (1 - \varepsilon)x^*(\theta)$. By construction, we have $d(x^\varepsilon, x^*) = \varepsilon$. Throughout the discussion in this section, we use this x^ε as the allocation rule which approximates x^*. Given x^ε, we define $\Theta^{**}_{2,\varepsilon} = \{\theta_2 \in \Theta_2 : x^\varepsilon(\theta_1, \theta_2) = 1 \text{ for all } \theta_1 \in \Theta_1\}$. Obviously, $\Theta^{**}_{2,\varepsilon} = \emptyset$ and thus, by the definition of θ^*_2, we set $\theta^*_2 = \overline{\theta}$. Taking into account that $\theta^*_2 = \overline{\theta}$, we define a two-stage mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$ as the *approximate* GS mechanism: for each type report $(\theta^r_1, \theta^r_2) \in \Theta$ and each payoff report $(u^r_1, u^r_2) \in \Pi_1 \times \Pi_2$,

$$t^{\varepsilon,GS}_1(\theta^r_1, \theta^r_2; u^r_1, u^r_2) = \begin{cases} \tilde{u}^\varepsilon_2(\theta^r_1, \theta^r_2) & \text{if } x^\varepsilon(\theta^r_1, \theta^r_2) \text{ dictates that the buyer receives the good} \\ & \text{after randomization and } u^r_2 \in \widetilde{\Pi}_2(\theta^r_1, \theta^r_2); \\ -\psi & \text{if } x^\varepsilon(\theta^r_1, \theta^r_2) \text{ dictates that the buyer receives the good} \\ & \text{after randomization and } u^r_2 \notin \widetilde{\Pi}_2(\theta^r_1, \theta^r_2); \\ 0 & \text{if } x^\varepsilon(\theta^r_1, \theta^r_2) \text{ dictates that the seller keeps the good} \\ & \text{after randomization,} \end{cases}$$

and

$$t^{\varepsilon,GS}_2(\theta^r_1, \theta^r_2; u^r_1, u^r_2) = \begin{cases} -\tilde{u}^\varepsilon_2(\theta^r_1, \theta^r_2) & \text{if } x^\varepsilon(\theta^r_1, \theta^r_2) \text{ dictates that the buyer receives the good} \\ & \text{after randomization;} \\ 0 & \text{if } x^\varepsilon(\theta^r_1, \theta^r_2) \text{ dictates that the seller keeps the good} \\ & \text{after randomization and } u^r_1 \in \widetilde{\Pi}_1(\theta^r_1, \theta^r_2); \\ -\psi & \text{if } x^\varepsilon(\theta^r_1, \theta^r_2) \text{ dictates that the seller keeps the good} \\ & \text{after randomization and } u^r_1 \notin \widetilde{\Pi}_1(\theta^r_1, \theta^r_2). \end{cases}$$

We introduce the *zero-trade* mechanism (Θ, Π, x^0, t^0) as follows:

$$x^0(\theta_1^r, \theta_2^r) = 0 \text{ for all } (\theta_1^r, \theta_2^r) \in \Theta$$

and for all $(\theta_1^r, \theta_2^r) \in \Theta$ and $(u_1^r, u_2^r) \in \Pi_1 \times \Pi_2$,

$$t_1^0(\theta_1^r, \theta_2^r; u_1^r, u_2^r) = 0,$$

and

$$t_2^0(\theta_1^r, \theta_2^r; u_1^r, u_2^r) = \begin{cases} 0 & \text{if } u_1^r \in \widetilde{\Pi}_1(\theta_1^r, \theta_2^r); \\ -\psi & \text{otherwise.} \end{cases}$$

As we argue below, this mechanism will be used to illustrate the relation between the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ and the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$.

Claim 5. The approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$ is equivalent to the following random mechanism: the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ is played with probability $1 - \varepsilon$ and the zero-trade mechanism (Θ, Π, x^0, t^0) is played with probability ε.

Proof. The proof is in the Appendix. □

For each $i \in \{1, 2\}$ and $\theta_i, \theta_i^r \in \Theta_i$, we denote by $U_i^{\varepsilon,\psi_0}(\theta_i; \theta_i^r)$ the expected utility of agent i of type θ_i when agent i reports θ_i^r and agent $j \neq i$ reports their type truthfully in the first stage in the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi_0)$, where $\psi_0 = 0$. We also denote by $x_i^\varepsilon(\theta)$ the probability that the good is allocated to agent i when θ is the reported type profile. Assumption 3 below is stated as Assumption 1 in verbatim, except that the efficient decision rule x^* is replaced by the allocation rule x^ε.

Assumption 3. For each agent $i \in \{1, 2\}$ and each pair of types $\theta_i, \theta_i^r \in \Theta_i$, if $U_i^{\varepsilon,\psi_0}(\theta_i; \theta_i^r) > U_i^{\varepsilon,\psi_0}(\theta_i; \theta_i)$, then there exists a positive-measure set $\Theta_j^+ \subseteq \Theta_j$ where $j \neq i$ such that (i) $\int_{\Omega^\psi(\theta_i,\theta_j;\theta_i^r)} dH\left(\omega|(\theta_i, \theta_j)\right) > 0$ for all $\theta_j \in \Theta_j^+$, and (ii) $\int_{\Theta_j^+} x_j^\varepsilon(\theta_i^r, \theta_j) dF(\theta_j) > 0$.

We establish below that Assumption 3 is a necessary and sufficient condition for the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$ to satisfy BIC. The proof of Proposition 5 is completed verbatim in the proof of Proposition 2, except that the efficient decision rule x^* is replaced by the allocation rule x^ε and $\theta_2^* = \bar{\theta}$ always hold under x^ε. Hence, we omit the proof.

Proposition 5. There exists a threshold $\psi_\varepsilon^* > 0$ such that the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$ with $\psi \geq \psi_\varepsilon^*$ satisfies BIC if and only if Assumption 3 is satisfied.

It is easy to show that the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$ always satisfies IIR.

Claim 6. The approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$ always satisfies IIR.

Proof. Consider the seller of type θ_1. If both agents report truthfully in both stages, the seller's expected utility in the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$ is computed below:

$$\begin{aligned}
& \int_{\Theta_2} \left[x^\varepsilon(\theta_1,\theta_2)\tilde{u}_2^e(\theta_1,\theta_2) + (1 - x^\varepsilon(\theta_1,\theta_2))\,\tilde{u}_1^e(\theta_1,\theta_2)\right] dF(\theta_2) \\
= & \int_{\Theta_2} \left[(1-\varepsilon)x^*(\theta_1,\theta_2)\tilde{u}_2^e(\theta_1,\theta_2) + (1 - (1-\varepsilon)x^*(\theta_1,\theta_2))\,\tilde{u}_1^e(\theta_1,\theta_2)\right] dF(\theta_2) \\
= & \int_{\Theta_2} (1-\varepsilon)x^*(\theta_1,\theta_2)\left[\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_1^e(\theta_1,\theta_2)\right] dF(\theta_2) + \int_{\Theta_2} \tilde{u}_1^e(\theta_1,\theta_2) dF(\theta_2) \\
\geq & \int_{\Theta_2} \tilde{u}_1^e(\theta_1,\theta_2) dF(\theta_2),
\end{aligned}$$

where the inequality follows because $x^*(\theta_1,\theta_2) = 1$ implies $\tilde{u}_2^e(\theta_1,\theta_2) \geq \tilde{u}_1^e(\theta_1,\theta_2)$. Recall that the seller's expected utility from the outside option equals $\int_{\Theta_2} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2)$. Hence, the seller receives a higher expected utility from participation.

Consider the buyer of type θ_2. If both agents report truthfully in both stages, the buyer's expected utility in the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$ is computed below:

$$\int_{\Theta_1} x^\varepsilon(\theta_1,\theta_2)\left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2)\right) dF(\theta_1) = 0.$$

Recall that the buyer receives zero expected utility from his outside option. Hence, the buyer is indifferent between participating in the mechanism or his outside option. We conclude that IIR is satisfied. This completes the proof. □

Combining Proposition 5 and Claim 6, we obtain the following result.

Proposition 6. The GS mechanism is virtually implementable if and only if Assumption 3 holds.

Finally, we establish the following relationship between Assumptions 1 and 3.

Claim 7. The following relationships are established:

1. Assume $\int_{\Theta_2^{**}} dF(\theta_2) = 0$. If Assumption 1 holds for x^*, then Assumption 3 holds for x^ε. However, the converse does not necessarily hold.

2. Assume $\int_{\Theta_2^{**}} dF(\theta_2) > 0$. Then, Assumption 1 for x^* neither implies nor is implied by Assumption 3 for x^ε.

Proof. The proof is in the Appendix. □

We nonetheless stress that Assumption 3 is not a trivial assumption. Recall in Subsection 4.5 that Assumption 1 is violated in Example 2 because the buyer has an incentive to deviate in the first stage and all possible allocation payoffs of the seller are feasible regardless of the agents' types. Indeed, the same argument can be extended to virtual implementation to show that Assumption 3 is violated in Example 2 as well, except that x^* is replaced by x^ε.

7.2 The Relation with Galavotti, Muto, and Oyama (2011)

Galavotti, Muto, and Oyama (2011) (hereafter, GMO) consider the problem of efficient partnership dissolution in a deterministic model with interdependent values. Their model includes our bilateral trade model as a special case of it. In their Theorem 4, GMO (2011) provide their Assumption 5.1 under which the shoot-the-liar mechanism of Mezzetti (2007) satisfies BIC, EFF, BB as well as ex post individual rationality (EPIR), which is stronger than our IIR, for any ownership structure in the deterministic model.

Here we restrict our attention to the deterministic version of our Example 1 to assess the permissiveness of GMO's (2011) Assumption 5.1. Lemma 6 of Kunimoto and Zhang (2022) shows that GMO's (2011) Assumption 5.1 holds if and only if $\gamma_1 = \gamma_2$. Therefore, our Assumption 2 is more permissive than GMO's Assumption 5.1 because, as shown in Figure 5, there are many pairs of (γ_1, γ_2) in $(0,1)^2$ for which the GS mechanism satisfies BIC, IIR, EFF, and BB.

7.3 The Relation with Makowski and Mezzetti (1994)

Makowski and Mezzetti (1994) consider general mechanism design problems in a deterministic model with private values. In their Theorem 3.1, they show that the Groves mechanism is canonical in the sense that there exists a mechanism that satisfies BIC, IIR, EFF, and BB if and only if the ex ante budget deficit generated by the Groves mechanism be less than or equal to the sum of the maximum charges one can impose without violating anyone's IIR constraint.[13]

[13] See also Krishna and Perry (2000) and Williams (1999) for the similar results.

In this paper, we restrict attention to bilateral trade, while we allow for a random model with interdependent values and two-stage mechanisms. Kunimoto and Zhang (2022) show by a stylized model of bilateral trade with interdependent values that the generalized two-stage Groves mechanism, arguably a natural analogue of the Groves mechanism, always violates IIR, while the shoot-the-liar mechanism sometimes satisfies it. Therefore, we cannot extend the canonicality of the Groves mechanism to our setup. However, we establish in our Theorem 1 that the GS mechanism, not the two-stage Groves mechanism, is rather canonical in the sense that, under Assumption 1, if there exists a two-stage mechanism satisfying BIC, IIR, EFF and BB, the GS mechanism satisfies the same properties. We show in our Theorem 2 that, when Assumption 1 holds, there exists a BIC, IIR, EFF, BB mechanism if and only if Assumption 2 is satisfied. Our Assumption 2 requires that the ex ante gains from trade in the GS mechanism be at least as high as the buyer's ex ante expected utility.

8 Conclusion

This paper characterizes the conditions under which efficient, voluntary bilateral trades are implementable in an interdependent values environment in which agents' information is ex ante symmetric. Acknowledging some existing impossibility results by Gresik (1991) and FKM (2003), we obtain more positive results by looking at two-stage mechanisms proposed by Mezzetti (2004). The main results of this paper are summarized under Assumption 1: (i) there exists a two-stage mechanism satisfying BIC, EFF, BB, and IIR if and only if the GS mechanism satisfies BIC, EFF, BB, and IIR; and (ii) the GS mechanism satisfies BIC, EFF, BB, and IIR if and only if Assumption 2 is satisfied. Therefore, as long as Assumption 1 holds, the GS mechanism is proposed as a canonical two-stage mechanism with the desired property. We argue by means of examples that there are some scenarios in which Assumption 1 holds, whereas there are other scenarios in which Assumption 1 fails. In the context of Example 1, Assumption 2 is satisfied as long as the buyer's degree of interdependence of preferences is not too high relative to the seller's counterpart. We also argue by the same example that Assumption 2 is restrictive but can still be satisfied for a large number of cases. By expanding our scope into two-stage mechanisms, we further push the boundary between when efficient, voluntary bilateral trades are implementable and when they are not. We argue that two-stage mechanisms can be justified in situations where two parties in a bilateral trade setup invite a trusted mediator (a third party) to their contractual relationship: the mediator asks both agents to put a large sum of money as a deposit in the mediator's account and the mediator pays back the remaining deposit to each agent after the two-stage mechanism has played out. Although positive results of our paper ultimately rely on the plausibility for the use of two-stage mechanisms, we argue that two-stage mechanisms can sometimes be implemented through a long-term relationship and/or a smart contract based on the blockchain technology.

9 Appendix

In the Appendix, we provide the proofs omitted from the main text of the paper.

9.1 Proof of Claim 1

Proof. In this claim, we establish two results, which we call Results 1 and 2, respectively.

Result 1: Suppose $\int_{\Theta_2^{**}} dF(\theta_2) > 0$. Suppose that for each $\theta_1 \in \Theta_1$, there exists $C \geq 0$ such that

$$\int_{\Theta_2^{**}} g(\theta_1)dF(\theta_2) = -\int_{\Theta_2^*(\theta_1)\backslash\Theta_2^{**}} \tilde{u}_2^e(\theta_1,\theta_2)dF(\theta_2) + \int_{\Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2) + C,$$

where $\tilde{u}_i^e(\theta_i,\theta_{-i}) = \int_\Omega \tilde{u}_i(\theta_i,\theta_{-i};\omega)dH\left(\omega|(\theta_i,\theta_{-i})\right)$ for each agent i, $\Theta_2^*(\theta_1) = \{\theta_2 \in \Theta_2 : x^*(\theta_1,\theta_2) = 1\}$ and $\Theta_2^{**} = \{\theta_2 \in \Theta_2 : x^*(\theta_1,\theta_2) = 1 \text{ for all } \theta_1 \in \Theta_1\}$. Consider the seller of type θ_1. Suppose that both agents report truthfully in both stages. The seller's interim expected utility after participating in the GS mechanism $(\Theta,\Pi,x^*,t^{GS},\psi)$, denoted by $U_1^{GS}(\theta_1)$, is computed as follows:

$$U_1^{GS}(\theta_1) = \int_{\Theta_2\backslash\Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2) + \int_{\Theta_2^*(\theta_1)\backslash\Theta_2^{**}} \tilde{u}_2^e(\theta_1,\theta_2)dF(\theta_2) + \int_{\Theta_2^{**}} g(\theta_1)dF(\theta_2).$$

Substituting the expression of $\int_{\Theta_2^{**}} g(\theta_1)dF(\theta_2)$ into $U_1^{GS}(\theta_1)$, we obtain

$$\begin{aligned} U_1^{GS}(\theta_1) = & \int_{\Theta_2\backslash\Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2) + \int_{\Theta_2^*(\theta_1)\backslash\Theta_2^{**}} \tilde{u}_2^e(\theta_1,\theta_2)dF(\theta_2) \\ & - \int_{\Theta_2^*(\theta_1)\backslash\Theta_2^{**}} \tilde{u}_2^e(\theta_1,\theta_2)dF(\theta_2) + \int_{\Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2) + C. \end{aligned}$$

After rearrangement, we have

$$\begin{aligned} U_1^{GS}(\theta_1) &= \int_{\Theta_2} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2) + C \\ &= U_1^O(\theta_1) + C \quad \left(\because\ U_1^O(\theta_1) = \int_{\Theta_2} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2)\right), \end{aligned}$$

where $U_1^O(\theta_1)$ denotes the outside option utility of seller of type θ_1. Since $C \geq 0$, we conclude that $U_1^{GS}(\theta_1) \geq U_1^O(\theta_1)$, implying that the seller's IIR constraint is satisfied.

Result 2: Suppose $\int_{\Theta_2^{**}} dF(\theta_2) > 0$. We set

$$\begin{aligned} C &= \int_{\Theta_1} \int_{\Theta_2^*(\theta_1)} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_1^e(\theta_1, \theta_2)\right) dF(\theta_2) dF(\theta_1) \\ &\quad - \int_{\Theta_1} \int_{\Theta_2^{**}} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^*)\right) dF(\theta_2) dF(\theta_1). \end{aligned}$$

Consider the buyer of type $\theta_2 \in \Theta_2^{**}$. Suppose that both agents report truthfully in both stages. The buyer's interim expected utility after participating in the GS mechanism, denoted by $U_2^{GS}(\theta_2)$, is computed as follows:

$$U_2^{GS}(\theta_2) = \int_{\Theta_1} \left(\tilde{u}_2^e(\theta_1, \theta_2) - g(\theta_1)\right) dF(\theta_1).$$

Substituting the expression of C into the expression of $\int_{\Theta_2^{**}} g(\theta_1) dF(\theta_2)$, we obtain

$$\begin{aligned} &\int_{\Theta_1} \int_{\Theta_2^{**}} g(\theta_1) dF(\theta_2) dF(\theta_1) \\ = &-\int_{\Theta_1} \int_{\Theta_2^*(\theta_1) \setminus \Theta_2^{**}} \tilde{u}_2^e(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1) + \int_{\Theta_1} \int_{\Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1) + C \\ = &-\int_{\Theta_1} \int_{\Theta_2^*(\theta_1) \setminus \Theta_2^{**}} \tilde{u}_2^e(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1) + \int_{\Theta_1} \int_{\Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1) \\ &+ \int_{\Theta_1} \int_{\Theta_2^*(\theta_1)} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_1^e(\theta_1, \theta_2)\right) dF(\theta_2) dF(\theta_1) \\ &- \int_{\Theta_1} \int_{\Theta_2^{**}} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^*)\right) dF(\theta_2) dF(\theta_1). \end{aligned}$$

After rearranging terms, we have

$$\begin{aligned} &\int_{\Theta_1} \int_{\Theta_2^{**}} g(\theta_1) dF(\theta_2) dF(\theta_1) \\ = &-\int_{\Theta_1} \int_{\Theta_2^*(\theta_1) \setminus \Theta_2^{**}} \tilde{u}_2^e(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1) + \int_{\Theta_1} \int_{\Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1) \\ &+ \int_{\Theta_1} \int_{\Theta_2^*(\theta_1) \setminus \Theta_2^{**}} \tilde{u}_2^e(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1) - \int_{\Theta_1} \int_{\Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1) \\ &+ \int_{\Theta_1} \int_{\Theta_2^{**}} \tilde{u}_2^e(\theta_1, \theta_2^*) dF(\theta_2) dF(\theta_1). \end{aligned}$$

Noticing that the first four terms are cancelled out, we obtain

$$\begin{aligned}\int_{\Theta_1}\int_{\Theta_2^{**}} g(\theta_1)dF(\theta_2)dF(\theta_1) &= \int_{\Theta_1}\int_{\Theta_2^{**}} \tilde{u}_2^e(\theta_1,\theta_2^*)dF(\theta_2)dF(\theta_1) \\ \Rightarrow \left(\int_{\Theta_2^{**}} dF(\theta_2)\right)\int_{\Theta_1} g(\theta_1)dF(\theta_1) &= \left(\int_{\Theta_2^{**}} dF(\theta_2)\right)\int_{\Theta_1} \tilde{u}_2^e(\theta_1,\theta_2^*)dF(\theta_1).\end{aligned}$$

Therefore, we obtain

$$\int_{\Theta_1} g(\theta_1)dF(\theta_1) = \int_{\Theta_1} \tilde{u}_2^e(\theta_1,\theta_2^*)dF(\theta_1).$$

Substituting this back into $U_2^{GS}(\theta_2)$, we obtain

$$U_2^{GS}(\theta_2) = \int_{\Theta_1} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2^*)\right) dF(\theta_1),$$

which is nonnegative because $\theta_2 \in \Theta_2^{**}$ implies $\theta_2 \geq \theta_2^* = \inf \Theta_2^{**}$ and thus $\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2^*) \geq 0$ for each $\theta_1 \in \Theta_1$. Since the buyer's outside option utility is always zero, we conclude that the IIR constraint for the buyer of type $\theta_2 \in \Theta_2^{**}$ is satisfied. This completes the proof of Claim 1. □

9.2 Proof of Proposition 1

Step 1: When $\int_{\Theta_2^{**}} dF(\theta_2) = 0$, the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ satisfies IIR.

Proof. We first show that IIR is satisfied for the seller. Consider the seller of type θ_1. Recall that if both agents report truthfully and trade occurs, the seller's monetary transfer is $\tilde{u}_2^e(\theta_1,\theta_2)$. Then, the expected utility of the seller under truth-telling is

$$\int_{\Theta_2 \setminus \Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2) + \int_{\Theta_2^*(\theta_1)} \tilde{u}_2^e(\theta_1,\theta_2)dF(\theta_2),$$

where $\tilde{u}_i^e(\theta_i,\theta_{-i}) = \int_\Omega \tilde{u}_i(\theta_i,\theta_{-i};\omega)dH\left(\omega|(\theta_i,\theta_{-i})\right)$ for each agent $i \in \{1,2\}$ and $\Theta_2^*(\theta_1) = \{\theta_2 \in \Theta_2 : x^*(\theta_1,\theta_2) = 1\}$. we claim that this expected utility is at least as high as the outside option utility. To see this, we compute the difference between the seller's expected utility under truth-telling and her outside option utility:

$$
\begin{aligned}
& \int_{\Theta_2\setminus\Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2) + \int_{\Theta_2^*(\theta_1)} \tilde{u}_2^e(\theta_1,\theta_2)dF(\theta_2) - \int_{\Theta_2} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2) \\
= & \int_{\Theta_2\setminus\Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2) + \int_{\Theta_2^*(\theta_1)} \tilde{u}_2^e(\theta_1,\theta_2)dF(\theta_2) \\
& - \left[\int_{\Theta_2\setminus\Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2) + \int_{\Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2)\right] \\
= & \int_{\Theta_2^*(\theta_1)} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_1^e(\theta_1,\theta_2)\right) dF(\theta_2) \\
\geq & \ 0,
\end{aligned}
$$

where the weak inequality follows because whenever $\theta_2 \in \Theta_2^*(\theta_1)$, it is efficient to trade, implying that $\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_1^e(\theta_1,\theta_2) \geq 0$. We conclude that the seller's expected utility by participating in the mechanism is at least as high as that from the outside option. This implies that IIR is satisfied for the seller.

Consider the buyer of type θ_2. Recall that if both agents report truthfully and trade occurs, the buyer's monetary transfer is $-\tilde{u}_2^e(\theta_1,\theta_2)$. Then, the expected utility of the buyer under truth-telling is

$$
\int_{\Theta_1^*(\theta_2)} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2)\right) dF(\theta_1) = 0 = U_2^O(\theta_2),
$$

where $\Theta_1^*(\theta_2) = \{\theta_1 \in \Theta_1 : x^*(\theta_1,\theta_2) = 1\}$ and $U_2^O(\theta_2)$ denotes the outside option utility of the buyer of type θ_2. Hence, by participating in the mechanism, the buyer receives exactly the same expected utility as his outside option utility. We thus conclude that IIR is satisfied for the buyer. This completes the proof of Step 1. □

Step 2: When $\int_{\Theta_2^{**}} dF(\theta_2) > 0$, the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ satisfies the buyer's IIR constraint.

Proof. We show that IIR is satisfied for the buyer. Consider the buyer of type θ_2. If $\theta_2 \notin \Theta_2^{**}$ and both agents' reports are truthful in both stages, the expected utility of the buyer of type θ_2 after participating in the mechanism is

$$
\int_{\Theta_1^*(\theta_2)} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2)\right) dF(\theta_1) = 0 = U_2^O(\theta_2),
$$

where $U_2^O(\theta_2)$ denotes the outside option utility of the buyer of type θ_2. Hence, if $\theta_2 \notin \Theta_2^{**}$, by participating in the mechanism, the buyer receives exactly the same expected utility as his outside option utility.

Moreover, it follows from the proof of Claim 1 that the buyer of type $\theta_2 \in \Theta_2^{**}$ receives the following expected utility if both agents' reports are truthful in both stages:

$$\int_{\Theta_1} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^*)\right) dF(\theta_1) \geq 0 = U_2^O(\theta_2),$$

where the inequality follows because $\theta_2 \in \Theta_2^{**}$ implies $\theta_2 \geq \theta_2^* = \inf \Theta_2^{**}$. We thus conclude that IIR is satisfied for the buyer. This completes the proof of Step 2. $\square$

9.3 Proof of Sufficiency Part of Proposition 2

We divide the proof into two cases: Case 1: $\theta_2^* = \overline{\theta}$ and Case 2: $\theta_2^* \in (\underline{\theta}, \overline{\theta})$. In each case, the proof is completed by two steps.

Case 1: $\theta_2^* = \overline{\theta}$, i.e., $\int_{\Theta_1} x^*(\theta_1, \theta_2) dF(\theta_1) < 1$ for all $\theta_2 < \overline{\theta}$. In Case 1, we set

$$\psi_1^* = \max\{A_1, A_2\},$$

where

$$A_1 = \sup_{\substack{\theta_1 \in [\underline{\theta}, \overline{\theta}] \\ \theta_1^r \in \{\tilde{\theta}_1 : \tilde{\theta}_1 \in \Theta_1 \text{ and } U_1^{\psi_0}(\theta_1; \tilde{\theta}_1) > U_1^{\psi_0}(\theta_1; \theta_1)\}}} \frac{U_1^{\psi_0}(\theta_1; \theta_1^r) - U_1^{\psi_0}(\theta_1; \theta_1)}{\int_{\Theta_2^*(\theta_1^r)} \int_{\Omega^\psi(\theta_1, \theta_2; \theta_1^r)} dH\left(\omega|(\theta_1, \theta_2)\right) dF(\theta_2)}$$

and

$$A_2 = \sup_{\substack{\theta_2 \in [\underline{\theta}, \overline{\theta}] \\ \theta_2^r \in \{\tilde{\theta}_2 : \tilde{\theta}_2 \in \Theta_2 \text{ and } U_2^{\psi_0}(\theta_2; \tilde{\theta}_2) > U_2^{\psi_0}(\theta_2; \theta_2)\}}} \frac{U_2^{\psi_0}(\theta_2; \theta_2^r) - U_2^{\psi_0}(\theta_2; \theta_2)}{\int_{\Theta_1 \setminus \Theta_1^*(\theta_2^r)} \int_{\Omega^\psi(\theta_2, \theta_1; \theta_2^r)} dH\left(\omega|(\theta_1, \theta_2)\right) dF(\theta_1)},$$

where $\Theta_2^*(\theta_1^r) = \{\theta_2 \in \Theta_2 : x^*(\theta_1^r, \theta_2) = 1\}$, $\Theta_1^*(\theta_2^r) = \{\theta_1 \in \Theta_1 : x^*(\theta_1, \theta_2^r) = 1\}$, and $U_i^{\psi_0}(\theta_i; \theta_i^r)$ is the expected utility of agent i when agent i has type θ_i and reports θ_i^r in the GS mechanism with $\psi = 0$ (which will be clarified in the proof later).

There are two steps to be checked in this case.

Step 1-1: If the buyer always reports the truth in the first stage, the seller has no incentive to tell a lie in the first stage in the GS mechanism with $\psi \geq \psi_1^*$.

Proof. Consider the seller of type θ_1. Recall that if both agents report truthfully in both stages, the expected utility of the seller of type θ_1 after participating in the mechanism is

$$\int_{\Theta_2 \setminus \Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2) + \int_{\Theta_2^*(\theta_1)} \tilde{u}_2^e(\theta_1, \theta_2) dF(\theta_2),$$

where $\tilde{u}_i^e(\theta_i, \theta_{-i}) = \int_\Omega \tilde{u}_i(\theta_i, \theta_{-i}; \omega) dH(\omega|(\theta_i, \theta_{-i}))$ for each agent $i \in \{1, 2\}$ and each type profile $(\theta_i, \theta_{-i}) \in \Theta$, and $\Theta_2^*(\theta_1) = \{\theta_2 \in \Theta_2 : x^*(\theta_1, \theta_2) = 1\}$. Note that the expected utility of the seller under truth-telling is identical to that in the GS mechanism with $\psi = 0$, which is $U_1^{\psi_0}(\theta_1; \theta_1)$. On the other hand, if trade occurs after the seller deviates to $\theta_1^r \neq \theta_1$ and the buyer's allocation payoff report cannot occur given the type profile (θ_1^r, θ_2), then the seller must pay a penalty ψ according to the transfer rule t_1^{GS}. Therefore, the expected utility of the seller of type θ_1 becomes

$$\int_{\Theta_2 \setminus \Theta_2^*(\theta_1^r)} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2) + \int_{\Theta_2^*(\theta_1^r)} \tilde{u}_2^e(\theta_1^r, \theta_2) \left[1 - \int_{\Omega^\psi(\theta_1, \theta_2; \theta_1^r)} dH(\omega|(\theta_1, \theta_2)) \right] dF(\theta_2)$$
$$-\psi \int_{\Theta_2^*(\theta_1^r)} \int_{\Omega^\psi(\theta_1, \theta_2; \theta_1^r)} dH(\omega|(\theta_1, \theta_2)) dF(\theta_2),$$

where $\Omega^\psi(\theta_1, \theta_2; \theta_1^r) = \left\{ \omega \in \Omega : \tilde{u}_2(\theta_1, \theta_2; \omega) \in \widetilde{\Pi}_2(\theta_1, \theta_2) \text{ but } \tilde{u}_2(\theta_1, \theta_2; \omega) \notin \widetilde{\Pi}_2(\theta_1^r, \theta_2) \right\}$, $\widetilde{\Pi}_2(\theta_1^r, \theta_2) = \{\hat{u}_2 \in \mathbb{R} :$ there exists a $\omega \in \Omega$ such that $h(\omega|(\theta_1^r, \theta_2)) > 0$ and $\tilde{u}_2(\theta_1^r, \theta_2; \omega) = \hat{u}_2\}$, and $h(\omega|(\theta_1^r, \theta_2))$ is the probability density function of the residual uncertainty ω conditional on the type profile being (θ_1^r, θ_2). Note that the expected utility of the seller after deviation can be rewritten as follows:

$$U_1^{\psi_0}(\theta_1; \theta_1^r) - \psi \int_{\Theta_2^*(\theta_1^r)} \int_{\Omega^\psi(\theta_1, \theta_2; \theta_1^r)} dH(\omega|(\theta_1, \theta_2)) dF(\theta_2), \tag{8}$$

where

$$U_1^{\psi_0}(\theta_1; \theta_1^r) = \int_{\Theta_2 \setminus \Theta_2^*(\theta_1^r)} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2) + \int_{\Theta_2^*(\theta_1^r)} \tilde{u}_2^e(\theta_1^r, \theta_2) \left[1 - \int_{\Omega^\psi(\theta_1, \theta_2; \theta_1^r)} dH(\omega|(\theta_1, \theta_2)) \right] dF(\theta_2)$$

is the expected utility of the seller after deviating to θ_1^r in the GS mechanism with $\psi = 0$.

If $U_1^{\psi_0}(\theta_1; \theta_1) \geq U_1^{\psi_0}(\theta_1; \theta_1^r)$, the deviation is not profitable. If $U_1^{\psi_0}(\theta_1; \theta_1) < U_1^{\psi_0}(\theta_1; \theta_1^r)$, then, by Assumption 1, there exists a positive-measure set $\Theta_2^+ \subseteq \Theta_2$ such that (i) $\int_{\Omega^\psi(\theta_1, \theta_2; \theta_1^r)} dH(\omega|(\theta_1, \theta_2)) > 0$ for all $\theta_2 \in \Theta_2^+$, and (ii) $\int_{\Theta_2^+} x_2^*(\theta_1^r, \theta_2) dF(\theta_2) > 0$ where $x_2^*(\theta_1^r, \theta_2)$ denotes the probability that the good is allocated to the buyer after the type report (θ_1^r, θ_2). As a result, the opposite of the coefficient of ψ in (8), $\int_{\Theta_2^*(\theta_1^r)} \int_{\Omega^\psi(\theta_1, \theta_2; \theta_1^r)} dH(\omega|(\theta_1, \theta_2)) dF(\theta_2)$, is strictly positive.

Since $\psi \geq \psi_1^* = \max\{A_1, A_2\}$ where

$$A_1 = \sup_{\substack{\theta_1 \in [\underline{\theta}, \overline{\theta}] \\ \theta_1^r \in \{\tilde{\theta}_1 : \tilde{\theta}_1 \in \Theta_1 \text{ and } U_1^{\psi_0}(\theta_1; \tilde{\theta}_1) > U_1^{\psi_0}(\theta_1; \theta_1)\}}} \frac{U_1^{\psi_0}(\theta_1; \theta_1^r) - U_1^{\psi_0}(\theta_1; \theta_1)}{\int_{\Theta_2^*(\theta_1^r)} \int_{\Omega^\psi(\theta_1, \theta_2; \theta_1^r)} dH(\omega|(\theta_1, \theta_2)) dF(\theta_2)},$$

we obtain that, for each $\theta_1, \theta_1^r \in \Theta_1$ such that $U_1^{\psi_0}(\theta_1; \theta_1^r) > U_1^{\psi_0}(\theta_1; \theta_1)$,

$$\begin{aligned}
& \psi \geq \frac{U_1^{\psi_0}(\theta_1; \theta_1^r) - U_1^{\psi_0}(\theta_1; \theta_1)}{\int_{\Theta_2^*(\theta_1^r)} \int_{\Omega^\psi(\theta_1, \theta_2; \theta_1^r)} dH\left(\omega | (\theta_1, \theta_2)\right) dF(\theta_2)} \\
\Rightarrow \;& \psi \int_{\Theta_2^*(\theta_1^r)} \int_{\Omega^\psi(\theta_1, \theta_2; \theta_1^r)} dH\left(\omega | (\theta_1, \theta_2)\right) dF(\theta_2) \geq U_1^{\psi_0}(\theta_1; \theta_1^r) - U_1^{\psi_0}(\theta_1; \theta_1),
\end{aligned}$$

because the denominator is strictly positive. Then, we can compare the seller's expected utility under truth-telling and that after deviation:

$$\begin{aligned}
& U_1^{\psi_0}(\theta_1; \theta_1^r) - \psi \int_{\Theta_2^*(\theta_1^r)} \int_{\Omega^\psi(\theta_1, \theta_2; \theta_1^r)} dH\left(\omega | (\theta_1, \theta_2)\right) dF(\theta_2) \\
\leq \;& U_1^{\psi_0}(\theta_1; \theta_1^r) - U_1^{\psi_0}(\theta_1; \theta_1^r) + U_1^{\psi_0}(\theta_1; \theta_1) \\
= \;& U_1^{\psi_0}(\theta_1; \theta_1),
\end{aligned}$$

implying that the expected utility of the seller after deviating to θ_1^r is at most as high as that under truth-telling. Therefore, the seller has no incentive to deviate to θ_1^r in the GS mechanism with $\psi \geq \psi_1^*$. This completes the proof of Step 1-1. □

Step 1-2: If the seller always reports the truth in the first stage, the buyer has no incentive to tell a lie in the first stage in the GS mechanism with $\psi \geq \psi_1^*$.

Proof. Consider the buyer of type θ_2. Recall that if both agents report truthfully in both stages, the expected utility of the buyer after participating in the mechanism is

$$\int_{\Theta_1^*(\theta_2)} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2)\right) dF(\theta_1) = 0,$$

where $\tilde{u}_2^e(\theta_1, \theta_2) = \int_\Omega \tilde{u}_2(\theta_1, \theta_2; \omega) dH\left(\omega | (\theta_1, \theta_2)\right)$ and $\Theta_1^*(\theta_2) = \{\theta_1 \in \Theta_1 : x^*(\theta_1, \theta_2) = 1\}$. Note that the expected utility of the buyer under truth-telling is identical to that in the GS mechanism with $\psi = 0$, which is $U_2^{\psi_0}(\theta_2; \theta_2)$. On the other hand, if no trade occurs after the buyer of type θ_2 deviates to $\theta_2^r \neq \theta_2$ and the seller's allocation payoff report is not consistent with the type profile (θ_1, θ_2^r), then the buyer must pay a penalty ψ according to the transfer rule t_2^{GS}. Therefore, the expected utility of the buyer of type θ_2 becomes

$$\int_{\Theta_1^*(\theta_2^r)} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^r)\right) dF(\theta_1) - \psi \int_{\Theta_1 \setminus \Theta_1^*(\theta_2^r)} \int_{\Omega^\psi(\theta_2, \theta_1; \theta_2^r)} dH\left(\omega | (\theta_1, \theta_2)\right) dF(\theta_1),$$

where $\Omega^\psi(\theta_2, \theta_1; \theta_2^r) = \left\{\omega \in \Omega : \tilde{u}_1(\theta_1, \theta_2; \omega) \in \widetilde{\Pi}_1(\theta_1, \theta_2) \text{ but } \tilde{u}_1(\theta_1, \theta_2; \omega) \notin \widetilde{\Pi}_1(\theta_1, \theta_2^r)\right\}$, $\widetilde{\Pi}_1(\theta_1, \theta_2^r) = \{\hat{u}_1 \in \mathbb{R} :$ there exists a $\omega \in \Omega$ such that $h\left(\omega | (\theta_1, \theta_2^r)\right) > 0$ and $\tilde{u}_1(\theta_1, \theta_2^r; \omega) =$

$\hat{u}_1\}$, and $h\left(\omega|(\theta_1,\theta_2^r)\right)$ is the probability density function of the residual uncertainty ω conditional on the type profile being (θ_1,θ_2^r). Note that the expected utility of the buyer after deviation can be rewritten as follows:

$$U_2^{\psi_0}(\theta_2;\theta_2^r) - \psi \int_{\Theta_1\setminus\Theta_1^*(\theta_2^r)} \int_{\Omega^\psi(\theta_2,\theta_1;\theta_2^r)} dH\left(\omega|(\theta_1,\theta_2)\right) dF(\theta_1), \tag{9}$$

where $U_2^{\psi_0}(\theta_2;\theta_2^r) = \int_{\Theta_1^*(\theta_2^r)} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2^r)\right) dF(\theta_1)$ is the expected utility of the buyer after deviating to θ_2^r in the GS mechanism with $\psi = 0$.

If $U_2^{\psi_0}(\theta_2;\theta_2) \geq U_2^{\psi_0}(\theta_2;\theta_2^r)$, the deviation is not profitable. If $U_2^{\psi_0}(\theta_2;\theta_2) < U_2^{\psi_0}(\theta_2;\theta_2^r)$, then, by Assumption 1, there exists a positive-measure set $\Theta_1^+ \subseteq \Theta_1$ such that (i) $\int_{\Omega^\psi(\theta_2,\theta_1;\theta_2^r)} dH\left(\omega|(\theta_1,\theta_2)\right) > 0$ for all $\theta_1 \in \Theta_1^+$, and (ii) $\int_{\Theta_1^+} x_1^*(\theta_1,\theta_2^r) dF(\theta_1) > 0$, where $x_1^*(\theta_1,\theta_2^r)$ denotes the probability that the seller keeps the good after the type report (θ_1,θ_2^r). In other words, the opposite of the coefficient of ψ in (9) is strictly positive, i.e., $\int_{\Theta_1\setminus\Theta_1^*(\theta_2^r)} \int_{\Omega^\psi(\theta_2,\theta_1;\theta_2^r)} dH\left(\omega|(\theta_1,\theta_2)\right) dF(\theta_1) > 0$.

Since $\psi \geq \psi_1^* = \max\{A_1, A_2\}$ where

$$A_2 = \sup_{\substack{\theta_2\in[\underline{\theta},\bar{\theta}] \\ \theta_2^r\in\{\tilde{\theta}_2:\tilde{\theta}_2\in\Theta_2 \text{ and } U_2^{\psi_0}(\theta_2;\tilde{\theta}_2)>U_2^{\psi_0}(\theta_2;\theta_2)\}}} \frac{U_2^{\psi_0}(\theta_2;\theta_2^r) - U_2^{\psi_0}(\theta_2;\theta_2)}{\int_{\Theta_1\setminus\Theta_1^*(\theta_2^r)} \int_{\Omega^\psi(\theta_2,\theta_1;\theta_2^r)} dH\left(\omega|(\theta_1,\theta_2)\right) dF(\theta_1)},$$

we obtain that, for each $\theta_2, \theta_2^r \in \Theta_2$ such that $U_2^{\psi_0}(\theta_2;\theta_2^r) > U_2^{\psi_0}(\theta_2;\theta_2)$,

$$\begin{aligned}
& \psi \geq \frac{U_2^{\psi_0}(\theta_2;\theta_2^r) - U_2^{\psi_0}(\theta_2;\theta_2)}{\int_{\Theta_1\setminus\Theta_1^*(\theta_2^r)} \int_{\Omega^\psi(\theta_2,\theta_1;\theta_2^r)} dH\left(\omega|(\theta_1,\theta_2)\right) dF(\theta_1)} \\
\Rightarrow \quad & \psi \int_{\Theta_1\setminus\Theta_1^*(\theta_2^r)} \int_{\Omega^\psi(\theta_2,\theta_1;\theta_2^r)} dH\left(\omega|(\theta_1,\theta_2)\right) dF(\theta_1) \geq U_2^{\psi_0}(\theta_2;\theta_2^r) - U_2^{\psi_0}(\theta_2;\theta_2),
\end{aligned}$$

because the denominator is strictly positive. Then, we can compare the buyer's expected utility under truth-telling and that after deviation:

$$\begin{aligned}
& U_2^{\psi_0}(\theta_2;\theta_2^r) - \psi \int_{\Theta_1\setminus\Theta_1^*(\theta_2^r)} \int_{\Omega^\psi(\theta_2,\theta_1;\theta_2^r)} dH\left(\omega|(\theta_1,\theta_2)\right) dF(\theta_1) \\
\leq \; & U_2^{\psi_0}(\theta_2;\theta_2^r) - U_2^{\psi_0}(\theta_2;\theta_2^r) + U_2^{\psi_0}(\theta_2;\theta_2) \\
= \; & U_2^{\psi_0}(\theta_2;\theta_2),
\end{aligned}$$

implying that the expected utility of the buyer after deviating to θ_2^r is at most as high as that under truth-telling. Therefore, the buyer has no incentive to deviate to θ_2^r in the GS mechanism with $\psi \geq \psi_1^*$. This completes the proof of Step 1-2. $\square$

We complete the proof for Case 1 and next proceed to Case 2.

Case 2: $\theta_2^* \in (\underline{\theta}, \overline{\theta})$ such that for any $\theta_2 \in \Theta_2$,

$$\int_{\Theta_1} x^*(\theta_1, \theta_2) dF(\theta_1) \begin{cases} < 1 & \text{if } \theta_2 < \theta_2^* \\ = 1 & \text{if } \theta_2 \geq \theta_2^*. \end{cases}$$

In Case 2, we set

$$\psi_2^* = \max\{B_1, B_2\},$$

where

$$B_1 = \sup_{\substack{\theta_1 \in [\underline{\theta}, \overline{\theta}] \\ \theta_1^r \in \{\tilde{\theta}_1 : \tilde{\theta}_1 \in \Theta_1 \text{ and } U_1^{\psi_0}(\theta_1; \tilde{\theta}_1) > U_1^{\psi_0}(\theta_1; \theta_1)\}}} \frac{U_1^{\psi_0}(\theta_1; \theta_1^r) - U_1^{\psi_0}(\theta_1; \theta_1)}{\int_{\Theta_2^*(\theta_1^r)} \int_{\Omega^{\psi}(\theta_1, \theta_2; \theta_1^r)} dH\left(\omega | (\theta_1, \theta_2)\right) dF(\theta_2)}$$

and

$$B_2 = \sup_{\substack{\theta_2 \in [\underline{\theta}, \overline{\theta}] \\ \theta_2^r \in \{\tilde{\theta}_2 : \tilde{\theta}_2 \in \Theta_2 \text{ and } U_2^{\psi_0}(\theta_2; \tilde{\theta}_2) > U_2^{\psi_0}(\theta_2; \theta_2)\}}} \frac{U_2^{\psi_0}(\theta_2; \theta_2^r) - U_2^{\psi_0}(\theta_2; \theta_2)}{\int_{\Theta_1 \setminus \Theta_1^*(\theta_2^r)} \int_{\Omega^{\psi}(\theta_2, \theta_1; \theta_2^r)} dH\left(\omega | (\theta_1, \theta_2)\right) dF(\theta_1)}.$$

Since the transfer rule varies across the cutoff θ_2^*, the agents' expected utility in the GS mechanism with $\psi = 0$ in Case 2 is different from that in Case 1, implying that B_1 and B_2 are different from A_1 and A_2, respectively. There are two steps to be checked in this case as well.

Step 2-1: If the buyer always reports the truth in the first stage, the seller has no incentive to tell a lie in the first stage in the GS mechanism with $\psi \geq \psi_2^*$.

Proof. Consider the seller of type θ_1. Recall in the proof of Claim 1 that the seller receives the following expected utility under truth-telling in Case 2:

$$\int_{\Theta_2} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2) + \int_{\Theta_1} \int_{\Theta_2^*(\theta_1)} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_1^e(\theta_1, \theta_2)\right) dF(\theta_2) dF(\theta_1)$$
$$- \int_{\Theta_1} \int_{\Theta_2^{**}} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^*)\right) dF(\theta_2) dF(\theta_1),$$

where $\tilde{u}_i^e(\theta_i, \theta_{-i}) = \int_\Omega \tilde{u}_i(\theta_i, \theta_{-i}; \omega) dH\left(\omega | (\theta_i, \theta_{-i})\right)$ and $\Theta_2^*(\theta_1) = \{\theta_2 \in \Theta_2 : x^*(\theta_1, \theta_2) = 1\}$, $\Theta_2^{**} = [\theta_2^*, \overline{\theta}]$. Note that the expected utility of the seller under truth-telling is identical to that in the GS mechanism with $\psi = 0$, which is $U_1^{\psi_0}(\theta_1; \theta_1)$. On the other hand, if trade occurs after the seller deviates to $\theta_1^r \neq \theta_1$ and the buyer's allocation payoff report is not consistent with the first-stage type reports (θ_1^r, θ_2), the seller must pay a penalty ψ according to the transfer rule t_1^{GS}. Therefore, the expected utility of the seller of type θ_1 when announcing θ_1^r becomes

$$\int_{\Theta_2 \setminus \Theta_2^*(\theta_1^r)} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2) + \int_{\Theta_2^*(\theta_1^r) \setminus \Theta_2^{**}} \tilde{u}_2^e(\theta_1^r, \theta_2) \left[1 - \int_{\Omega^{\psi}(\theta_1, \theta_2; \theta_1^r)} dH\left(\omega | (\theta_1, \theta_2)\right)\right] dF(\theta_2)$$
$$+ \int_{\Theta_2^{**}} g(\theta_1^r) \left[1 - \int_{\Omega^{\psi}(\theta_1, \theta_2; \theta_1^r)} dH\left(\omega | (\theta_1, \theta_2)\right)\right] dF(\theta_2) - \psi \int_{\Theta_2^*(\theta_1^r)} \int_{\Omega^{\psi}(\theta_1, \theta_2; \theta_1^r)} dH\left(\omega | (\theta_1, \theta_2)\right) dF(\theta_2),$$

where $\Omega^{\psi}(\theta_1, \theta_2; \theta_1^r) = \Big\{\omega \in \Omega : \tilde{u}_2(\theta_1, \theta_2; \omega) \in \widetilde{\Pi}_2(\theta_1, \theta_2) \text{ but } \tilde{u}_2(\theta_1, \theta_2; \omega) \notin \widetilde{\Pi}_2(\theta_1^r, \theta_2)\Big\}$, $\widetilde{\Pi}_2(\theta_1^r, \theta_2) = \{\hat{u}_2 \in \mathbb{R} :$ there exists a $\omega \in \Omega$ such that $h\left(\omega|(\theta_1^r, \theta_2)\right) > 0$ and $\tilde{u}_2(\theta_1^r, \theta_2; \omega) = \hat{u}_2\}$, and $h\left(\omega|(\theta_1^r, \theta_2)\right)$ is the probability density function of the residual uncertainty ω conditional on the type profile being (θ_1^r, θ_2). Note that the expected utility of the seller after deviation can be rewritten as follows:

$$U_1^{\psi_0}(\theta_1; \theta_1^r) - \psi \int_{\Theta_2^*(\theta_1^r)} \int_{\Omega^{\psi}(\theta_1,\theta_2;\theta_1^r)} dH\left(\omega|(\theta_1, \theta_2)\right) dF(\theta_2), \tag{10}$$

where

$$\begin{aligned} U_1^{\psi_0}(\theta_1; \theta_1^r) &= \int_{\Theta_2 \setminus \Theta_2^*(\theta_1^r)} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2) + \int_{\Theta_2^*(\theta_1^r)\setminus\Theta_2^{**}} \tilde{u}_2^e(\theta_1^r, \theta_2) \left[1 - \int_{\Omega^{\psi}(\theta_1,\theta_2;\theta_1^r)} dH\left(\omega|(\theta_1, \theta_2)\right)\right] dF(\theta_2) \\ &\quad + \int_{\Theta_2^{**}} g(\theta_1^r) \left[1 - \int_{\Omega^{\psi}(\theta_1,\theta_2;\theta_1^r)} dH\left(\omega|(\theta_1, \theta_2)\right)\right] dF(\theta_2) \end{aligned}$$

is the expected utility of the seller after deviating to θ_1^r in the GS mechanism with $\psi = 0$.

If $U_1^{\psi_0}(\theta_1; \theta_1) \geq U_1^{\psi_0}(\theta_1; \theta_1^r)$, the deviation is not profitable. If $U_1^{\psi_0}(\theta_1; \theta_1) < U_1^{\psi_0}(\theta_1; \theta_1^r)$, then, by Assumption 1, there exists a positive-measure set $\Theta_2^+ \subseteq \Theta_2$ such that (i) $\int_{\Omega^{\psi}(\theta_1,\theta_2;\theta_1^r)} dH\left(\omega|(\theta_1, \theta_2)\right) > 0$ for all $\theta_2 \in \Theta_2^+$, and (ii) $\int_{\Theta_2^+} x_2^*(\theta_1^r, \theta_2) dF(\theta_2) > 0$. In other words, the opposite of the coefficient of ψ in (10) is strictly positive, i.e., $\int_{\Theta_2^*(\theta_1^r)} \int_{\Omega^{\psi}(\theta_1,\theta_2;\theta_1^r)} dH\left(\omega|(\theta_1, \theta_2)\right) dF(\theta_2) > 0$.

Since $\psi \geq \psi_2^* = \max\{B_1, B_2\}$ where

$$B_1 = \sup_{\substack{\theta_1 \in [\underline{\theta}, \overline{\theta}] \\ \theta_1^r \in \{\tilde{\theta}_1 : \tilde{\theta}_1 \in \Theta_1 \text{ and } U_1^{\psi_0}(\theta_1; \tilde{\theta}_1) > U_1^{\psi_0}(\theta_1; \theta_1)\}}} \frac{U_1^{\psi_0}(\theta_1; \theta_1^r) - U_1^{\psi_0}(\theta_1; \theta_1)}{\int_{\Theta_2^*(\theta_1^r)} \int_{\Omega^{\psi}(\theta_1,\theta_2;\theta_1^r)} dH\left(\omega|(\theta_1, \theta_2)\right) dF(\theta_2)},$$

we obtain that, for each $\theta_1, \theta_1^r \in \Theta_1$ such that $U_1^{\psi_0}(\theta_1; \theta_1^r) > U_1^{\psi_0}(\theta_1; \theta_1)$,

$$\begin{aligned} & \psi \geq \frac{U_1^{\psi_0}(\theta_1; \theta_1^r) - U_1^{\psi_0}(\theta_1; \theta_1)}{\int_{\Theta_2^*(\theta_1^r)} \int_{\Omega^{\psi}(\theta_1,\theta_2;\theta_1^r)} dH\left(\omega|(\theta_1, \theta_2)\right) dF(\theta_2)} \\ \Rightarrow \; & \psi \int_{\Theta_2^*(\theta_1^r)} \int_{\Omega^{\psi}(\theta_1,\theta_2;\theta_1^r)} dH\left(\omega|(\theta_1, \theta_2)\right) dF(\theta_2) \geq U_1^{\psi_0}(\theta_1; \theta_1^r) - U_1^{\psi_0}(\theta_1; \theta_1), \end{aligned}$$

because the denominator is strictly positive. Then, we can compare the seller's expected utility under truth-telling and that after deviation:

$$\begin{aligned} & U_1^{\psi_0}(\theta_1; \theta_1^r) - \psi \int_{\Theta_2^*(\theta_1^r)} \int_{\Omega^{\psi}(\theta_1,\theta_2;\theta_1^r)} dH\left(\omega|(\theta_1, \theta_2)\right) dF(\theta_2) \\ \leq \; & U_1^{\psi_0}(\theta_1; \theta_1^r) - U_1^{\psi_0}(\theta_1; \theta_1^r) + U_1^{\psi_0}(\theta_1; \theta_1) \\ = \; & U_1^{\psi_0}(\theta_1; \theta_1), \end{aligned}$$

implying that the expected utility of the seller after deviating to θ_1^r is at most as high as that under truth-telling. Therefore, the seller has no incentive to deviate to θ_1^r in the GS mechanism with $\psi \geq \psi_2^*$. This completes the proof of Step 2-1. $\square$

Step 2-2: If the seller always reports the truth in the first stage, the buyer has no incentive to tell a lie in the first stage in the GS mechanism with $\psi \geq \psi_2^*$.

Proof. There are two subcases we shall consider.

1. Consider the buyer of type $\theta_2 < \theta_2^*$. Then, the buyer's expected utility under truth-telling is

$$\int_{\Theta_1^*(\theta_2)} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2)\right) dF(\theta_1) = 0,$$

where $\tilde{u}_2^e(\theta_1, \theta_2) = \int_\Omega \tilde{u}_2(\theta_1, \theta_2; \omega) dH\left(\omega|(\theta_1, \theta_2)\right)$ and $\Theta_1^*(\theta_2) = \{\theta_1 \in \Theta_1 : x^*(\theta_1, \theta_2) = 1\}$. Note that the expected utility of the buyer under truth-telling is identical to that in the GS mechanism with $\psi = 0$, which is $U_2^{\psi_0}(\theta_2; \theta_2)$. On the other hand, if no trade occurs after the buyer of type θ_2 deviates to $\theta_2^r < \theta_2^*$ and the seller's allocation payoff report cannot occur given the first-stage type reports (θ_1, θ_2^r), then the buyer must pay a penalty ψ according to the transfer rule t_2^{GS}. Therefore, the expected utility of the buyer of type θ_2 when announcing θ_2^r becomes

$$\int_{\Theta_1^*(\theta_2^r)} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^r)\right) dF(\theta_1) - \psi \int_{\Theta_1 \setminus \Theta_1^*(\theta_2^r)} \int_{\Omega^\psi(\theta_2, \theta_1; \theta_2^r)} dH\left(\omega|(\theta_1, \theta_2)\right) dF(\theta_1),$$

where $\Omega^\psi(\theta_2, \theta_1; \theta_2^r) = \left\{\omega \in \Omega : \tilde{u}_1(\theta_1, \theta_2; \omega) \in \widetilde{\Pi}_1(\theta_1, \theta_2) \text{ but } \tilde{u}_1(\theta_1, \theta_2; \omega) \notin \widetilde{\Pi}_1(\theta_1, \theta_2^r)\right\}$, $\widetilde{\Pi}_1(\theta_1, \theta_2^r) = \{\hat{u}_1 \in \mathbb{R} : \text{ there exists a } \omega \in \Omega \text{ such that } h\left(\omega|(\theta_1, \theta_2^r)\right) > 0 \text{ and } \tilde{u}_1(\theta_1, \theta_2^r; \omega) = \hat{u}_1\}$, and $h\left(\omega|(\theta_1, \theta_2^r)\right)$ is the probability density function of the residual uncertainty ω conditional on the type profile being (θ_1, θ_2^r). Note that the expected utility of the buyer after deviation can be rewritten as follows:

$$U_2^{\psi_0}(\theta_2; \theta_2^r) - \psi \int_{\Theta_1 \setminus \Theta_1^*(\theta_2^r)} \int_{\Omega^\psi(\theta_2, \theta_1; \theta_2^r)} dH\left(\omega|(\theta_1, \theta_2)\right) dF(\theta_1), \tag{11}$$

where $U_2^{\psi_0}(\theta_2; \theta_2^r) = \int_{\Theta_1^*(\theta_2^r)} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^r)\right) dF(\theta_1)$ is the expected utility of the buyer after deviating to $\theta_2^r < \theta_2^*$ in the GS mechanism with $\psi = 0$.

If $U_2^{\psi_0}(\theta_2; \theta_2) \geq U_2^{\psi_0}(\theta_2; \theta_2^r)$, the deviation is not profitable. If $U_2^{\psi_0}(\theta_2; \theta_2) < U_2^{\psi_0}(\theta_2; \theta_2^r)$, then, by Assumption 1, there exists a positive-measure set $\Theta_1^+ \subseteq$

Θ_1 such that (i) $\int_{\Omega^\psi(\theta_2,\theta_1;\theta_2^r)} dH(\omega|(\theta_1,\theta_2)) > 0$ for all $\theta_1 \in \Theta_1^+$, and (ii) $\int_{\Theta_1^+} x_1^*(\theta_1,\theta_2^r) dF(\theta_1) > 0$. In other words, the opposite of the coefficient of ψ in (11) is strictly positive, i.e.,

$$\int_{\Theta_1 \setminus \Theta_1^*(\theta_2^r)} \int_{\Omega^\psi(\theta_2,\theta_1;\theta_2^r)} dH(\omega|(\theta_1,\theta_2))\, dF(\theta_1) > 0.$$

Since $\psi \geq \psi_2^* = \max\{B_1, B_2\}$ where

$$B_2 = \sup_{\substack{\theta_2 \in [\underline{\theta},\overline{\theta}] \\ \theta_2^r \in \{\tilde{\theta}_2 : \tilde{\theta}_2 \in \Theta_2 \text{ and } U_2^{\psi_0}(\theta_2;\tilde{\theta}_2) > U_2^{\psi_0}(\theta_2;\theta_2)\}}} \frac{U_2^{\psi_0}(\theta_2;\theta_2^r) - U_2^{\psi_0}(\theta_2;\theta_2)}{\int_{\Theta_1 \setminus \Theta_1^*(\theta_2^r)} \int_{\Omega^\psi(\theta_2,\theta_1;\theta_2^r)} dH(\omega|(\theta_1,\theta_2))\, dF(\theta_1)},$$

we obtain that, for each $\theta_2, \theta_2^r \in [\underline{\theta}, \theta_2^*)$ such that $U_2^{\psi_0}(\theta_2;\theta_2^r) > U_2^{\psi_0}(\theta_2;\theta_2)$,

$$\begin{aligned} & \psi \geq \frac{U_2^{\psi_0}(\theta_2;\theta_2^r) - U_2^{\psi_0}(\theta_2;\theta_2)}{\int_{\Theta_1 \setminus \Theta_1^*(\theta_2^r)} \int_{\Omega^\psi(\theta_2,\theta_1;\theta_2^r)} dH(\omega|(\theta_1,\theta_2))\, dF(\theta_1)} \\ \Rightarrow\ & \psi \int_{\Theta_1 \setminus \Theta_1^*(\theta_2^r)} \int_{\Omega^\psi(\theta_2,\theta_1;\theta_2^r)} dH(\omega|(\theta_1,\theta_2))\, dF(\theta_1) \geq U_2^{\psi_0}(\theta_2;\theta_2^r) - U_2^{\psi_0}(\theta_2;\theta_2), \end{aligned}$$

because the denominator is strictly positive. Then, we can compare the buyer's expected utility under truth-telling and that after deviation:

$$\begin{aligned} & U_2^{\psi_0}(\theta_2;\theta_2^r) - \psi \int_{\Theta_1 \setminus \Theta_1^*(\theta_2^r)} \int_{\Omega^\psi(\theta_2,\theta_1;\theta_2^r)} dH(\omega|(\theta_1,\theta_2))\, dF(\theta_1) \\ \leq\ & U_2^{\psi_0}(\theta_2;\theta_2^r) - U_2^{\psi_0}(\theta_2;\theta_2^r) + U_2^{\psi_0}(\theta_2;\theta_2) \\ =\ & U_2^{\psi_0}(\theta_2;\theta_2), \end{aligned}$$

implying that the expected utility of the buyer after deviating to θ_2^r is at most as high as that under truth-telling. Therefore, the buyer has no incentive to deviate to $\theta_2^r < \theta_2^*$ in the GS mechanism with $\psi \geq \psi_2^*$.

Moreover, if the buyer deviates to $\theta_2^r \geq \theta_2^*$, trade always occurs and the expected utility of the buyer of type θ_2 when announcing θ_2^r becomes

$$\int_{\Theta_1} (\tilde{u}_2^e(\theta_1,\theta_2) - g(\theta_1))\, dF(\theta_1) = \int_{\Theta_1} (\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2^*))\, dF(\theta_1) < 0.$$

The equality follows because by the proof of Claim 1, we have $\int_{\Theta_1} g(\theta_1) dF(\theta_1) = \int_{\Theta_1} \tilde{u}_2^e(\theta_1,\theta_2^*) dF(\theta_1)$. The inequality follows because $\theta_2 < \theta_2^*$ implies $\tilde{u}_2^e(\theta_1,\theta_2) < \tilde{u}_2^e(\theta_1,\theta_2^*)$ for all $\theta_1 \in \Theta_1$. Recall that the buyer of type $\theta_2 < \theta_2^*$ always receives zero expected utility under truth-telling. Therefore, the buyer is never better off after a deviation to $\theta_2^r \geq \theta_2^*$ so that he has no incentive to deviate from truth-telling to $\theta_2^r \geq \theta_2^*$.

2. Consider the buyer of type $\theta_2 \geq \theta_2^*$. In this case, it is always efficient to trade the good regardless of the seller's type. Therefore, the expected utility of the buyer of type θ_2 under truth-telling is

$$\int_{\Theta_1} \left(\tilde{u}_2^e(\theta_1, \theta_2) - g(\theta_1)\right) dF(\theta_1) = \int_{\Theta_1} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^*)\right) dF(\theta_1).$$

The equality follows because by the proof of Claim 1, we have $\int_{\Theta_1} g(\theta_1) dF(\theta_1) = \int_{\Theta_1} \tilde{u}_2^e(\theta_1, \theta_2^*) dF(\theta_1)$. On the other hand, if the buyer deviates to $\theta_2^r \neq \theta_2$ such that $\theta_2^r \geq \theta_2^*$, then trade still occurs with probability one regardless of the seller's type. Thus, the expected utility of the buyer of type θ_2 after deviating to θ_2^r is

$$\int_{\Theta_1} \left(\tilde{u}_2^e(\theta_1, \theta_2) - g(\theta_1)\right) dF(\theta_1),$$

which is the same as the expected utility under truth-telling. Therefore, the buyer of type $\theta_2 \geq \theta_2^*$ has no incentive to deviate to $\theta_2^r \geq \theta_2^*$.

Moreover, if trade does not occur after the buyer of type θ_2 deviates to $\theta_2^r < \theta_2^*$ and the seller's allocation payoff report cannot occur given the first-stage type reports (θ_1, θ_2^r), the buyer must pay a penalty ψ according to the transfer rule t_2^{GS}. Therefore, the expected utility of the buyer of type θ_2 when announcing θ_2^r becomes

$$\int_{\Theta_1^*(\theta_2^r)} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^r)\right) dF(\theta_1) - \psi \int_{\Theta_1 \setminus \Theta_1^*(\theta_2^r)} \int_{\Omega^\psi(\theta_2, \theta_1; \theta_2^r)} dH\left(\omega | (\theta_1, \theta_2)\right) dF(\theta_1).$$

Note that the expected utility of the buyer after deviation can be rewritten as follows:

$$U_2^{\psi_0}(\theta_2; \theta_2^r) - \psi \int_{\Theta_1 \setminus \Theta_1^*(\theta_2^r)} \int_{\Omega^\psi(\theta_2, \theta_1; \theta_2^r)} dH\left(\omega | (\theta_1, \theta_2)\right) dF(\theta_1), \qquad (12)$$

where $U_2^{\psi_0}(\theta_2; \theta_2^r) = \int_{\Theta_1^*(\theta_2^r)} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^r)\right) dF(\theta_1)$ is the expected utility of the buyer after deviating to $\theta_2^r < \theta_2^*$ in the GS mechanism with $\psi = 0$.

If $U_2^{\psi_0}(\theta_2; \theta_2) \geq U_2^{\psi_0}(\theta_2; \theta_2^r)$, the deviation is not profitable. If $U_2^{\psi_0}(\theta_2; \theta_2) < U_2^{\psi_0}(\theta_2; \theta_2^r)$, then, by Assumption 1, there exists a positive-measure set $\Theta_1^+ \subseteq \Theta_1$ such that (i) $\int_{\Omega^\psi(\theta_2, \theta_1; \theta_2^r)} dH\left(\omega | (\theta_1, \theta_2)\right) > 0$ for all $\theta_1 \in \Theta_1^+$, and (ii) $\int_{\Theta_1^+} x_1^*(\theta_1, \theta_2^r) dF(\theta_1) > 0$. In other words, the opposite of the coefficient of ψ in (12),

$$\int_{\Theta_1 \setminus \Theta_1^*(\theta_2^r)} \int_{\Omega^\psi(\theta_2, \theta_1; \theta_2^r)} dH\left(\omega | (\theta_1, \theta_2)\right) dF(\theta_1),$$

is strictly positive.

Since $\psi \geq \psi_2^* = \max\{B_1, B_2\}$ where

$$B_2 = \sup_{\substack{\theta_2 \in [\underline{\theta}, \bar{\theta}] \\ \theta_2^r \in \{\tilde{\theta}_2 : \tilde{\theta}_2 \in \Theta_2 \text{ and } U_2^{\psi_0}(\theta_2; \tilde{\theta}_2) > U_2^{\psi_0}(\theta_2; \theta_2)\}}} \frac{U_2^{\psi_0}(\theta_2; \theta_2^r) - U_2^{\psi_0}(\theta_2; \theta_2)}{\int_{\Theta_1 \backslash \Theta_1^*(\theta_2^r)} \int_{\Omega^{\psi}(\theta_2, \theta_1; \theta_2^r)} dH\left(\omega | (\theta_1, \theta_2)\right) dF(\theta_1)},$$

we obtain that, for each $\theta_2 \geq \theta_2^*$ and each $\theta_2^r < \theta_2^*$ such that $U_2^{\psi_0}(\theta_2; \theta_2^r) > U_2^{\psi_0}(\theta_2; \theta_2)$,

$$\begin{aligned} & \psi \geq \frac{U_2^{\psi_0}(\theta_2; \theta_2^r) - U_2^{\psi_0}(\theta_2; \theta_2)}{\int_{\Theta_1 \backslash \Theta_1^*(\theta_2^r)} \int_{\Omega^{\psi}(\theta_2, \theta_1; \theta_2^r)} dH\left(\omega | (\theta_1, \theta_2)\right) dF(\theta_1)} \\ \Rightarrow \quad & \psi \int_{\Theta_1 \backslash \Theta_1^*(\theta_2^r)} \int_{\Omega^{\psi}(\theta_2, \theta_1; \theta_2^r)} dH\left(\omega | (\theta_1, \theta_2)\right) dF(\theta_1) \geq U_2^{\psi_0}(\theta_2; \theta_2^r) - U_2^{\psi_0}(\theta_2; \theta_2), \end{aligned}$$

because the denominator is strictly positive. Then, we can compare the buyer's expected utility under truth-telling and that after deviation:

$$\begin{aligned} & U_2^{\psi_0}(\theta_2; \theta_2^r) - \psi \int_{\Theta_1 \backslash \Theta_1^*(\theta_2^r)} \int_{\Omega^{\psi}(\theta_2, \theta_1; \theta_2^r)} dH\left(\omega | (\theta_1, \theta_2)\right) dF(\theta_1) \\ \leq \quad & U_2^{\psi_0}(\theta_2; \theta_2^r) - U_2^{\psi_0}(\theta_2; \theta_2^r) + U_2^{\psi_0}(\theta_2; \theta_2) \\ = \quad & U_2^{\psi_0}(\theta_2; \theta_2), \end{aligned}$$

implying that the expected utility of the buyer after deviating to θ_2^r is at most as high as that under truth-telling. Therefore, the buyer has no incentive to deviate to $\theta_2^r < \theta_2^*$ in the GS mechanism with $\psi \geq \psi_2^*$. This completes the proof of Step 2-2.

□

Set $\psi^* = \max\{\psi_1^*, \psi_2^*\}$. Then, the GS mechanism with $\psi \geq \psi^*$ satisfies BIC. This completes the proof of the sufficiency part.

9.4 Proof of Proposition 3

9.4.1 Proof of Step II

Proof. We establish two results here. We first show that in any two-stage mechanism $(\Theta, \Pi, x^*, t, \psi)$ satisfying BIC, EFF, and BB, there exists $\hat{g} : \Theta_1 \to \mathbb{R}$ such that for all $\theta_1 \in \Theta_1$ and all $\theta_2 \geq \theta_2^*$,

$$t_1^e(\theta_1, \theta_2; u_1, u_2) = -t_2^e(\theta_1, \theta_2; u_1, u_2) = \hat{g}(\theta_1),$$

where, for each agent $i \in \{1, 2\}$, $t_i^e(\theta_i, \theta_{-i}; u_i, u_{-i})$ is the expected monetary transfer of agent i conditional on the type profile being (θ_i, θ_{-i}) when both agents report truthfully in both stages. It is given by

$$\int_{\Omega} t_i\left(\theta_i, \theta_{-i}; u_i(x^*(\theta_i, \theta_{-i}), \theta_i, \theta_{-i}, \omega), u_{-i}(x^*(\theta_i, \theta_{-i}), \theta_i, \theta_{-i}, \omega)\right) dH\left(\omega|(\theta_i, \theta_{-i})\right).$$

Consider $\theta_2, \theta_2' \in \Theta_2$ such that $\theta_2 > \theta_2' \geq \theta_2^*$. By Assumption 0, we know that each agent has no incentive to deviate from the truth-telling in their payoff report in the second stage. Suppose that the seller always reports her true type in the first stage. If the buyer of type θ_2 reports his type truthfully in the first stage, then it is always efficient to trade and his allocation payoff is $u_2(x^*(\theta_1, \theta_2), \theta_1, \theta_2, \omega) = \tilde{u}_2(\theta_1, \theta_2; \omega)$. On the other hand, if he deviates to θ_2' in the first stage, then trade still occurs with probability one and the buyer's allocation payoff is $u_2(x^*(\theta_1, \theta_2'), \theta_1, \theta_2, \omega) = \tilde{u}_2(\theta_1, \theta_2; \omega)$. Since trade always occurs in both cases, the seller's allocation payoff is zero. Thus, to stop the buyer of type θ_2 from deviating to θ_2' in the first stage, the following BIC constraint must be satisfied:

$$\begin{aligned}
& \int_{\Theta_1} \tilde{u}_2^e(\theta_1, \theta_2) dF(\theta_1) + \int_{\Theta_1} t_2^e\left(\theta_1, \theta_2; 0, \tilde{u}_2(\theta_1, \theta_2; \omega)\right) dF(\theta_1) \\
\geq \; & \int_{\Theta_1} \tilde{u}_2^e(\theta_1, \theta_2) dF(\theta_1) + \int_{\Theta_1} \int_{\Omega} t_2\left(\theta_1, \theta_2'; 0, \tilde{u}_2(\theta_1, \theta_2; \omega)\right) dH\left(\omega|(\theta_1, \theta_2)\right) dF(\theta_1) \\
\Rightarrow \quad & \int_{\Theta_1} t_2^e\left(\theta_1, \theta_2; 0, \tilde{u}_2(\theta_1, \theta_2; \omega)\right) dF(\theta_1) \geq \int_{\Theta_1} \int_{\Omega} t_2\left(\theta_1, \theta_2'; 0, \tilde{u}_2(\theta_1, \theta_2; \omega)\right) dH\left(\omega|(\theta_1, \theta_2)\right) dF(\theta_1),
\end{aligned}$$

where $\tilde{u}_2^e(\theta_1, \theta_2) = \int_\Omega \tilde{u}_2(\theta_1, \theta_2; \omega) dH\left(\omega|(\theta_1, \theta_2)\right)$ and

$$t_2^e\left(\theta_1, \theta_2; 0, \tilde{u}_2(\theta_1, \theta_2; \omega)\right) = \int_{\Omega} t_2\left(\theta_1, \theta_2; 0, \tilde{u}_2(\theta_1, \theta_2; \omega)\right) dH\left(\omega|(\theta_1, \theta_2)\right).$$

Analogously, to stop the buyer of type θ_2' from deviating to θ_2, the following BIC constraint must be satisfied:

$$\begin{aligned}
& \int_{\Theta_1} \tilde{u}_2^e(\theta_1, \theta_2') dF(\theta_1) + \int_{\Theta_1} t_2^e\left(\theta_1, \theta_2'; 0, \tilde{u}_2(\theta_1, \theta_2'; \omega)\right) dF(\theta_1) \\
\geq \; & \int_{\Theta_1} \tilde{u}_2^e(\theta_1, \theta_2') dF(\theta_1) + \int_{\Theta_1} \int_{\Omega} t_2\left(\theta_1, \theta_2; 0, \tilde{u}_2(\theta_1, \theta_2'; \omega)\right) dH\left(\omega|(\theta_1, \theta_2')\right) dF(\theta_1) \\
\Rightarrow \quad & \int_{\Theta_1} t_2^e\left(\theta_1, \theta_2'; 0, \tilde{u}_2(\theta_1, \theta_2'; \omega)\right) dF(\theta_1) \geq \int_{\Theta_1} \int_{\Omega} t_2\left(\theta_1, \theta_2; 0, \tilde{u}_2(\theta_1, \theta_2'; \omega)\right) dH\left(\omega|(\theta_1, \theta_2')\right) dF(\theta_1),
\end{aligned}$$

where $\tilde{u}_2^e(\theta_1, \theta_2') = \int_\Omega \tilde{u}_2(\theta_1, \theta_2'; \omega) dH\left(\omega|(\theta_1, \theta_2')\right)$ and

$$t_2^e\left(\theta_1, \theta_2'; 0, \tilde{u}_2(\theta_1, \theta_2'; \omega)\right) = \int_{\Omega} t_2\left(\theta_1, \theta_2'; 0, \tilde{u}_2(\theta_1, \theta_2'; \omega)\right) dH\left(\omega|(\theta_1, \theta_2')\right).$$

Since, conditional on the type profile, the residual uncertainty affects the buyer's transfer above the cutoff only through his own second-stage report and, by Assumption 0, the buyer's transfer is independent of his own second-stage report, we obtain that the buyer's transfer is independent of the residual uncertainty and we write

$$\begin{aligned}
& \int_{\Theta_1} \int_{\Omega} t_2\left(\theta_1, \theta_2; 0, \tilde{u}_2(\theta_1, \theta_2'; \omega)\right) dH\left(\omega|(\theta_1, \theta_2')\right) dF(\theta_1) \\
= & \int_{\Theta_1} t_2\left(\theta_1, \theta_2; 0, \tilde{u}_2(\theta_1, \theta_2'; \omega)\right) dF(\theta_1) \\
= & \int_{\Theta_1} t_2\left(\theta_1, \theta_2; 0, \tilde{u}_2(\theta_1, \theta_2; \omega)\right) dF(\theta_1) \\
= & \int_{\Theta_1} \int_{\Omega} t_2\left(\theta_1, \theta_2; 0, \tilde{u}_2(\theta_1, \theta_2; \omega)\right) dH\left(\omega|(\theta_1, \theta_2)\right) dF(\theta_1) = \int_{\Theta_1} t_2^e\left(\theta_1, \theta_2; 0, \tilde{u}_2(\theta_1, \theta_2; \omega)\right) dF(\theta_1),
\end{aligned}$$

where the second equality follows because the buyer's transfer is independent of his own second-stage report. As a result, the BIC constraint which stops the buyer of type θ_2' from deviating to θ_2 can be rewritten as follows:

$$\int_{\Theta_1} t_2^e\left(\theta_1, \theta_2'; 0, \tilde{u}_2(\theta_1, \theta_2'; \omega)\right) dF(\theta_1) \geq \int_{\Theta_1} t_2^e\left(\theta_1, \theta_2; 0, \tilde{u}_2(\theta_1, \theta_2; \omega)\right) dF(\theta_1). \quad (13)$$

Analogously, the BIC constraint which stops the buyer of type θ_2 from deviating to θ_2' can be rewritten as follows:

$$\int_{\Theta_1} t_2^e\left(\theta_1, \theta_2; 0, \tilde{u}_2(\theta_1, \theta_2; \omega)\right) dF(\theta_1) \geq \int_{\Theta_1} t_2^e\left(\theta_1, \theta_2'; 0, \tilde{u}_2(\theta_1, \theta_2'; \omega)\right) dF(\theta_1). \quad (14)$$

Combining (13) and (14), we obtain

$$\int_{\Theta_1} t_2^e\left(\theta_1, \theta_2; 0, \tilde{u}_2(\theta_1, \theta_2; \omega)\right) dF(\theta_1) = \int_{\Theta_1} t_2^e\left(\theta_1, \theta_2'; 0, \tilde{u}_2(\theta_1, \theta_2'; \omega)\right) dF(\theta_1)$$

for all $\theta_2, \theta_2' \in [\theta_2^*, \overline{\theta}]$, implying that the buyer's transfer above the cutoff θ_2^* is independent of his first-stage report. Once again, we know from Assumption 0 that the buyer's transfer is independent of his own second-stage report. In addition, for the buyer of type $\theta_2 \geq \theta_2^*$, it is always efficient to trade and the seller's allocation payoff is always zero. So, the transfer for the buyer of type $\theta_2 \geq \theta_2^*$ depends only on the seller's first-stage report. Hence, we can define $\hat{g} : \Theta_1 \to \mathbb{R}$ such that for all $\theta_1 \in \Theta_1$ and all $\theta_2 \geq \theta_2^*$, we have

$$t_2^e(\theta_1, \theta_2; 0, \tilde{u}_2(\theta_1, \theta_2; \omega)) = -\hat{g}(\theta_1).$$

Since the two-stage mechanism $(\Theta, \Pi, x^*, t, \psi)$ satisfies BB, the total expected transfer equals zero on the equilibrium path. So,

$$t_1^e(\theta_1, \theta_2; 0, \tilde{u}_2(\theta_1, \theta_2; \omega)) = -t_2^e(\theta_1, \theta_2; 0, \tilde{u}_2(\theta_1, \theta_2; \omega)) = \hat{g}(\theta_1).$$

We next show that in any two-stage mechanism $(\Theta, \Pi, x^*, t, \psi)$ satisfying EFF, BB as well as the buyer's IIR, the following inequality hold:

$$\int_{\Theta_1} \hat{g}(\theta_1) dF(\theta_1) \leq \int_{\Theta_1} \tilde{u}_2^e(\theta_1, \theta_2^*) dF(\theta_1).$$

Consider the buyer of type $\theta_2 \geq \theta_2^*$. Then his IIR constraint is the following:

$$\int_{\Theta_1} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \hat{g}(\theta_1)\right) dF(\theta_1) \geq 0 \Rightarrow \int_{\Theta_1} \hat{g}(\theta_1) dF(\theta_1) \leq \int_{\Theta_1} \tilde{u}_2^e(\theta_1, \theta_2) dF(\theta_1),$$

where $\tilde{u}_2^e(\theta_1, \theta_2) = \int_\Omega \tilde{u}_2(\theta_1, \theta_2; \omega) dH\left(\omega | (\theta_1, \theta_2)\right)$. Since the above inequality is true for any $\theta_2 \geq \theta_2^*$, we can set $\theta_2 = \theta_2^*$ so that

$$\int_{\Theta_1} \hat{g}(\theta_1) dF(\theta_1) \leq \int_{\Theta_1} \tilde{u}_2^e(\theta_1, \theta_2^*) dF(\theta_1).$$

This completes the proof of Step II. □

9.4.2 Proof of Step III

Proof. In any two-stage mechanism $(\Theta, \Pi, x^*, t, \psi)$, the seller's ex ante expected utility on the equilibrium path is computed as follows:

$$\int_{\Theta_1} \int_{\Theta_2 \setminus \Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1) + \int_{\Theta_1} \int_{\Theta_2} t_1^e(\theta_1, \theta_2; u_1, u_2) dF(\theta_2) dF(\theta_1),$$

where $t_1^e(\theta_1, \theta_2; u_1, u_2)$ is the expected monetary transfer of the seller conditional on the type profile being (θ_1, θ_2) when both agents report truthfully in both stages, that is,

$$t_1^e(\theta_1, \theta_2; u_1, u_2) = \int_\Omega t_1\left(\theta_1, \theta_2; u_1(x^*(\theta_1, \theta_2), \theta_1, \theta_2, \omega), u_2(x^*(\theta_1, \theta_2), \theta_1, \theta_2, \omega)\right) dH\left(\omega | (\theta_1, \theta_2)\right),$$

and for each $\theta_1 \in \Theta_1$,

$$\Theta_2^*(\theta_1) = \begin{cases} \{\overline{\theta}\} & \text{if } \{\theta_2 \in \Theta_2 : x^*(\theta_1, \theta_2) = 1\} = \emptyset; \\ \{\theta_2 \in \Theta_2 : x^*(\theta_1, \theta_2) = 1\} & \text{otherwise,} \end{cases}$$

It follows from Step II that the seller's equilibrium transfer when $\theta_2 \geq \theta_2^*$ depends only on the seller's type. Then, the seller's ex ante expected utility on the equilibrium path is rewritten as follows:

$$\int_{\Theta_1} \int_{\Theta_2 \setminus \Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1) + \int_{\Theta_1} \int_{\Theta_2 \setminus \Theta_2^{**}} t_1^e(\theta_1, \theta_2; u_1, u_2) dF(\theta_2) dF(\theta_1)$$
$$+ \int_{\Theta_1} \int_{\Theta_2^{**}} \hat{g}(\theta_1) dF(\theta_2) dF(\theta_1).$$

Changing the order of integration, we rewrite the seller's ex ante expected utility as follows:

$$\begin{aligned}
&\int_{\Theta_1}\int_{\Theta_2\setminus\Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2)dF(\theta_1) + \int_{\Theta_2\setminus\Theta_2^{**}}\int_{\Theta_1} t_1^e(\theta_1,\theta_2;u_1,u_2)dF(\theta_1)dF(\theta_2) \\
&+\int_{\Theta_1}\int_{\Theta_2^{**}} \hat{g}(\theta_1)dF(\theta_2)dF(\theta_1) \\
\leq\ &\int_{\Theta_1}\int_{\Theta_2\setminus\Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2)dF(\theta_1) + \int_{\Theta_2\setminus\Theta_2^{**}}\int_{\Theta_1^*(\theta_2)} \tilde{u}_2^e(\theta_1,\theta_2)dF(\theta_1)dF(\theta_2) \\
&+\int_{\Theta_1}\int_{\Theta_2^{**}} \tilde{u}_2^e(\theta_1,\theta_2^*)dF(\theta_2)dF(\theta_1),
\end{aligned}$$

where the inequality follows because Step I requires that for any $\theta_2 < \theta_2^*$,

$$\int_{\Theta_1} t_1^e(\theta_1,\theta_2;u_1,u_2)dF(\theta_1) \leq \int_{\Theta_1^*(\theta_2)} \tilde{u}_2^e(\theta_1,\theta_2)dF(\theta_1),$$

and Step II requires $\int_{\Theta_1}\hat{g}(\theta_1)dF(\theta_1) \leq \int_{\Theta_1}\tilde{u}_2^e(\theta_1,\theta_2^*)dF(\theta_1)$.

Changing the order of integration, we rewrite the upper bound of the seller's maximum ex ante expected utility as follows:

$$\begin{aligned}
&\int_{\Theta_1}\int_{\Theta_2\setminus\Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2)dF(\theta_1) + \int_{\Theta_1}\int_{\Theta_2^*(\theta_1)\setminus\Theta_2^{**}} \tilde{u}_2^e(\theta_1,\theta_2)dF(\theta_2)dF(\theta_1) \\
&+\int_{\Theta_1}\int_{\Theta_2^{**}} \tilde{u}_2^e(\theta_1,\theta_2^*)dF(\theta_2)dF(\theta_1).
\end{aligned}$$

After rearrangement, we further rewrite the upper bound of the seller's maximum ex ante expected utility as follows:

$$\begin{aligned}
&\int_{\Theta_1}\int_{\Theta_2} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2)dF(\theta_1) + \int_{\Theta_1}\int_{\Theta_2^*(\theta_1)} \left(\tilde{u}_2^e(\theta_1,\theta_2)-\tilde{u}_1^e(\theta_1,\theta_2)\right)dF(\theta_2)dF(\theta_1) \\
&-\int_{\Theta_1}\int_{\Theta_2^{**}} \left(\tilde{u}_2^e(\theta_1,\theta_2)-\tilde{u}_2^e(\theta_1,\theta_2^*)\right)dF(\theta_2)dF(\theta_1),
\end{aligned}$$

which is identical to (4). This completes the proof of Step III. □

9.4.3 Proof of Step IV

Proof. If both agents report their type and allocation payoff truthfully, the proof of Claim 1 allows us to obtain the following interim expected utility for the seller of type θ_1 after participating in the GS mechanism $(\Theta,\Pi,x^*,t^{GS},\psi)$:

$$\begin{aligned}
U_1^{GS}(\theta_1) \;=\;& U_1^O(\theta_1) + \int_{\Theta_1}\int_{\Theta_2^*(\theta_1)} \left(\tilde{u}_2^e(\theta_1,\theta_2)-\tilde{u}_1^e(\theta_1,\theta_2)\right)dF(\theta_2)dF(\theta_1) \\
&-\int_{\Theta_1}\int_{\Theta_2^{**}} \left(\tilde{u}_2^e(\theta_1,\theta_2)-\tilde{u}_2^e(\theta_1,\theta_2^*)\right)dF(\theta_2)dF(\theta_1),
\end{aligned}$$

where $U_1^O(\theta_1)$ denotes the outside option utility for the seller of type θ_1 and is defined as follows:

$$U_1^O(\theta_1) = \int_{\Theta_2} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2),$$

and $\tilde{u}_i^e(\theta_i, \theta_{-i}) = \int_\Omega \tilde{u}_i(\theta_i, \theta_{-i}; \omega) dH(\omega|(\theta_i, \theta_{-i}))$ for each agent $i \in \{1, 2\}$ and each type profile $(\theta_i, \theta_{-i}) \in \Theta$. Then, we compute the seller's ex ante expected utility as follows:

$$\begin{aligned}\int_{\Theta_1} U_1^{GS}(\theta_1) dF(\theta_1) &= \int_{\Theta_1}\int_{\Theta_2} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1) \\ &\quad + \int_{\Theta_1}\int_{\Theta_2^*(\theta_1)} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_1^e(\theta_1, \theta_2)\right) dF(\theta_2) dF(\theta_1) \\ &\quad - \int_{\Theta_1}\int_{\Theta_2^{**}} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^*)\right) dF(\theta_2) dF(\theta_1).\end{aligned}$$

This implies that the seller's ex ante expected utility in the GS mechanism is identical to the upper bound in (4). This completes the proof of Step IV. □

9.5 Proof of Claim 2

Proof. The proof is completed by two steps.

Step 1: Condition (i) of Assumption 1 is always satisfied, that is, for each agent $i \in \{1, 2\}$, each $\theta_i, \theta_i^r \in \Theta_i$ such that $\theta_i \neq \theta_i^r$ and each $\theta_j \in \Theta_j$ where $j \neq i$, we have $\int_{\Omega^\psi(\theta_i, \theta_j; \theta_i^r)} dH(\omega|(\theta_i, \theta_j)) > 0$.

Proof. For each agent $i \in \{1, 2\}$, each $\theta_i, \theta_i^r \in \Theta_i$ such that $\theta_i \neq \theta_i^r$, and each $\theta_j \in \Theta_j$ where $j \neq i$, the set of feasible allocation payoffs of agent $j \neq i$ given the true type profile (θ_i, θ_j) is

$$\begin{aligned}\widetilde{\Pi}_j(\theta_i, \theta_j) &= \{\hat{u}_j \in \mathbb{R} : \text{ there exists } \omega \in \Omega \text{ s.t. } h(\omega|(\theta_i, \theta_j)) > 0 \text{ and } \tilde{u}_j(\theta_i, \theta_j; \omega) = \hat{u}_j\} \\ &= \{\hat{u}_j \in \mathbb{R} : \hat{u}_j = (1 + \omega_j)(\theta_j + \gamma_j\theta_i) \text{ for some } \omega_j \in [-0.1, 0.1]\} \\ &= [0.9(\theta_j + \gamma_j\theta_i), 1.1(\theta_j + \gamma_j\theta_i)].\end{aligned}$$

Similarly, the set of feasible allocation payoffs of agent j given the deviation (θ_i^r, θ_j) becomes

$$\begin{aligned}\widetilde{\Pi}_j(\theta_i^r, \theta_j) &= \{\hat{u}_j \in \mathbb{R} : \text{ there exists } \omega \in \Omega \text{ s.t. } h(\omega|(\theta_i^r, \theta_j)) > 0 \text{ and } \tilde{u}_j(\theta_i^r, \theta_j; \omega) = \hat{u}_j\} \\ &= \{\hat{u}_j \in \mathbb{R} : \hat{u}_j = (1 + \omega_j)(\theta_j + \gamma_j\theta_i^r) \text{ where } \omega_j \in [-0.1, 0.1]\} \\ &= [0.9(\theta_j + \gamma_j\theta_i^r), 1.1(\theta_j + \gamma_j\theta_i^r)].\end{aligned}$$

Suppose $\theta_i^r > \theta_i$ without loss of generality. Then, there exist some allocation payoffs of agent j such that they are feasible given the true type profile (θ_i, θ_j), but they are not feasible given the deviation (θ_i^r, θ_j), which is,

$$\begin{cases} [0.9(\theta_j + \gamma_j \theta_i), 1.1(\theta_j + \gamma_j \theta_i)] & \text{if } 1.1(\theta_j + \gamma_j \theta_i) < 0.9(\theta_j + \gamma_j \theta_i^r), \\ [0.9(\theta_j + \gamma_j \theta_i), 0.9(\theta_j + \gamma_j \theta_i^r)) & \text{otherwise.} \end{cases}$$

As a result, the set of residual uncertainties under which agent j's allocation payoffs are feasible given the true type profile (θ_i, θ_j), but not feasible given the deviation (θ_i^r, θ_j), is as follows:

$$\Omega^{\psi}(\theta_i, \theta_j; \theta_i^r) = \begin{cases} [-0.1, 0.1]^2 & \text{if } 1.1(\theta_j + \gamma_j \theta_i) < 0.9(\theta_j + \gamma_j \theta_i^r), \\ \left\{(\omega_i, \omega_j) : \omega_i \in [-0.1, 0.1] \text{ and } \omega_j \in \left[-0.1, \frac{0.9(\theta_j + \gamma_j \theta_i^r)}{\theta_j + \gamma_j \theta_i} - 1\right)\right\} & \text{otherwise.} \end{cases}$$

Therefore, we obtain $\int_{\Omega^{\psi}(\theta_i, \theta_j; \theta_i^r)} dH\left(\omega | (\theta_i, \theta_j)\right) > 0$ for each $i \in \{1, 2\}$, each $\theta_i, \theta_i^r \in \Theta_i$ such that $\theta_i \neq \theta_i^r$, and each $\theta_j \in \Theta_j$ where $j \neq i$, implying that Condition (i) in Assumption 1 is always satisfied in Example 1. This completes the proof of Step 1. □

Step 2: Condition (ii) of Assumption 1 is satisfied, that is, for each agent $i \in \{1, 2\}$ and $\theta_i, \theta_i^r \in \Theta_i$, if $U_i^{\psi_0}(\theta_i; \theta_i^r) > U_i^{\psi_0}(\theta_i; \theta_i)$, there exists a positive-measure set $\Theta_j^+ \subseteq \Theta_j$ where $j \neq i$ such that $\int_{\Theta_j^+} x_j^*(\theta_i^r, \theta_j) dF(\theta_j) > 0$.

Proof. We prove this by its contrapositive: for each $i \in \{1, 2\}$ and $\theta_i, \theta_i^r \in \Theta_i$, if $\int_{\Theta_j^+} x_j^*(\theta_i^r, \theta_j) dF(\theta_j) = 0$ for all subsets $\Theta_j^+ \subseteq \Theta_j$, then $U_i^{\psi_0}(\theta_i; \theta_i^r) \leq U_i^{\psi_0}(\theta_i; \theta_i)$ must be satisfied. In words, if the good is never allocated to agent $j \neq i$ after agent i deviates to θ_i^r, or equivalently, the good is always allocated to agent i, then agent i has no incentive to deviate. Recall that there are two cases in Example 1: Case (i): $0 < \gamma_2 \leq \gamma_1 < 1$ and Case (ii): $0 < \gamma_1 < \gamma_2 < 1$.

Case (i): $0 < \gamma_2 \leq \gamma_1 < 1$

There are two steps to be checked in this case.

Step 2-1-1: For each $\theta_1, \theta_1^r \in \Theta_1$, if $\int_{\Theta_2^+} x_2^*(\theta_1^r, \theta_2) dF(\theta_2) = 0$ for all subsets $\Theta_2^+ \subseteq \Theta_2$, then it follows that $U_1^{\psi_0}(\theta_1; \theta_1^r) \leq U_1^{\psi_0}(\theta_1; \theta_1)$.

Proof. Suppose that $\int_{\Theta_2^+} x_2^*(\theta_1^r, \theta_2) dF(\theta_2) = 0$ for all subsets $\Theta_2^+ \subseteq \Theta_2$, that is, the seller always keeps the good after she deviates to θ_1^r. Consider the seller of type θ_1. The seller obtains the following expected utility under truth-telling in the GS mechanism with $\psi = 0$:

$$U_1^{\psi_0}(\theta_1; \theta_1) = \int_{\Theta_2^*(\theta_1)} \tilde{u}_2^e(\theta_1, \theta_2) dF(\theta_2) + \int_{\Theta_2 \setminus \Theta_2^*(\theta_1)} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2),$$

where $\Theta_2^*(\theta_1) = \{\theta_2 \in \Theta_2 : x^*(\theta_1, \theta_2) = 1\}$ and $\tilde{u}_i^e(\theta_i, \theta_{-i}) = \int_\Omega \tilde{u}_i(\theta_i, \theta_{-i}; \omega) dH(\omega | (\theta_i, \theta_{-i}))$ for each agent $i \in \{1, 2\}$ and each type profile $(\theta_i, \theta_{-i}) \in \Theta$. On the other hand, if the seller deviates to θ_1^r, she always keeps the good and her expected utility becomes the following:

$$U_1^{\psi_0}(\theta_1; \theta_1^r) = \int_{\Theta_2} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2).$$

Recall in the proof of Proposition 1 that the seller's expected utility under truth-telling is at least as high that from the outside option which is $\int_{\Theta_2} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2)$. Since the seller's expected utility after deviation is equivalent to $\int_{\Theta_2} \tilde{u}_1^e(\theta_1, \theta_2) dF(\theta_2)$, we conclude that the seller will never be better off after such a deviation so that she has no incentive to deviate from truth-telling in the GS mechanism with $\psi = 0$. This completes the proof of Step 2-1-1. □

Step 2-1-2: For each $\theta_2, \theta_2^r \in \Theta_2$, if $\int_{\Theta_1^+} x_1^*(\theta_1, \theta_2^r) dF(\theta_1) = 0$ for all subsets $\Theta_1^+ \subseteq \Theta_1$, then it follows that $U_2^{\psi_0}(\theta_2; \theta_2^r) \leq U_2^{\psi_0}(\theta_2; \theta_2)$.

Proof. Suppose that $\int_{\Theta_1^+} x_1^*(\theta_1, \theta_2^r) dF(\theta_1) = 0$ for all subsets $\Theta_1^+ \subseteq \Theta_1$, or equivalently, $\int_{\Theta_1^+} x_2^*(\theta_1, \theta_2^r) dF(\theta_1) = 1$ for all positive-measure subsets $\Theta_1^+ \subseteq \Theta_1$. Recall that in Case (i), we have $\int_\Theta x_2^*(\theta_1, \theta_2^r) dF(\theta_1) < 1$ for all $\theta_2^r < \overline{\theta}$. Hence, there is only one possible value of θ_2^r which could satisfy $\int_{\Theta_1^+} x_2^*(\theta_1, \theta_2^r) dF(\theta_1) = 1$ for all positive-measure subsets $\Theta_1^+ \subseteq \Theta_1$, that is, $\theta_2^r = \overline{\theta}$. Consider the buyer of type θ_2. The buyer receives the following expected utility under truth-telling in the GS mechanism with $\psi = 0$:

$$U_2^{\psi_0}(\theta_2; \theta_2) = \int_{\Theta_1^*(\theta_2)} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2)\right) dF(\theta_1) = 0,$$

where $\Theta_1^*(\theta_2) = \{\theta_1 \in \Theta_1 : x^*(\theta_1, \theta_2) = 1\}$. On the other hand, if the buyer deviates to $\theta_2^r = \overline{\theta}$, his expected utility becomes the following:

$$U_2^{\psi_0}(\theta_2; \theta_2^r) = \int_{\Theta_1^*(\overline{\theta})} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \overline{\theta})\right) dF(\theta_1) = \int_{\Theta_1} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \overline{\theta})\right) dF(\theta_1) < 0,$$

The second equality follows because we assume the buyer receives the good with probability one after he deviates to $\overline{\theta}$, and the inequality follows because $\theta_2 < \overline{\theta}$ implies $\tilde{u}_2^e(\theta_1, \theta_2) < \tilde{u}_2^e(\theta_1, \overline{\theta})$ for all $\theta_1 \in \Theta_1$. As a result, we obtain $U_2^{\psi_0}(\theta_2; \theta_2^r) < U_2^{\psi_0}(\theta_2; \theta_2)$. This completes the proof the Step 2-1-2. □

We complete the proof for Case (i) and proceed to Case (ii).

Case (ii): $0 < \gamma_1 < \gamma_2 < 1$

We first argue that for the seller, there does not exist $\theta_1^r \in \Theta_1$ such that $\int_{\Theta_2^+} x_2^*(\theta_1^r, \theta_2) dF(\theta_2) = 0$ for all subsets $\Theta_2^+ \subseteq \Theta_2$ in Case (ii). Recall that in Case (ii), $\theta_2^* = (1-\gamma_2)/(1-\gamma_1) < \overline{\theta}$ and the good is always allocated to the buyer whenever the buyer reports a type higher than the cutoff θ_2^*. Hence, for all $\theta_1^r \in \Theta_1$, we have $\int_{\Theta_2} x_2^*(\theta_1^r, \theta_2) dF(\theta_2) > 0$, implying that $\int_{\Theta_2^+} x_2^*(\theta_1^r, \theta_2) dF(\theta_2) = 0$ for all subsets $\Theta_2^+ \subseteq \Theta_2$ is impossible. As a result, it is sufficient to check whether condition (ii) holds for the buyer.

Step 2-2: For each $\theta_2, \theta_2^r \in \Theta_2$, if $\int_{\Theta_1^+} x_1^*(\theta_1, \theta_2^r) dF(\theta_1) = 0$ for all subsets $\Theta_1^+ \subseteq \Theta_1$, then it follows that $U_2^{\psi_0}(\theta_2; \theta_2^r) \leq U_2^{\psi_0}(\theta_2; \theta_2)$.

Proof. Suppose that $\int_{\Theta_1^+} x_1^*(\theta_1, \theta_2^r) dF(\theta_1) = 0$ for all subsets $\Theta_1^+ \subseteq \Theta_1$. That is, the good is never allocated to the seller after the buyer deviates, or equivalently, the buyer always receives the good. As a result, $\theta_2^r \geq \theta_2^*$ must hold. There are two subcases, depending on the value of the true type θ_2.

We first consider the buyer of type $\theta_2 < \theta_2^*$. The buyer receives the following expected utility under truth-telling in the GS mechanism with $\psi = 0$:

$$U_2^{\psi_0}(\theta_2; \theta_2) = \int_{\Theta_1^*(\theta_2)} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2)\right) dF(\theta_1) = 0,$$

where $\Theta_1^*(\theta_2) = \{\theta_1 \in \Theta_1 : x^*(\theta_1, \theta_2) = 1\}$. On the other hand, if the buyer deviates to $\theta_2^r \geq \theta_2^*$, his expected utility becomes the following:

$$U_2^{\psi_0}(\theta_2; \theta_2^r) = \int_{\Theta_1} \left(\tilde{u}_2^e(\theta_1, \theta_2) - g(\theta_1)\right) dF(\theta_1) = \int_{\Theta_1} \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^*)\right) dF(\theta_1) < 0.$$

The second equality follows because by the proof of Claim 1, we have $\int_{\Theta_1} g(\theta_1) dF(\theta_1) = \int_{\Theta_1} \tilde{u}_2^e(\theta_1, \theta_2^*) dF(\theta_1)$, and the inequality follows because $\theta_2 < \theta_2^*$ implies $\tilde{u}_2^e(\theta_1, \theta_2) < \tilde{u}_2^e(\theta_1, \theta_2^*)$ for all $\theta_1 \in \Theta_1$. Therefore, we obtain $U_2^{\psi_0}(\theta_2; \theta_2^r) \leq U_2^{\psi_0}(\theta_2; \theta_2)$ in this case.

Next, we consider the buyer of type $\theta_2 \geq \theta_2^*$. The buyer receives the following expected utility under truth-telling in the GS mechanism with $\psi = 0$:

$$U_2^{\psi_0}(\theta_2; \theta_2) = \int_{\Theta_1} \left(\tilde{u}_2^e(\theta_1, \theta_2) - g(\theta_1)\right) dF(\theta_1).$$

On the other hand, if the buyer deviates to $\theta_2^r \geq \theta_2^*$, his expected utility becomes the following:

$$U_2^{\psi_0}(\theta_2; \theta_2^r) = \int_{\Theta_1} \left(\tilde{u}_2^e(\theta_1, \theta_2) - g(\theta_1)\right) dF(\theta_1),$$

which is identical to $U_2^{\psi_0}(\theta_2; \theta_2)$. Therefore, we conclude that $U_2^{\psi_0}(\theta_2; \theta_2^r) \leq U_2^{\psi_0}(\theta_2; \theta_2)$ is satisfied. This completes the proof of Step 2-2. □

Therefore, we obtain that if $U_i^{\psi_0}(\theta_i;\theta_i^r) > U_i^{\psi_0}(\theta_i;\theta_i)$, there exists a positive-measure set $\Theta_j^+ \subseteq \Theta_j$ with $j \neq i$ such that Condition (ii) in Assumption 1 is satisfied. This completes the proof of Step 2. $\square$

We conclude that Assumption 1 is satisfied in Example 1. This completes the proof of this claim. $\square$

9.6 Proof of Proposition 4

Proof. We first prove the sufficiency of Assumption 2. By Lemma 2, we divide the proof into the following two cases. In Case 1 where $\theta_2^* = \overline{\theta}$, it follows from Proposition 1 that the GS mechanism satisfies the seller's IIR. In Case 2 where $\theta_2^* \in (\underline{\theta}, \overline{\theta})$, it is clear from Claim 1 that the GS mechanism satisfies the seller's IIR constraint if the expression for C, which is identified in (3), is nonnegative. That is,

$$\begin{aligned}
&\int_{\Theta_1}\int_{\Theta_2^*(\theta_1)} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_1^e(\theta_1,\theta_2)\right) dF(\theta_2)dF(\theta_1) \\
- &\int_{\Theta_1}\int_{\Theta_2^{**}} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2^*)\right) dF(\theta_2)dF(\theta_1) \geq 0.
\end{aligned}$$

This is identical to Assumption 2. Therefore, we complete the proof of the sufficiency part.

We now prove the necessity. Suppose, on the contrary, that Assumption 2 is violated. Recall that Assumption 2 is always satisfied in Case 1: $\theta_2^* = \overline{\theta}$. Then Assumption 2 must be violated in Case 2: $\theta_2^* \in (\underline{\theta}, \overline{\theta})$. If both agents report their type and allocation payoff truthfully, the proof of Step IV of Proposition 3 allows us to obtain the following ex ante expected utility for the seller after participating in the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$:

$$\begin{aligned}
&\int_{\Theta_1} U_1^O(\theta_1)dF(\theta_1) + \int_{\Theta_1}\int_{\Theta_2^*(\theta_1)} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_1^e(\theta_1,\theta_2)\right) dF(\theta_2)dF(\theta_1) \\
&- \int_{\Theta_1}\int_{\Theta_2^{**}} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2^*)\right) dF(\theta_2)dF(\theta_1),
\end{aligned}$$

where $U_1^O(\theta_1) = \int_{\Theta_2} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2)$ is the outside option utility for the seller of type θ_1 and $\tilde{u}_i^e(\theta_i,\theta_{-i}) = \int_\Omega \tilde{u}_i(\theta_i,\theta_{-i};\omega)dH\left(\omega|(\theta_i,\theta_{-i})\right)$ for each agent $i \in \{1,2\}$ and each $(\theta_i,\theta_{-i}) \in \Theta$. As we assume that Assumption 2 is violated, the sum of the last two terms of the above expression is negative, implying that the seller's ex ante expected utility after participating in the mechanism is lower than her outside option utility. Hence, the seller's ex ante individual rationality constraint is violated in the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$. Since IIR implies ex ante individual rationality constraint, we further conclude that the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ violates the seller's IIR constraint. This completes the proof of the proposition. $\square$

9.7 Proof of Lemma 3

Proof. Recall our Assumption 2 says that

$$\begin{aligned}&\int_{\Theta_1}\int_{\Theta_2^*(\theta_1)} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_1^e(\theta_1,\theta_2)\right) dF(\theta_2)dF(\theta_1)\\ &-\int_{\Theta_1}\int_{\Theta_2^{**}} \left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2^*)\right) dF(\theta_2)dF(\theta_1) \geq 0,\end{aligned}$$

where for each agent $i \in \{1,2\}$ and each type profile $(\theta_i, \theta_{-i}) \in \Theta$,

$$\tilde{u}_i^e(\theta_i,\theta_{-i}) = \int_{\Omega} \tilde{u}_i(\theta_i,\theta_{-i};\omega) dH\left(\omega|(\theta_i,\theta_{-i})\right),$$

and for each $\theta_1 \in \Theta_1$,

$$\Theta_2^*(\theta_1) = \begin{cases} \{\overline{\theta}\} & \text{if } \{\theta_2 \in \Theta_2 : x^*(\theta_1,\theta_2) = 1\} = \emptyset \\ \{\theta_2 \in \Theta_2 : x^*(\theta_1,\theta_2) = 1\} & \text{otherwise,}\end{cases}$$

$\Theta_2^{**} = [\theta_2^*, \overline{\theta}]$, and $\theta_2^* \in (\underline{\theta}, \overline{\theta}]$ is the cutoff point identified in Lemma 2. From Figure 4, we know that in Case (ii): $\theta_2^* \in (\underline{\theta}, \overline{\theta})$,

$$\Theta_2^*(\theta_1) = \left[\frac{1-\gamma_2}{1-\gamma_1}\theta_1, 1\right]$$

for all $\theta_1 \in [0,1]$, and $\theta_2^* = (1-\gamma_2)/(1-\gamma_1)$; hence,

$$\Theta_2^{**} = \left[\frac{1-\gamma_2}{1-\gamma_1}, 1\right].$$

Reflecting the type space $\Theta = [0,1]^2$ and each agent i's expected valuation $\tilde{u}_i^e(\theta_i,\theta_{-i}) = \theta_i + \gamma_i\theta_{-i}$ in Assumption 2, we obtain

$$\int_0^1 \int_{\frac{1-\gamma_2}{1-\gamma_1}\theta_1}^1 \left((1-\gamma_1)\theta_2 - (1-\gamma_2)\theta_1\right) d\theta_2 d\theta_1 - \int_0^1 \int_{\frac{1-\gamma_2}{1-\gamma_1}}^1 \left(\theta_2 - \frac{1-\gamma_2}{1-\gamma_1}\right) d\theta_2 d\theta_1 \geq 0.$$

We compute the left-hand side of the above inequality:

$$
\begin{aligned}
& \int_0^1 \left[\frac{1}{2}(1-\gamma_1)\left(1-\left(\frac{1-\gamma_2}{1-\gamma_1}\theta_1\right)^2\right)-(1-\gamma_2)\theta_1\left(1-\frac{1-\gamma_2}{1-\gamma_1}\theta_1\right)\right]d\theta_1 \\
& -\int_0^1 \left[\frac{1}{2}\left(1-\left(\frac{1-\gamma_2}{1-\gamma_1}\right)^2\right)-\frac{1-\gamma_2}{1-\gamma_1}\left(1-\frac{1-\gamma_2}{1-\gamma_1}\right)\right]d\theta_1 \\
= & \int_0^1 \left[\frac{1}{2}(1-\gamma_1)-\frac{1}{2}\frac{(1-\gamma_2)^2}{1-\gamma_1}(\theta_1)^2-(1-\gamma_2)\theta_1+\frac{(1-\gamma_2)^2}{1-\gamma_1}(\theta_1)^2\right]d\theta_1 \\
& -\int_0^1 \left[\frac{1}{2}-\frac{1}{2}\left(\frac{1-\gamma_2}{1-\gamma_1}\right)^2-\frac{1-\gamma_2}{1-\gamma_1}+\left(\frac{1-\gamma_2}{1-\gamma_1}\right)^2\right]d\theta_1.
\end{aligned}
$$

We continue our computation below:

$$
\begin{aligned}
& \int_0^1 \left[\frac{1}{2}(1-\gamma_1)-(1-\gamma_2)\theta_1+\frac{1}{2}\frac{(1-\gamma_2)^2}{1-\gamma_1}(\theta_1)^2\right]d\theta_1-\int_0^1 \left[\frac{1}{2}-\frac{1-\gamma_2}{1-\gamma_1}+\frac{1}{2}\left(\frac{1-\gamma_2}{1-\gamma_1}\right)^2\right]d\theta_1 \\
= & \frac{1}{2}(1-\gamma_1)-\frac{1}{2}(1-\gamma_2)+\frac{1}{6}\frac{(1-\gamma_2)^2}{1-\gamma_1}-\frac{1}{2}+\frac{1-\gamma_2}{1-\gamma_1}-\frac{1}{2}\left(\frac{1-\gamma_2}{1-\gamma_1}\right)^2 \\
= & \frac{1}{6}\frac{(1-\gamma_2)^2}{1-\gamma_1}+\frac{1-\gamma_2}{1-\gamma_1}-\frac{1}{2}\left(\frac{1-\gamma_2}{1-\gamma_1}\right)^2+\frac{1}{2}(\gamma_2-\gamma_1-1).
\end{aligned}
$$

Therefore, our Assumption 2 is reduced to

$$
\frac{1}{6}\frac{(1-\gamma_2)^2}{1-\gamma_1}+\frac{1-\gamma_2}{1-\gamma_1}-\frac{1}{2}\left(\frac{1-\gamma_2}{1-\gamma_1}\right)^2+\frac{1}{2}(\gamma_2-\gamma_1-1)\geq 0.
$$

This completes the proof of Lemma 3. □

9.8 Proof of Claim 5

Proof. We define the mechanism $(\Theta, \Pi, \hat{x}, \hat{t}, \psi)$ as follows: for each type report $(\theta_1^r, \theta_2^r) \in \Theta$ and each payoff report $(u_1^r, u_2^r) \in \Pi_1 \times \Pi_2$,

$$
\hat{x}(\theta_1^r, \theta_2^r) = (1-\varepsilon)x^*(\theta_1^r, \theta_2^r) + \varepsilon x^0(\theta_1^r, \theta_2^r)
$$

and

$$
\hat{t}_i(\theta_1^r, \theta_2^r; u_1^r, u_2^r) = (1-\varepsilon)t_i^{GS}(\theta_1^r, \theta_2^r; u_1^r, u_2^r) + \varepsilon t_i^0(\theta_1^r, \theta_2^r; u_1^r, u_2^r)
$$

for each agent $i \in \{1, 2\}$. It suffices to check that the mechanism $(\Theta, \Pi, \hat{x}, \hat{t}, \psi)$ is equivalent to the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$. We first compare the decision rules $\hat{x}$ and x^ε.

1. If $x^*(\theta_1^r, \theta_2^r) = 1$, then $\hat{x}$ dictates that the good is delivered to the buyer with probability $(1 - \varepsilon)$, which is the same as x^ε;

2. If $x^*(\theta_1^r, \theta_2^r) = 0$, then $\hat{x}$ dictates that the seller keeps the good with probability one, which is the same as x^ε.

Therefore, the decision rule $\hat{x}$ is equivalent to x^ε.

Finally, we compare the transfer rules $\hat{t}$ and $t^{\varepsilon,GS}$.

1. If $x^*(\theta_1^r, \theta_2^r) = 1$ and trade occurs after randomization, then $\hat{t}$ dictates

   (a) the buyer pays an amount equal to $\tilde{u}_2^e(\theta_1^r, \theta_2^r)$;

   (b) the seller receives an amount equal to $\tilde{u}_2^e(\theta_1^r, \theta_2^r)$ if $u_2^r \in \widetilde{\Pi}_2(\theta_1^r, \theta_2^r)$;

   (c) the seller pays a penalty equal to ψ if $u_2^r \notin \widetilde{\Pi}_2(\theta_1^r, \theta_2^r)$.

   Hence, the realizations of transfers under the transfer rule $\hat{t}$ is the same as those under $t^{\varepsilon,GS}$.

2. If trade does not occur after randomization, then $\hat{t}$ dictates

   (a) the seller makes no transfer;

   (b) the buyer makes no transfer if $u_1^r \in \widetilde{\Pi}_1(\theta_1^r, \theta_2^r)$;

   (c) the buyer pays a penalty equal to ψ if $u_1^r \notin \widetilde{\Pi}_1(\theta_1^r, \theta_2^r)$.

   Hence, the realizations of transfers under the transfer rule $\hat{t}$ is the same as those under $t^{\varepsilon,GS}$.

Thus, we conclude that the mechanism $(\Theta, \Pi, \hat{x}, \hat{t}, \psi)$ is equivalent to the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$. □

9.9 Proof of Claim 7

Proof. The proof consists of four steps. In Step 1, we show that if $\int_{\Theta_2^{**}} dF(\theta_2) = 0$ and Assumption 1 holds for x^*, then Assumption 3 holds for x^ε. In Step 2, we provide an example where $\int_{\Theta_2^{**}} dF(\theta_2) = 0$, Assumption 3 holds for x^ε, but Assumption 1 is violated for x^*. Combining Steps 1 and 2, we obtain that when $\int_{\Theta_2^{**}} dF(\theta_2) = 0$, Assumption 3 is weaker than Assumption 1, as desired. On the other hand, Step 3 provides an example where Assumption 1 holds for x^*, but Assumption 3 is violated for x^ε when $\int_{\Theta_2^{**}} dF(\theta_2) > 0$. Step 4 provides an example where Assumption 3 holds for x^ε, but Assumption 1 is violated for x^* when

$\int_{\Theta_2^{**}} dF(\theta_2) > 0$. Combining Steps 3 and 4, we obtain that when $\int_{\Theta_2^{**}} dF(\theta_2) > 0$, Assumption 1 neither implies nor is implied by Assumption 3, as desired.

Step 1: If $\int_{\Theta_2^{**}} dF(\theta_2) = 0$ and Assumption 1 holds for x^*, then Assumption 3 holds for x^ε.

Proof. Observe that Condition (i) in Assumption 3 is identical to Condition (i) in Assumption 1. Then, it suffices to check (a) the relation between $U_i^{\varepsilon,\psi_0}(\theta_i;\theta_i^r) - U_i^{\varepsilon,\psi_0}(\theta_i;\theta_i)$ in Assumption 3 and $U_i^{\psi_0}(\theta_i;\theta_i^r) - U_i^{\psi_0}(\theta_i;\theta_i)$ in Assumption 1 and, (b) the relation between Condition (ii) in Assumption 3 and Condition (ii) in Assumption 1. We complete the proof of Step 1 by considering the following two subcases, each of which dictates who deviates in the first stage.

Case 1-A: If the seller deviates in the first stage while the buyer reports truthfully,

Step 1-A-1: We show

$$U_1^{\varepsilon,\psi_0}(\theta_1;\theta_1^r) - U_1^{\varepsilon,\psi_0}(\theta_1;\theta_1) = (1-\varepsilon)\left(U_1^{\psi_0}(\theta_1;\theta_1^r) - U_1^{\psi_0}(\theta_1;\theta_1)\right)$$

for all $\theta_1, \theta_1^r \in \Theta_1$.

Proof. Consider the seller of type θ_1. Recall that in the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$, if both agents report truthfully and trade occurs after randomization, the seller's monetary transfer is $\tilde{u}_2^e(\theta_1,\theta_2)$. Then, the expected utility of the seller under truth-telling is

$$\int_{\Theta_2} \left[x^\varepsilon(\theta_1,\theta_2)\tilde{u}_2^e(\theta_1,\theta_2) + (1 - x^\varepsilon(\theta_1,\theta_2))\, \tilde{u}_1^e(\theta_1,\theta_2)\right] dF(\theta_2),$$

where $\tilde{u}_i^e(\theta_i,\theta_{-i}) = \int_\Omega \tilde{u}_i(\theta_i,\theta_{-i};\omega) dH\left(\omega|(\theta_i,\theta_{-i})\right)$ for each agent i. Note that the expected utility of the seller under truth-telling is identical to that in the approximate GS mechanism with $\psi = 0$, which is $U_1^{\varepsilon,\psi_0}(\theta_1;\theta_1)$. The seller's expected utility under truth-telling can be rewritten as follows:

$$\begin{aligned} U_1^{\varepsilon,\psi_0}(\theta_1;\theta_1) &= \int_{\Theta_2} \left[(1-\varepsilon)x^*(\theta_1,\theta_2)\tilde{u}_2^e(\theta_1,\theta_2) + (1 - x^*(\theta_1,\theta_2) + \varepsilon x^*(\theta_1,\theta_2))\, \tilde{u}_1^e(\theta_1,\theta_2)\right] dF(\theta_2) \\ &= \int_{\Theta_2} \left[x^*(\theta_1,\theta_2)\tilde{u}_2^e(\theta_1,\theta_2) + (1 - x^*(\theta_1,\theta_2))\, \tilde{u}_1^e(\theta_1,\theta_2)\right] dF(\theta_2) \\ &\quad -\varepsilon \int_{\Theta_2} \left[x^*(\theta_1,\theta_2)\tilde{u}_2^e(\theta_1,\theta_2) - x^*(\theta_1,\theta_2)\tilde{u}_1^e(\theta_1,\theta_2)\right] dF(\theta_2). \end{aligned}$$

Recall that the seller's expected utility under truth-telling in the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ with $\psi = 0$ is

$$U_1^{\psi_0}(\theta_1;\theta_1) = \int_{\Theta_2} \left[x^*(\theta_1,\theta_2)\tilde{u}_2^e(\theta_1,\theta_2) + (1 - x^*(\theta_1,\theta_2))\, \tilde{u}_1^e(\theta_1,\theta_2)\right] dF(\theta_2).$$

Hence, the seller's expected utility under truth-telling in the approximate GS mechanism $(\Theta, \Pi, x^{\varepsilon}, t^{\varepsilon,GS}, \psi)$ with $\psi = 0$ can be further rewritten as follows:

$$\begin{aligned} U_1^{\varepsilon,\psi_0}(\theta_1;\theta_1) &= U_1^{\psi_0}(\theta_1;\theta_1) - \varepsilon\left[U_1^{\psi_0}(\theta_1;\theta_1) - \int_{\Theta_2} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2)\right] \\ &= (1-\varepsilon)U_1^{\psi_0}(\theta_1;\theta_1) + \varepsilon\int_{\Theta_2} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2). \end{aligned} \tag{15}$$

On the other hand, if the seller deviates to θ_1^r in the approximate GS mechanism $(\Theta, \Pi, x^{\varepsilon}, t^{\varepsilon,GS}, \psi)$ with $\psi = 0$, her expected utility becomes

$$\begin{aligned} & U_1^{\varepsilon,\psi_0}(\theta_1;\theta_1^r) \\ = & \int_{\Theta_2}\left[x^{\varepsilon}(\theta_1^r,\theta_2)\tilde{u}_2^e(\theta_1^r,\theta_2)\left(1-\int_{\Omega^{\psi}(\theta_1,\theta_2;\theta_1^r)} dH\left(\omega|(\theta_1,\theta_2)\right)\right) + \left(1-x^{\varepsilon}(\theta_1^r,\theta_2)\right)\tilde{u}_1^e(\theta_1,\theta_2)\right]dF(\theta_2), \end{aligned}$$

which can be rewritten as

$$\begin{aligned} & U_1^{\varepsilon,\psi_0}(\theta_1;\theta_1^r) \\ = & \int_{\Theta_2}(1-\varepsilon)x^*(\theta_1^r,\theta_2)\tilde{u}_2^e(\theta_1^r,\theta_2)\left(1-\int_{\Omega^{\psi}(\theta_1,\theta_2;\theta_1^r)} dH\left(\omega|(\theta_1,\theta_2)\right)\right)dF(\theta_2) \\ & + \int_{\Theta_2}\left(1-x^*(\theta_1^r,\theta_2)+\varepsilon x^*(\theta_1^r,\theta_2)\right)\tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2) \\ = & \int_{\Theta_2}\left[x^*(\theta_1^r,\theta_2)\tilde{u}_2^e(\theta_1^r,\theta_2)\left(1-\int_{\Omega^{\psi}(\theta_1,\theta_2;\theta_1^r)} dH\left(\omega|(\theta_1,\theta_2)\right)\right) + \left(1-x^*(\theta_1^r,\theta_2)\right)\tilde{u}_1^e(\theta_1,\theta_2)\right]dF(\theta_2) \\ & -\varepsilon\int_{\Theta_2}\left[x^*(\theta_1^r,\theta_2)\tilde{u}_2^e(\theta_1^r,\theta_2)\left(1-\int_{\Omega^{\psi}(\theta_1,\theta_2;\theta_1^r)} dH\left(\omega|(\theta_1,\theta_2)\right)\right) - x^*(\theta_1^r,\theta_2)\tilde{u}_1^e(\theta_1,\theta_2)\right]dF(\theta_2). \end{aligned}$$

Recall that the seller's expected utility after deviation in the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ with $\psi = 0$ is

$$\begin{aligned} & U_1^{\psi_0}(\theta_1;\theta_1^r) \\ = & \int_{\Theta_2}\left[x^*(\theta_1^r,\theta_2)\tilde{u}_2^e(\theta_1^r,\theta_2)\left(1-\int_{\Omega^{\psi}(\theta_1,\theta_2;\theta_1^r)} dH\left(\omega|(\theta_1,\theta_2)\right)\right) + \left(1-x^*(\theta_1^r,\theta_2)\right)\tilde{u}_1^e(\theta_1,\theta_2)\right]dF(\theta_2). \end{aligned}$$

Hence, the seller's expected utility after deviation in the approximate GS mechanism $(\Theta, \Pi, x^{\varepsilon}, t^{\varepsilon,GS}, \psi)$ with $\psi = 0$ can be further rewritten as follows:

$$\begin{aligned} U_1^{\varepsilon,\psi_0}(\theta_1;\theta_1^r) &= U_1^{\psi_0}(\theta_1;\theta_1^r) - \varepsilon\left[U_1^{\psi_0}(\theta_1;\theta_1^r) - \int_{\Theta_2} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2)\right] \\ &= (1-\varepsilon)U_1^{\psi_0}(\theta_1;\theta_1^r) + \varepsilon\int_{\Theta_2} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2). \end{aligned} \tag{16}$$

Then, using (15) and (16), we compute

$$U_1^{\varepsilon,\psi_0}(\theta_1;\theta_1^r) - U_1^{\varepsilon,\psi_0}(\theta_1;\theta_1) = (1-\varepsilon)\left[U_1^{\psi_0}(\theta_1;\theta_1^r) - U_1^{\psi_0}(\theta_1;\theta_1)\right].$$

As a result, $U_1^{\varepsilon,\psi_0}(\theta_1;\theta_1^r) > U_1^{\varepsilon,\psi_0}(\theta_1;\theta_1)$ if and only if $U_1^{\psi_0}(\theta_1;\theta_1^r) > U_1^{\psi_0}(\theta_1;\theta_1)$, implying that the seller has an incentive to deviate in the approximate GS mechanism $(\Theta,\Pi,x^\varepsilon,t^{\varepsilon,GS},\psi)$ if and only if she has an incentive to deviate in the GS mechanism $(\Theta,\Pi,x^*,t^{GS},\psi)$. This completes the proof of Step 1-A-1. □

In words, the seller has an incentive to deviate in the approximate GS mechanism $(\Theta,\Pi,x^\varepsilon,t^{\varepsilon,GS},\psi_0)$ if and only if she has an incentive to deviate in the GS mechanism $(\Theta,\Pi,x^*,t^{GS},\psi_0)$.

Step 1-A-2: We show that if the seller deviates in the first stage while the buyer reports truthfully, then Condition (ii) of Assumption 3 is equivalent to Condition (ii) of Assumption 1.

Proof. If the seller deviates to θ_1^r in the approximate GS mechanism $(\Theta,\Pi,x^\varepsilon,t^{\varepsilon,GS},\psi_0)$, then for any positive-measure set $\Theta_2^+ \subseteq \Theta_2$, we have

$$\begin{aligned}\int_{\Theta_2^+} x_2^\varepsilon(\theta_1^r,\theta_2)dF(\theta_2) &= \int_{\Theta_2^+} x^\varepsilon(\theta_1^r,\theta_2)dF(\theta_2)\\ &= \int_{\Theta_2^+}(1-\varepsilon)x^*(\theta_1^r,\theta_2)dF(\theta_2) = \int_{\Theta_2^+}(1-\varepsilon)x_2^*(\theta_1^r,\theta_2)dF(\theta_2),\end{aligned}$$

implying that $\int_{\Theta_2^+} x_2^\varepsilon(\theta_1^r,\theta_2)dF(\theta_2) > 0$ if and only if $\int_{\Theta_2^+} x_2^*(\theta_1^r,\theta_2)dF(\theta_2) > 0$. Therefore, Condition (ii) of Assumption 3 and Condition (ii) of Assumption 1 are equivalent. This completes the proof of Step 1-A-2. □

Combining Steps 1-A-1 and 1-A-2, we obtain that if the seller deviates in the first stage, then Assumption 3 under x^ε is equivalent to Assumption 1 under x^*.

Case 1-B: If the buyer deviates in the first stage while the seller reports truthfully,

Step 1-B-1: We show

$$U_2^{\varepsilon,\psi_0}(\theta_2;\theta_2^r) - U_2^{\varepsilon,\psi_0}(\theta_2;\theta_2) = (1-\varepsilon)\left(U_2^{\psi_0}(\theta_2;\theta_2^r) - U_2^{\psi_0}(\theta_2;\theta_2)\right)$$

for all $\theta_2,\theta_2^r \in \Theta_2$.

Proof. Consider the buyer of type θ_2. Recall that if both agents report truthfully and trade occurs after randomization, the buyer's monetary transfer is $-\tilde{u}_2^e(\theta_1,\theta_2)$. Then, the expected utility of the buyer under truth-telling is

$$\int_{\Theta_1} x^\varepsilon(\theta_1,\theta_2)\left(\tilde{u}_2^e(\theta_1,\theta_2) - \tilde{u}_2^e(\theta_1,\theta_2)\right)dF(\theta_1) = 0,$$

where $\tilde{u}_2^e(\theta_1, \theta_2) = \int_\Omega \tilde{u}_2(\theta_1, \theta_2; \omega) dH(\omega|(\theta_1, \theta_2))$. Note that the expected utility of the buyer under truth-telling is identical to that in the approximate GS mechanism with $\psi = 0$, which is $U_2^{\varepsilon,\psi_0}(\theta_2; \theta_2)$. Recall that the buyer's expected utility under truth-telling in the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ with $\psi = 0$, $U_2^{\psi_0}(\theta_2; \theta_2)$, is also zero. On the other hand, if the buyer deviates to θ_2^r in the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$ with $\psi = 0$, his expected utility becomes

$$\begin{aligned} U_2^{\varepsilon,\psi_0}(\theta_2; \theta_2^r) &= \int_{\Theta_1} x^\varepsilon(\theta_1, \theta_2^r) \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^r)\right) dF(\theta_1) \\ &= \int_{\Theta_1} (1-\varepsilon) x^*(\theta_1, \theta_2^r) \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^r)\right) dF(\theta_1). \end{aligned}$$

Recall that the buyer's expected utility after deviation in the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ with $\psi = 0$ is

$$U_2^{\psi_0}(\theta_2; \theta_2^r) = \int_{\Theta_1} x^*(\theta_1, \theta_2^r) \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^r)\right) dF(\theta_1).$$

Hence, the buyer's expected utility after deviation in the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$ with $\psi = 0$ can be rewritten as follows:

$$U_2^{\varepsilon,\psi_0}(\theta_2; \theta_2^r) = (1-\varepsilon) U_2^{\psi_0}(\theta_2; \theta_2^r),$$

Finally, we compute

$$U_2^{\varepsilon,\psi_0}(\theta_2; \theta_2^r) - U_2^{\varepsilon,\psi_0}(\theta_2; \theta_2) = (1-\varepsilon) U_2^{\psi_0}(\theta_2; \theta_2^r) = (1-\varepsilon)\left(U_2^{\psi_0}(\theta_2; \theta_2^r) - U_2^{\psi_0}(\theta_2; \theta_2)\right).$$

The first equality follows because $U_2^{\varepsilon,\psi_0}(\theta_2; \theta_2) = 0$. The second equality holds because $U_2^{\psi_0}(\theta_2; \theta_2) = 0$. As a result, $U_2^{\varepsilon,\psi_0}(\theta_2; \theta_2^r) > U_2^{\varepsilon,\psi_0}(\theta_2; \theta_2)$ if and only if $U_2^{\psi_0}(\theta_2; \theta_2^r) > U_2^{\psi_0}(\theta_2; \theta_2)$, implying that the buyer has an incentive to deviate in the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$ with $\psi = 0$ if and only if he has an incentive to deviate in the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ with $\psi = 0$. This completes the proof of Step 1-B-1. □

In words, the buyer has an incentive to deviate in the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi_0)$ if and only if he has an incentive to deviate in the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi_0)$.

Step 1-B-2: We show that if the buyer deviates in the first stage while the seller reports truthfully, Condition (ii) of Assumption 3 is satisfied for any positive-measure set $\Theta_1^+ \subseteq \Theta_1$.

Proof. If the buyer deviates to θ_2^r in the approximate GS mechanism $(\Theta, \Pi, x^{\varepsilon}, t^{\varepsilon,GS}, \psi)$ with $\psi = 0$, then for any positive-measure set $\Theta_1^+ \subseteq \Theta_1$, we have

$$\int_{\Theta_1^+} x_1^{\varepsilon}(\theta_1, \theta_2^r) dF(\theta_1) = \int_{\Theta_1^+} (1 - x^{\varepsilon}(\theta_1, \theta_2^r))\, dF(\theta_1) = \int_{\Theta_1^+} [1 - (1-\varepsilon)x^*(\theta_1, \theta_2^r)]\, dF(\theta_1) > 0.$$

The inequality follows because for all $\theta_1 \in \Theta_1$, $1 - (1-\varepsilon)x^*(\theta_1, \theta_2^r) \geq 1 - (1-\varepsilon) = \varepsilon > 0$. Hence, Condition (ii) of Assumption 3 is always satisfied if the buyer deviates. This completes the proof of Step 1-B-2. $\square$

Recall $\int_{\Theta_2^{**}} dF(\theta_2) = 0$. If $\Theta_2^{**} = \emptyset$ and the buyer deviates to some θ_2^r, there always exists a positive-measure set $\Theta_1^+ \subseteq \Theta_1$ such that $\int_{\Theta_1^+} x_1^*(\theta_1, \theta_2^r) dF(\theta_1) > 0$. In other words, there always exists a positive-measure set $\Theta_1^+ \subseteq \Theta_1$ such that Condition (ii) of Assumption 1 is satisfied. Moreover, if Θ_2^{**} is a singleton, say, $\Theta_2^{**} = \{\hat{\theta}_2\}$, and the buyer has an incentive to deviate to $\hat{\theta}_2$ in the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$, then $\int_{\Theta_1^+} x_1^*(\theta_1, \hat{\theta}_2) dF(\theta_1) = 0$ for all $\Theta_1^+ \subseteq \Theta_1$, implying that Condition (ii) of Assumption 1 is violated. We thus conclude that Condition (ii) of Assumption 1 is stronger than Condition (ii) of Assumption 3. This completes the proof of Step 1. $\square$

Step 2: We provide an example in which $\int_{\Theta_2^{**}} dF(\theta_2) = 0$; Assumption 3 holds for x^{ε}; but Assumption 1 fails for x^*.

Proof. The proposed example is based on Example 1, except that we specify the valuation functions differently.

Example 4. Suppose that each agent i's type θ_i is drawn from the uniform distribution on the unit interval $[0, 1]$. Conditional on the type profile θ, ω_1 and ω_2 are independently drawn from the uniform distribution on the closed interval $[-0.1, 0.1]$. Then, we define $\Omega = \{(\omega_1, \omega_2) |\ \omega_1, \omega_2 \in [-0.1, 0.1]\}$ as the set of residual uncertainties. Each agent's valuation for the object when the true type profile is (θ_1, θ_2) and the residual uncertainty is $\omega = (\omega_1, \omega_2) \in \Omega$ is given as follows:

$$\tilde{u}_1(\theta_1, \theta_2; \omega) \;=\; \begin{cases} 1.5\theta_1 & \text{if } \theta_1 \geq 0.3; \\ \theta_1 + 0.2\theta_2 + \omega_1 & \text{otherwise} \end{cases}$$

and

$$\tilde{u}_2(\theta_1, \theta_2; \omega) \;=\; \theta_2 + 0.2\theta_1 + \omega_2.$$

We compute

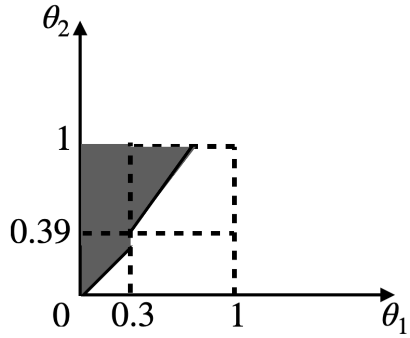
$$\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_1^e(\theta_1, \theta_2) \;=\; \begin{cases} \theta_2 - 1.3\theta_1 & \text{if } \theta_1 \geq 0.3; \\ 0.8(\theta_2 - \theta_1) & \text{otherwise,} \end{cases}$$

implying that the efficient decision rule is

$$x^*(\theta_1, \theta_2) = \begin{cases} 1 & \text{if either } \theta_2 \geq 1.3\theta_1 \text{ for every } \theta_1 \geq 0.3 \text{ or } \theta_2 \geq \theta_1 \text{ for every } \theta_1 < 0.3; \\ 0 & \text{otherwise} \end{cases}$$

The following figure illustrates the decision at different type profiles in this example; in particular, the shaded region represents $\Theta_* = \{(\theta_1, \theta_2) \in [0,1]^2 : x^*(\theta_1, \theta_2) = 1\}$.

Figure 6



The example exhibits the following features regarding the buyer's incentive constraints. Under the efficient decision rule x^*, there exist some type profiles where it is efficient not to trade (i.e., the seller keeps the good with probability one) but the seller's allocation payoff does not reveal to the seller herself any information about the buyer's type. In that case, the buyer's deviation cannot be detected in the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$, implying that Assumption 1 is violated. In all the type profiles, the seller keeps the good with probability $\varepsilon > 0$ under the decision rule x^ε. Since the seller's valuation function is strictly increasing in the buyer's type, we can find a set of states with positive probability in which the buyer's deviation can be detected in the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$, implying that Assumption 3 can be satisfied. The rest of the proof formally establish these points.

The proof of Step 2 is completed by two steps.

Step 2-1: Assumption 1 is violated for x^* in Example 4.

Proof. Consider the buyer of type $\theta_2 > 0.39$. If both agents report truthfully in both stages, he receives zero expected utility in the GS mechanism $(\Theta, \Pi, x^*, t^{GS})$ with $\psi = 0$. If he deviates to $\theta_2^r \in (0.39, \theta_2)$, then his expected utility in the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ with $\psi = 0$ becomes

$$\int_{\Theta_1} x^*(\theta_1, \theta_2^r) \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^r)\right) dF(\theta_2) > 0,$$

because the buyer's expected utility is strictly increasing in his own type. As a result, the buyer of type $\theta_2 > 0.39$ has an incentive to deviate to $\theta_2^r \in (0.39, \theta_2)$ in the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$ with $\psi = 0$. Moreover, after the deviation, the good will be allocated to the seller only when $\theta_1 > 0.3$; in that region, the seller's allocation payoff is always equal to $1.5\theta_1$, which is independent of the buyer's type. In other words, all feasible allocation payoffs of the seller conditional on the true type profile (θ_1, θ_2) are also feasible given the deviation (θ_1, θ_2^r), implying that Condition (i) of Assumption 1 is violated. Therefore, Assumption 1 is violated in this example. This completes the proof of Step 2-1. □

Step 2-2: Assumption 3 holds for x^ε in Example 4.

Proof. The proof of Step 2-2 is further divided into two steps.

Step 2-2-1: Assumption 3 is satisfied for the seller.

Proof. Consider the seller of type θ_1. If both agents report truthfully in both stages, the seller's expected utility after particiating in the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$ is

$$\int_{\Theta_2} \left[x^\varepsilon(\theta_1,\theta_2)\tilde{u}_2^e(\theta_1,\theta_2) + (1 - x^\varepsilon(\theta_1,\theta_2))\,\tilde{u}_1^e(\theta_1,\theta_2)\right] dF(\theta_2) \geq \int_{\Theta_2} \tilde{u}_1^e(\theta_1,\theta_2) dF(\theta_2),$$

where the inequality follows from the proof of Claim 6. Note that the seller's expected utility under truth-telling is identical to that in the approximate GS mechanism with $\psi = 0$, which is $U_1^{\varepsilon,\psi_0}(\theta_1;\theta_1)$. If the seller keeps the good with probability one after she deviates to θ_1^r, her expected utility in the approximate GS mechanism with $\psi = 0$ equals $\int_{\Theta_2} \tilde{u}_1^e(\theta_1,\theta_2)dF(\theta_2)$ and this is not a profitable deviation. Therefore, the seller has an incentive to deviate in the approximate GS mechanism with $\psi = 0$ only if the good is allocated to the buyer with positive probability after the deviation. In other words, if $U_1^{\varepsilon,\psi_0}(\theta_1;\theta_1^r) > U_1^{\varepsilon,\psi_0}(\theta_1;\theta_1)$, then there must exist a positive-measure set $\Theta_2^+ \subseteq \Theta_2$ such that $\int_{\Theta_2^+} x^\varepsilon(\theta_1^r,\theta_2)dF(\theta_2) > 0$, implying that Condition (ii) of Assumption 3 is satisfied.

It remains to verify that Condition (i) of Assumption 3 is satisfied, that is, for every $\theta_2 \in \Theta_2^+$, $\int_{\Omega^\psi(\theta_1,\theta_2;\theta_1^r)} dH\left(\omega|(\theta_1,\theta_2)\right) > 0$. The set of feasible allocation payoffs of the buyer given the true type profile (θ_1,θ_2) is

$$\begin{aligned}
\widetilde{\Pi}_2(\theta_1,\theta_2) &= \{\hat{u}_2 \in \mathbb{R} : \text{ there exists } \omega \in \Omega \text{ s.t. } h\left(\omega|(\theta_1,\theta_2)\right) > 0 \text{ and } \tilde{u}_2(\theta_1,\theta_2;\omega) = \hat{u}_2\} \\
&= \{\hat{u}_2 \in \mathbb{R} : \hat{u}_2 = \theta_2 + 0.2\theta_1 + \omega_2 \text{ for some } \omega_2 \in [-0.1, 0.1]\} \\
&= [\theta_2 + 0.2\theta_1 - 0.1, \theta_2 + 0.2\theta_1 + 0.1].
\end{aligned}$$

Similarly, the set of feasible allocation payoffs of the buyer given the deviation (θ_1^r, θ_2) becomes

$$\begin{aligned}\widetilde{\Pi}_2(\theta_1^r, \theta_2) &= \{\hat{u}_2 \in \mathbb{R}: \text{ there exists } \omega \in \Omega \text{ s.t. } h(\omega|(\theta_1^r, \theta_2)) > 0 \text{ and } \tilde{u}_2(\theta_1^r, \theta_2; \omega) = \hat{u}_2\} \\ &= \{\hat{u}_2 \in \mathbb{R} : \hat{u}_2 = \theta_2 + 0.2\theta_1^r + \omega_2 \text{ where } \omega_2 \in [-0.1, 0.1]\} \\ &= [\theta_2 + 0.2\theta_1^r - 0.1, \theta_2 + 0.2\theta_1^r + 0.1].\end{aligned}$$

Suppose $\theta_1^r > \theta_1$ without loss of generality. Then, there exist some allocation payoffs of the buyer such that they are feasible given the true type profile (θ_1, θ_2), but they are not feasible given the deviation (θ_1^r, θ_2), which is,

$$\begin{cases}[\theta_2 + 0.2\theta_1 - 0.1, \theta_2 + 0.2\theta_1 + 0.1] & \text{if } \theta_2 + 0.2\theta_1^r - 0.1 > \theta_2 + 0.2\theta_1 + 0.1, \\ [\theta_2 + 0.2\theta_1 - 0.1, \theta_2 + 0.2\theta_1^r - 0.1) & \text{otherwise.}\end{cases}$$

As a result, the set of residual uncertainties under which the buyer's allocation payoffs are feasible given the true type profile (θ_1, θ_2), but not feasible given the deviation (θ_1^r, θ_2), is as follows:

$$\Omega^{\psi}(\theta_1, \theta_2; \theta_1^r) = \begin{cases}[-0.1, 0.1]^2 & \text{if } \theta_2 + 0.2\theta_1^r - 0.1 > \theta_2 + 0.2\theta_1 + 0.1, \\ \{(\omega_1, \omega_2) : \omega_1 \in [-0.1, 0.1] \text{ and } \omega_2 \in [-0.1, 0.2(\theta_1^r - \theta_1) - 0.1)\} & \text{otherwise.}\end{cases}$$

Therefore, for every $\theta_1, \theta_1^r \in \Theta_1$ and every $\theta_2 \in \Theta_2$, there exists a positive-measure set of residual uncertainties under which the buyer's allocation payoffs are feasible given the true type profile (θ_1, θ_2), but not feasible given the deviation (θ_1^r, θ_2), implying that Condition (i) of Assumption 3 is satisfied for all $\theta_1, \theta_1^r \in \Theta_1$ and all $\theta_2 \in \Theta_2$. Recall that Condition (ii) of Assumption 3 is also satisfied. We conclude that Assumption 3 is satisfied for the seller. $\square$

Step 2-2-2: Assumption 3 is satisfied for the buyer.

Proof. Consider the buyer of type θ_2. Recall $x^{\varepsilon}(\theta_1, \theta_2) = (1 - \varepsilon)x^*(\theta_1, \theta_2)$ for all $(\theta_1, \theta_2) \in \Theta$. Then, for all $(\theta_1, \theta_2) \in \Theta$, the probability that the seller keeps the good is at least as high as ε. Moreover, when $\theta_1 < 0.3$, the set of feasible allocation payoffs of the seller given the true type profile (θ_1, θ_2) is

$$\begin{aligned}\widetilde{\Pi}_1(\theta_1, \theta_2) &= \{\hat{u}_1 \in \mathbb{R}: \text{ there exists } \omega \in \Omega \text{ s.t. } h(\omega|(\theta_1, \theta_2)) > 0 \text{ and } \tilde{u}_1(\theta_1, \theta_2; \omega) = \hat{u}_1\} \\ &= \{\hat{u}_1 \in \mathbb{R} : \hat{u}_1 = \theta_1 + 0.2\theta_2 + \omega_1 \text{ for some } \omega_1 \in [-0.1, 0.1]\} \\ &= [\theta_1 + 0.2\theta_2 - 0.1, \theta_1 + 0.2\theta_2 + 0.1].\end{aligned}$$

Similarly, the set of feasible allocation payoffs of the buyer given the deviation (θ_1, θ_2^r) becomes

$$\begin{aligned}\widetilde{\Pi}_1(\theta_1, \theta_2^r) &= \{\hat{u}_1 \in \mathbb{R} : \text{ there exists } \omega \in \Omega \text{ s.t. } h\left(\omega|(\theta_1, \theta_2^r)\right) > 0 \text{ and } \tilde{u}_1(\theta_1, \theta_2^r; \omega) = \hat{u}_1\} \\ &= \{\hat{u}_1 \in \mathbb{R} : \hat{u}_1 = \theta_1 + 0.2\theta_2^r + \omega_1 \text{ where } \omega_1 \in [-0.1, 0.1]\} \\ &= [\theta_1 + 0.2\theta_2^r - 0.1, \theta_1 + 0.2\theta_2^r + 0.1].\end{aligned}$$

Suppose $\theta_2^r < \theta_2$ without loss of generality. Then, there exist some allocation payoffs of the seller such that they are feasible given the true type profile (θ_1, θ_2), but they are not feasible given the deviation (θ_1, θ_2^r), which is,

$$\begin{cases} [\theta_1 + 0.2\theta_2 - 0.1, \theta_1 + 0.2\theta_2 + 0.1] & \text{if } \theta_1 + 0.2\theta_2^r + 0.1 < \theta_1 + 0.2\theta_2 - 0.1 \\ (\theta_1 + 0.2\theta_2^r + 0.1, \theta_1 + 0.2\theta_2 + 0.1] & \text{otherwise.} \end{cases}$$

As a result, the set of residual uncertainties under which the seller's allocation payoffs are feasible given the true type profile (θ_1, θ_2), but not feasible given the deviation (θ_1, θ_2^r), is as follows:

$$\Omega^{\psi}(\theta_2, \theta_1; \theta_2^r) = \begin{cases} [-0.1, 0.1]^2 & \text{if } \theta_1 + 0.2\theta_2^r + 0.1 < \theta_1 + 0.2\theta_2 - 0.1, \\ \{(\omega_1, \omega_2) : \omega_1 \in (0.2(\theta_2^r - \theta_2) + 0.1, 0.1] \text{ and } \omega_2 \in [-0.1, 0.1]\} & \text{otherwise.} \end{cases}$$

Therefore, for every $\theta_2, \theta_2^r \in \Theta_2$, if $\theta_1 < 0.3$, there exists a positive-measure set of residual uncertainties under which the seller's allocation payoffs are feasible given the true type profile (θ_1, θ_2), but not feasible given the deviation (θ_1, θ_2^r). Since the seller keeps the good with positive probability for every $(\theta_1, \theta_2) \in \Theta$, we conclude that Assumption 3 is satisfied for the buyer. □

We conclude that Assumption 3 is satisfied in Example 4. This completes the proof of Step 2-2. □

This completes the proof of Step 2. □

Step 3: We provide an example in which Assumption 1 holds for x^*, whereas Assumption 3 is violated for x^ε when $\int_{\Theta_2^{**}} dF(\theta_2) > 0$.

Proof. The proposed example is built on Example 1, except that we specify the valuation functions differently.

Example 5. Suppose that each agent i's type θ_i is drawn from the uniform distribution on the unit interval $[0, 1]$. Conditional on the type profile θ, ω_1 and ω_2 are independently drawn from the uniform distribution on the closed interval

$[-0.1, 0.1]$. Then, we define $\Omega = \{(\omega_1, \omega_2)|\ \omega_1, \omega_2 \in [-0.1, 0.1]\}$ as the set of residual uncertainties. Each agent's valuation for the object when the true type profile is (θ_1, θ_2) and the residual uncertainty is $\omega = (\omega_1, \omega_2) \in \Omega$ is given as follows:

$$\tilde{u}_1(\theta_1, \theta_2; \omega) = \begin{cases} \theta_1 + \omega_1 & \text{if } \theta_2 \geq 0.4; \\ \theta_1 + 0.5\theta_2 + \omega_1 & \text{otherwise} \end{cases}$$

and

$$\tilde{u}_2(\theta_1, \theta_2; \omega) = \theta_2 + 0.8\theta_1 + \omega_2.$$

We compute

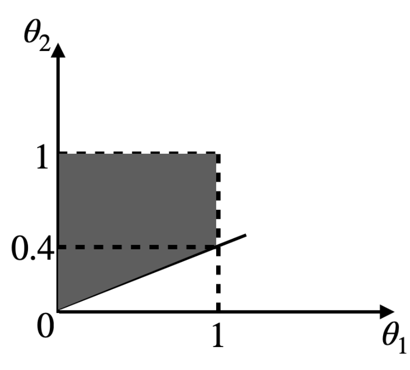
$$\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_1^e(\theta_1, \theta_2) = \begin{cases} \theta_2 - 0.2\theta_1 & \text{if } \theta_2 \geq 0.4; \\ 0.5\theta_2 - 0.2\theta_1 & \text{otherwise,} \end{cases}$$

implying that the efficient decision rule is

$$x^*(\theta_1, \theta_2) = \begin{cases} 1 & \text{if } \theta_2 \geq 0.4\theta_1; \\ 0 & \text{otherwise} \end{cases}$$

The following figure illustrates the decision at different type profiles in this example; in particular, the shaded region represents $\Theta_* = \{(\theta_1, \theta_2) \in [0, 1]^2 : x^*(\theta_1, \theta_2) = 1\}$.

Figure 7



We will show that in this example, Assumption 1 holds for x^*, while Assumption 3 is violated for x^ε. The example exhibits the following features. Under the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$, the buyer has no incentive to deviate because the payments are always the same. However, if we consider the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$, the buyer can lower his payment by deviating to a lower type so that he has an incentive to deviate. We present an example in which such a profitable deviation cannot be detected through the seller's allocation payoff. Hence, Assumption 3 is violated for x^{ε}.

Step 3-1: Assumption 1 holds for x^* in Example 5.

Proof. Note that the efficient decision rule in Example 5 corresponds to Case (ii): $0 < \gamma_1 < \gamma_2 < 1$ in Example 1. It follows from the proof of Claim 2 that if $U_i^{\psi_0}(\theta_i;\theta_i^r) > U_i^{\psi_0}(\theta_i;\theta_i)$, then there exists a positive-measure set $\Theta_j^+ \subseteq \Theta_j$ with $j \neq i$ such that $\int_{\Theta_j^+} x_j^*(\theta_i^r,\theta_j)dF(\theta_j) > 0$ where $x_j^*(\cdot)$ denotes the probability that agent j is allocated the good. In other words, Condition (ii) of Assumption 1 is satisfied.

It remains to verify that Condition (i) of Assumption 1 is satisfied as well, that is, $\int_{\Omega^{\psi}(\theta_i,\theta_j;\theta_i^r)} dH\left(\omega|(\theta_i,\theta_j)\right) > 0$ for every $\theta_j \in \Theta_j^+$. In doing so, we further divide the proof into two steps.

Step 3-1-1: Condition (i) of Assumption 1 is satisfied for the seller.

Proof. Consider the seller of type θ_1. The set of feasible allocation payoffs of the buyer given the true type profile (θ_1,θ_2) is

$$\begin{aligned}
\widetilde{\Pi}_2(\theta_1,\theta_2) &= \{\hat{u}_2 \in \mathbb{R}: \text{ there exists } \omega \in \Omega \text{ s.t. } h\left(\omega|(\theta_1,\theta_2)\right) > 0 \text{ and } \tilde{u}_2(\theta_1,\theta_2;\omega) = \hat{u}_2\} \\
&= \{\hat{u}_2 \in \mathbb{R}: \hat{u}_2 = \theta_2 + 0.8\theta_1 + \omega_2 \text{ for some } \omega_2 \in [-0.1,0.1]\} \\
&= [\theta_2 + 0.8\theta_1 - 0.1, \theta_2 + 0.8\theta_1 + 0.1].
\end{aligned}$$

Similarly, the set of feasible allocation payoffs of the buyer given the deviation (θ_1^r,θ_2) becomes

$$\begin{aligned}
\widetilde{\Pi}_2(\theta_1^r,\theta_2) &= \{\hat{u}_2 \in \mathbb{R}: \text{ there exists } \omega \in \Omega \text{ s.t. } h\left(\omega|(\theta_1^r,\theta_2)\right) > 0 \text{ and } \tilde{u}_2(\theta_1^r,\theta_2;\omega) = \hat{u}_2\} \\
&= \{\hat{u}_2 \in \mathbb{R}: \hat{u}_2 = \theta_2 + 0.8\theta_1^r + \omega_2 \text{ where } \omega_2 \in [-0.1,0.1]\} \\
&= [\theta_2 + 0.8\theta_1^r - 0.1, \theta_2 + 0.8\theta_1^r + 0.1].
\end{aligned}$$

Suppose $\theta_1^r > \theta_1$ without loss of generality. Then, there exist some allocation payoffs of the buyer such that they are feasible given the true type profile (θ_1,θ_2), but they are not feasible given the deviation (θ_1^r,θ_2), which is,

$$\begin{cases}
[\theta_2 + 0.8\theta_1 - 0.1, \theta_2 + 0.8\theta_1 + 0.1] & \text{if } \theta_2 + 0.8\theta_1^r - 0.1 > \theta_2 + 0.8\theta_1 + 0.1 \\
[\theta_2 + 0.8\theta_1 - 0.1, \theta_2 + 0.8\theta_1^r - 0.1) & \text{otherwise.}
\end{cases}$$

As a result, the set of residual uncertainties under which the buyer's allocation payoffs are feasible given the true type profile (θ_1,θ_2), but not feasible given the deviation (θ_1^r, θ_2), is as follows:

$$\Omega^{\psi}(\theta_1, \theta_2; \theta_1^r) = \begin{cases} [-0.1, 0.1]^2 & \text{if } \theta_2 + 0.8\theta_1^r - 0.1 > \theta_2 + 0.8\theta_1 + 0.1, \\ \{(\omega_1, \omega_2) : \omega_1 \in [-0.1, 0.1] \text{ and } \omega_2 \in [-0.1, 0.8(\theta_1^r - \theta_1) - 0.1)\} & \text{otherwise.} \end{cases}$$

Therefore, for every $\theta_1, \theta_1^r \in \Theta_1$ and every $\theta_2 \in \Theta_2$, there exists a positive-measure set of residual uncertainties under which the buyer's allocation payoffs are feasible given the true type profile (θ_1, θ_2), but not feasible given the deviation (θ_1^r, θ_2), implying that Condition (i) of Assumption 1 is satisfied for all $\theta_1, \theta_1^r \in \Theta_1$ and all $\theta_2 \in \Theta_2$. We thus conclude that Assumption 1 is satisfied for the seller. $\square$

Step 3-1-2: Condition (i) of Assumption 1 is satisfied for the buyer.

Proof. Consider the buyer of type θ_2. Recall that Condition (ii) of Assumption 1 is satisfied, that is, if $U_2^{\psi_0}(\theta_2; \theta_2^r) > U_2^{\psi_0}(\theta_2; \theta_2)$, then there exists a positive-measure set $\Theta_1^+ \subseteq \Theta_1$ such that $\int_{\Theta_1^+} x_1^*(\theta_1, \theta_2^r) dF(\theta_1) > 0$ where $x_1^*(\cdot)$ denotes the probability that the seller is allocated the good. Since the seller never keeps the good when $\theta_2^r \geq 0.4$ in Example 5, it suffices to consider $\theta_2^r < 0.4$.

When $\theta_2^r < 0.4$, the set of feasible allocation payoffs of the seller given the true type profile is

$$\begin{aligned} \widetilde{\Pi}_1(\theta_1, \theta_2) &= \{\hat{u}_1 \in \mathbb{R} : \text{ there exists } \omega \in \Omega \text{ s.t. } h(\omega|(\theta_1, \theta_2)) > 0 \text{ and } \tilde{u}_1(\theta_1, \theta_2; \omega) = \hat{u}_1\} \\ &= \{\hat{u}_1 \in \mathbb{R} : \hat{u}_1 = \theta_1 + 0.5\theta_2 + \omega_1 \text{ for some } \omega_1 \in [-0.1, 0.1]\} \\ &= [\theta_1 + 0.5\theta_2 - 0.1, \theta_1 + 0.5\theta_2 + 0.1]. \end{aligned}$$

Similarly, the set of feasible allocation payoffs of the buyer given the deviation (θ_1, θ_2^r) becomes

$$\begin{aligned} \widetilde{\Pi}_1(\theta_1, \theta_2^r) &= \{\hat{u}_1 \in \mathbb{R} : \text{ there exists } \omega \in \Omega \text{ s.t. } h(\omega|(\theta_1, \theta_2^r)) > 0 \text{ and } \tilde{u}_1(\theta_1, \theta_2^r; \omega) = \hat{u}_1\} \\ &= \{\hat{u}_1 \in \mathbb{R} : \hat{u}_1 = \theta_1 + 0.5\theta_2^r + \omega_1 \text{ where } \omega_1 \in [-0.1, 0.1]\} \\ &= [\theta_1 + 0.5\theta_2^r - 0.1, \theta_1 + 0.5\theta_2^r + 0.1]. \end{aligned}$$

Suppose $\theta_2^r < \theta_2$ without loss of generality. Then, there exist some allocation payoffs of the seller such that they are feasible given the true type profile (θ_1, θ_2), but they are not feasible given the deviation (θ_1, θ_2^r), which is,

$$\begin{cases} [\theta_1 + 0.5\theta_2 - 0.1, \theta_1 + 0.5\theta_2 + 0.1] & \text{if } \theta_1 + 0.5\theta_2^r + 0.1 < \theta_1 + 0.5\theta_2 - 0.1 \\ (\theta_1 + 0.5\theta_2^r + 0.1, \theta_1 + 0.5\theta_2 + 0.1] & \text{otherwise.} \end{cases}$$

As a result, the set of residual uncertainties under which the seller's allocation payoffs are feasible given the true type profile (θ_1, θ_2), but not feasible given the deviation (θ_1, θ_2^r), is as follows:

$$\Omega^\psi(\theta_2, \theta_1; \theta_2^r) = \begin{cases} [-0.1, 0.1]^2 & \text{if } \theta_1 + 0.5\theta_2^r + 0.1 < \theta_1 + 0.5\theta_2 - 0.1, \\ \{(\omega_1, \omega_2) : \omega_1 \in (0.5(\theta_2^r - \theta_2) + 0.1, 0.1] \text{ and } \omega_2 \in [-0.1, 0.1]\} & \text{otherwise.} \end{cases}$$

Therefore, for every $\theta_2 \in \Theta_2$, if $\theta_2^r < 0.4$, there exists a positive-measure set of residual uncertainties under which the seller's allocation payoffs are feasible given the true type profile (θ_1, θ_2), but not feasible given the deviation (θ_1, θ_2^r). Recall that Condition (ii) of Assumption 1 is also satisfied. we conclude that Assumption 1 is satisfied for the buyer. $\square$

We conclude that Assumption 1 is satisfied in Example 5. This completes the proof of Step 3-1. $\square$

Step 3-2: Assumption 3 is violated for x^ε in Example 5.

Proof. Consider the buyer of type $\theta_2 \geq 0.4$. If both agents report truthfully, then the buyer's expected utility in the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$ is

$$\int_{\Theta_1} x^\varepsilon(\theta_1, \theta_2) \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2)\right) dF(\theta_1) = 0.$$

Note that the expected utility of the buyer under truth-telling is identical to that in the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$ with $\psi = 0$, which is $U_2^{\varepsilon,\psi_0}(\theta_2; \theta_2)$. If he deviates to $\theta_2^r \in [0.4, \theta_2)$, his expected utility in the approximate GS mechanism $(\Theta, \Pi, x^\varepsilon, t^{\varepsilon,GS}, \psi)$ becomes

$$U_2^{\varepsilon,\psi_0}(\theta_2; \theta_2^r) = \int_{\Theta_1} x^\varepsilon(\theta_1, \theta_2^r) \left(\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_2^e(\theta_1, \theta_2^r)\right) dF(\theta_1) > 0,$$

where the inequality follows because the buyer's valuation is strictly increasing in his own type. Hence, the buyer of type $\theta_2 \geq 0.4$ has an incentive to deviate to $\theta_2^r \in [0.4, \theta_2)$. However, the seller's true allocation payoffs equal $\theta_1 + \omega_1$, which is independent of the buyer's type. In other words, all feasible allocation payoffs of the seller given the true type profile (θ_1, θ_2) are feasible given the deviation (θ_1, θ_2^r). Hence, Assumption 3 is violated. $\square$

This completes the proof of Step 3. $\square$

Step 4: We provide an example where Assumption 3 holds for x^ε, but Assumption 1 is violated for x^* when $\int_{\Theta_2^{**}} dF(\theta_2) > 0$.

Proof. We use Example 3 in Section 4.5. Recall that in Example 3, each agent's valuation for the object when the true type profile is (θ_1, θ_2) and the residual uncertainty is $\omega = (\omega_1, \omega_2) \in \Omega$ is given as follows:

$$\begin{aligned} \tilde{u}_1(\theta_1, \theta_2; \omega) &= \theta_1 + 2\theta_2 + \omega_1; \\ \tilde{u}_2(\theta_1, \theta_2; \omega) &= 0.5 + 2\theta_1 + \theta_2 + \omega_2. \end{aligned}$$

We compute

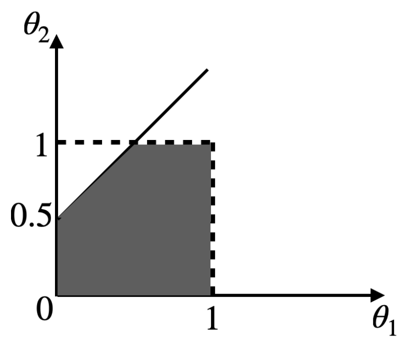
$$\tilde{u}_2^e(\theta_1, \theta_2) - \tilde{u}_1^e(\theta_1, \theta_2) = 0.5 + \theta_1 - \theta_2,$$

implying that the efficient decision rule dictates

$$x^*(\theta_1, \theta_2) = \begin{cases} 1 & \text{if } \theta_2 \leq 0.5 + \theta_1; \\ 0 & \text{otherwise} \end{cases}$$

The following figure illustrates the decision at different type profiles in this example; in particular, the shaded region represents $\Theta_* = \{(\theta_1, \theta_2) \in [0,1]^2 : x^*(\theta_1, \theta_2) = 1\}$.

Figure 8



We already show in Claim 4 that Assumption 1 is violated in Example 3. We will show below, by means of an example, that Assumption 3 holds for x^ε. The proposed example exhibits the following features. The buyer of type $\theta_2 > 0.5$ has an incentive to deviate to $\theta_2^r < 0.5$ in the GS mechanism $(\Theta, \Pi, x^*, t^{GS}, \psi)$. This deviation induces that the good is never allocated to the seller after the deviation, which guarantees that the buyer's deviation cannot be detected. Thus, Assumption 1 for x^* is violated. On the contrary, this deviation can be detected under x^ε because the good is always allocated to the seller with probability at least $\varepsilon > 0$ under x^ε.

We divide the proof of Step 4 into two substeps.

Step 4-1: Assumption 3 is satisfied for the seller in Example 3.

Proof. Observe from Figure 8 that in Example 3, for all $\theta_1 \in \Theta_1$, there exists a positive-measure set $\Theta_2^+ \subseteq \Theta_2$ such that $x^\varepsilon(\theta_1, \theta_2) = (1-\varepsilon)x^*(\theta_1, \theta_2) > 0$ for all $\theta_2 \in \Theta_2^+$. Hence, Condition (ii) of Assumption 3 is satisfied for the seller. It remains to verify that Condition (i) of Assumption 1 is satisfied as well, that is, $\int_{\Omega^\psi(\theta_1,\theta_2;\theta_1^r)} dH\left(\omega|(\theta_1, \theta_2)\right) > 0$ for every $\theta_1, \theta_1^r \in \Theta_1$ and every $\theta_2 \in \Theta_2^+$.

Consider the seller of type θ_1. The set of feasible allocation payoffs of the buyer given the true type profile (θ_1, θ_2) is

$$\begin{aligned} \widetilde{\Pi}_2(\theta_1, \theta_2) &= \{\hat{u}_2 \in \mathbb{R} : \text{ there exists } \omega \in \Omega \text{ s.t. } h\left(\omega|(\theta_1, \theta_2)\right) > 0 \text{ and } \tilde{u}_2(\theta_1, \theta_2; \omega) = \hat{u}_2\} \\ &= \{\hat{u}_2 \in \mathbb{R} : \hat{u}_2 = 0.5 + 2\theta_1 + \theta_2 + \omega_2 \text{ for some } \omega_2 \in [-0.1, 0.1]\} \\ &= [0.5 + 2\theta_1 + \theta_2 - 0.1, 0.5 + 2\theta_1 + \theta_2 + 0.1]. \end{aligned}$$

Similarly, the set of feasible allocation payoffs of the buyer given the deviation (θ_1^r, θ_2) becomes

$$\begin{aligned} \widetilde{\Pi}_2(\theta_1^r, \theta_2) &= \{\hat{u}_2 \in \mathbb{R} : \text{ there exists } \omega \in \Omega \text{ s.t. } h\left(\omega|(\theta_1^r, \theta_2)\right) > 0 \text{ and } \tilde{u}_2(\theta_1^r, \theta_2; \omega) = \hat{u}_2\} \\ &= \{\hat{u}_2 \in \mathbb{R} : \hat{u}_2 = 0.5 + 2\theta_1^r + \theta_2 + \omega_2 \text{ where } \omega_2 \in [-0.1, 0.1]\} \\ &= [0.5 + 2\theta_1^r + \theta_2 - 0.1, 0.5 + 2\theta_1^r + \theta_2 + 0.1]. \end{aligned}$$

Suppose $\theta_1^r > \theta_1$ without loss of generality. Then, there exist some allocation payoffs of the buyer such that they are feasible given the true type profile (θ_1, θ_2), but they are not feasible given the deviation (θ_1^r, θ_2), which is,

$$\begin{cases} [0.5 + 2\theta_1 + \theta_2 - 0.1, 0.5 + 2\theta_1 + \theta_2 + 0.1] & \text{if } 0.5 + 2\theta_1^r + \theta_2 - 0.1 > 0.5 + 2\theta_1 + \theta_2 + 0.1 \\ [0.5 + 2\theta_1 + \theta_2 - 0.1, 0.5 + 2\theta_1^r + \theta_2 - 0.1) & \text{otherwise.} \end{cases}$$

As a result, the set of residual uncertainties under which the buyer's allocation payoffs are feasible given the true type profile (θ_1, θ_2), but not feasible given the deviation (θ_1^r, θ_2), is as follows:

$$\Omega^\psi(\theta_1, \theta_2; \theta_1^r) = \begin{cases} [-0.1, 0.1]^2 \\ \quad \text{if } 0.5 + 2\theta_1^r + \theta_2 - 0.1 > 0.5 + 2\theta_1 + \theta_2 + 0.1, \\ \{(\omega_1, \omega_2) : \omega_1 \in [-0.1, 0.1] \text{ and } \omega_2 \in [-0.1, 2(\theta_1^r - \theta_1) - 0.1)\} \\ \quad \text{otherwise.} \end{cases}$$

Therefore, for every $\theta_1, \theta_1^r \in \Theta_1$ and every $\theta_2 \in \Theta_2$, there exists a positive-measure set of residual uncertainties under which the buyer's allocation payoffs are feasible given the true type profile (θ_1, θ_2), but not feasible given the deviation (θ_1^r, θ_2), implying that condition (i) of Assumption 3 is satisfied for all $\theta_1, \theta_1^r \in \Theta_1$ and all $\theta_2 \in \Theta_2$. We conclude that Assumption 3 is satisfied for the seller. $\square$

Step 4-2: Assumption 3 is satisfied for the buyer in Example 3.

Proof. Recall that for all $(\theta_1, \theta_2) \in \Theta$, the probability that the seller keeps the good under decision rule x^ε is at least as high as ε. Hence, Condition (ii) of Assumption 3 is always satisfied for the buyer. It remains to verify that Condition (i) of Assumption 3 is satisfied as well, that is, for all $\theta_2, \theta_2^r \in \Theta_2$, there exists a positive-measure set $\Theta_1^+ \subseteq \Theta_1$ such that $\int_{\Omega^\psi(\theta_2,\theta_1;\theta_2^r)} dF(\theta_2) > 0$ for all $\theta_1 \in \Theta_1^+$.

Consider the buyer of type θ_2. The set of feasible allocation payoffs of the seller given the true type profile (θ_1, θ_2) is

$$\begin{aligned}
\widetilde{\Pi}_1(\theta_1, \theta_2) &= \{\hat{u}_1 \in \mathbb{R} : \text{ there exists } \omega \in \Omega \text{ s.t. } h\left(\omega|(\theta_1, \theta_2)\right) > 0 \text{ and } \tilde{u}_1(\theta_1, \theta_2; \omega) = \hat{u}_1\} \\
&= \{\hat{u}_1 \in \mathbb{R} : \hat{u}_1 = \theta_1 + 2\theta_2 + \omega_1 \text{ for some } \omega_1 \in [-0.1, 0.1]\} \\
&= [\theta_1 + 2\theta_2 - 0.1, \theta_1 + 2\theta_2 + 0.1].
\end{aligned}$$

Similarly, the set of feasible allocation payoffs of the seller given the deviation (θ_1, θ_2^r) becomes

$$\begin{aligned}
\widetilde{\Pi}_1(\theta_1, \theta_2^r) &= \{\hat{u}_1 \in \mathbb{R} : \text{ there exists } \omega \in \Omega \text{ s.t. } h\left(\omega|(\theta_1, \theta_2^r)\right) > 0 \text{ and } \tilde{u}_1(\theta_1, \theta_2^r; \omega) = \hat{u}_1\} \\
&= \{\hat{u}_1 \in \mathbb{R} : \hat{u}_1 = \theta_1 + 2\theta_2^r + \omega_1 \text{ where } \omega_1 \in [-0.1, 0.1]\} \\
&= [\theta_1 + 2\theta_2^r - 0.1, \theta_1 + 2\theta_2^r + 0.1].
\end{aligned}$$

Suppose $\theta_2^r < \theta_2$ without loss of generality. Then, there exist some allocation payoffs of the seller such that they are feasible given the true type profile (θ_1, θ_2), but they are not feasible given the deviation (θ_1, θ_2^r), which is,

$$\begin{cases}
[\theta_1 + 2\theta_2 - 0.1, \theta_1 + 2\theta_2 + 0.1] & \text{if } \theta_1 + 2\theta_2^r + 0.1 < \theta_1 + 2\theta_2 - 0.1, \\
(\theta_1 + 2\theta_2^r + 0.1, \theta_1 + 2\theta_2 + 0.1] & \text{otherwise.}
\end{cases}$$

As a result, the set of residual uncertainties under which the seller's allocation payoffs are feasible given the true type profile (θ_1, θ_2), but not feasible given the deviation (θ_1, θ_2^r), is as follows:

$$\Omega^\psi(\theta_2, \theta_1; \theta_2^r) = \begin{cases}
[-0.1, 0.1]^2 & \text{if } \theta_1 + 2\theta_2^r + 0.1 < \theta_1 + 2\theta_2 - 0.1, \\
\{(\omega_1, \omega_2) : \omega_1 \in (2(\theta_2^r - \theta_2) + 0.1, 0.1] \text{ and } \omega_2 \in [-0.1, 0.1]\} & \text{otherwise.}
\end{cases}$$

Therefore, for every $\theta_1 \in \Theta_1$ and every $\theta_2, \theta_2^r \in \Theta_2$, there exists a positive-measure set of residual uncertainties under which the seller's allocation payoffs are feasible given the true type profile (θ_1, θ_2), but not feasible given the deviation (θ_1, θ_2^r). Recall that Condition (ii) of Assumption 3 is also satisfied. We conclude that Assumption 3 is satisfied for the buyer. □

We conclude that Assumption 3 holds for x^{ε} in Example 3. This completes the proof of Step 4. $\square$

This completes the proof of the claim. $\square$

References

[1] Fieseler, K., T. Kittsteiner, and B. Moldovanu, "Partnerships, Lemons, and Efficient Trade," *Journal of Economic Theory*, vol. 113, (2003), 223-234.

[2] Galavotti, S., N. Muto, and D. Oyama, "On efficient partnership dissolution under ex post individual rationality," *Economic Theory*, vol. 48, (2011), 87-123.

[3] Gresik, T.A., "Ex Ante Incentive Efficient Trading Mechanisms without the Private Valuation Restriction," *Journal of Economic Theory*, vol. 55, (1991), 41-63.

[4] Jehiel, P. and B. Moldovanu, "Allocative and informational externalities in auctions and related mechanisms," in *The Proceedings of the 9th World Congress of the Econometric Society*, edited by Richard Blundell, Whitney Newey, and Torsten Persson, Cambridge University Press, 2006.

[5] Krishna, V., and M. Perry, "Efficient Mechanism Design," *Working Paper*, 2000.

[6] Kunimoto, T., and C. Zhang, "Efficient Bilateral Trade via Two-Stage Mechanisms: Comparison between One-Sided and Two-Sided Asymmetric Information Environments," *Journal of Mathematical Economics*, vol. 101, (2022), 102714.

[7] Makowski, L. and C. Mezzetti, "Bayesian and Weakly Robust First Best Mechanisms: Characterizations," *Journal of Economic Theory*, vol. 64, (1994), 500-519.

[8] Matsushima, H. and S. Noda, "Mechanism Design with Blockchain Enforcement," *Working Paper*, 2023.

[9] Mezzetti, C., "Auction Design with Interdependent Valuations: The Generalized Revelation Principle, Efficiency, Full Surplus Extraction and Information Acquisition," *Working Paper*, 2003.

[10] Mezzetti, C., "Mechanism Design with Interdependent Valuations: Efficiency," *Econometrica*, vol. 72, (2004), 1617-1626.

[11] Mezzetti, C., "Mechanism Design with Interdependent Valuations: Surplus Extraction," *Economic Theory*, vol. 31, (2007), 473-488.

[12] Myerson, R.B. and M.A. Satterthwaite, "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, vol. 29, (1983), 265-281.

[13] Nath, S. and O. Zoeter, "A strict ex-post incentive compatible mechanism for interdependent valuations," *Economics Letters*, vol. 121, (2013), 321-325.

[14] Osborne, M.J. and A. Rubinstein, *A Course in Game Theory*, The MIT Press, 1994.

[15] Williams, S.R., "A Characterization of Efficient, Bayesian Incentive Compatible Mechanisms," *Economic Theory*, vol. 14, (1999), 155-180.