4-2020

Essays on a mechanism design approach to the problem of bilateral trade and public good provision

Cuiling ZHANG
Singapore Management University

SINGAPORE MANAGEMENT UNIVERSITY

PHD DISSERTATION

Essays on a Mechanism Design Approach to the Problem of Bilateral Trade and Public Good Provision

Cuiling Zhang

supervised by

Professor TAKASHI KUNIMOTO

June 22, 2020

Essays on a Mechanism Design Approach to the Problem of Bilateral Trade and Public Good Provision

by

Cuiling Zhang

Submitted to School of Economics in partial fulfillment of

the requirements for the Degree of Doctor of Philosophy in Economics

Dissertation Committee:

Takashi Kunimoto (Supervisor/Chair)
Associate Professor of Economics
Singapore Management University

Shurojit Chatterji
Professor of Economics
Lee Kong Chian Fellow
Singapore Management University

Atsushi Kajii
Distinguished Term Professor of Economics
Singapore Management University

Yi-Chun Chen
Associate Professor of Economics
National University of Singapore

Singapore Management University

2020

Declaration

I hereby declare that this PhD dissertation is my original work and it has been written by me in its entirety.

I have duly acknowledged all the sources of information which have been used in this dissertation.

This PhD dissertation has also not been submitted for any degree in any university previously.

Cuiling Zhang

22 June 2020

Abstract

The dissertation consists of three chapters which studies a mechanism design approach to the problem of bilateral trade and public good provision.

Chapter 1 characterizes mechanisms satisfying Bayesian incentive compatibility (BIC) and interim individual rationality (IIR) in the classical public good provision problem. We propose a stress test for the results in the standard continuum type space by subjecting them to a finite type space. The main contribution of this paper is to propose a set of techniques that allow us to characterize the efficient and optimal mechanisms in a discrete setup. Using these techniques, we conclude that many of the known results gained within the standard continuum type space also hold when it is replaced by a discrete type space.

Chapter 2 seeks for more positive results by employing two-stage mechanisms (Mezzetti (2004)), as efficient, voluntary bilateral trade is generally not incentive compatible in an interdependent-value environment (Fieseler, Kittsteiner, Moldovanu (2003) and Gresik (1991)). First, we show by means of a stylized example that the generalized two-stage Groves mechanism never guarantees voluntary trade, while it satisfies efficiency and incentive compatibility. In a general environment, we next propose Condition α under which there exists a two-stage incentive compatible mechanism implementing an efficient, voluntary trade. Third, within the same example, we confirm that our Condition α is very weak because it holds as long as the buyer's degree of interdependence of preferences is not too high relative to the seller's counterpart. Finally, we show by the same example that if Condition α is violated, our proposed two-stage mechanism fails to achieve voluntary trade.

Chapter 3 clarifies how the interdependence in valuations and correlation of types across agents affect the possibility of efficient, voluntary bilateral trade in a model with discrete types, as efficient, voluntary bilateral trades are generally not incentive compatible in a private-value model with independently distributed continuous types (Myerson and Satterthwaite (1983)). First, we identify a necessary condition for the existence of incentive compatible mechanisms inducing an efficient and voluntary trade in a finite type model. Second, we show that the identified necessary condition becomes sufficient for a two-type model. Using this characterization in a model with linear valuations and two

types, we next conduct the comparative statics for how possibility results rely on the interdependence and correlation. Third, using the linear programming approach, we establish the general existence of an efficient, incentive compatible trade in a model with two types. This suggests that voluntary trade becomes a stringent requirement in an interdependent values model with correlated signals.

Contents

Acknowledgement

First and foremost I would like to express my deepest gratitude to my supervisor Professor Takashi Kunimoto. It has been an honor to be his Ph.D. student. He spent endless hours discussing my research papers and giving me excellent suggestions. I will always be indebted to Professor Kunimoto for his contributions of time, patience, and ideas to make my Ph.D. experience productive.

I would also like to thank my thesis committee: Professor Shurojit Chatterji, Professor Atsushi Kajii, and Professor Yi-Chun Chen for their invaluable comments every time I presented the paper. Many thanks also to Professor Jingyi Xue who has been continuously supportive all these years, and Professor Jiangtao Li who has provided extremely helpful comments on my dissertation.

I am indebted to all my friends and fellow students in SMU. In particular, I wish to thank my seniors Huaxia Zeng and Peng Liu for their kindness and willingness to share experience with me.

I am grateful to our programme director Professor Anthony Tay for his great leadership and to Professor Jun Yu for his support when I was on the market. I am also grateful to all those at the General Office, especially Ms. Qiu Ling Thor, Ms. Amelia Tan and others who were always so helpful and provided me with their assistance throughout my Ph.D. journey.

Last but not least, I would like to thank my family for all their love and encouragement. They have always had faith in me and encouraged me to follow my dreams. Special thanks go to Yiu Lim Lui whose accompany is so appreciated. Thank you.

1 On Incentive Compatible, Individually Rational Public Good Provision Mechanisms

1.1 Introduction

This paper revisits the classical public good provision problem in which a group of agents have to decide whether to produce some indivisible and non-excludable public good. This has been a central application of the theory of mechanism design (See, for example, Krishna (Chapter 5, 2009), Mas-Colell, Whinston, and Green (Chapter 23, 1995), Börgers (Chapters 3, 2015) for this). To analyze this problem, many papers in the literature consider the model of ex ante identical agents with a continuous, closed interval of types.[1] In what follows, we call such a model the *standard model*. One practical benefit of using the standard model is that we can appeal to the revenue equivalence theorem, which reduces the search for an appropriate mechanism to the class of the well-known Vickrey-Clarke-Groves (VCG) mechanism.[2] For example, we see this power of reduction in Krishna and Perry (2000) and Williams (1999). This paper, on the other hand, proposes a "stress test" for the known results by considering a *finite* discretization of the standard model of a continuous, closed interval of types. The difficulty of conducting this test lies in the fact that we cannot appeal to the optimality of the VCG mechanism established in the literature, which exploits the continuum type space assumption. The main contribution of this paper is to re-establish the optimality of a variant of the VCG mechanism in a discrete setup.[3] Exploiting the optimality of the VCG-like mechanism, we conclude that many of the standard results in the classical public good provision problem gained within a continuum type space also hold when it is replaced by a discrete type space.

As in the standard model, we assume that each agent's type, i.e., preferences for public good, is chosen independently from an identical distribution over finitely many values.

[1] The reader is referred to Chapter 3.3 of Börgers (2015) for the textbook treatment of the classical public good problem with identical agents whose type space constitutes a continuous closed interval on the real line.

[2] The VCG mechanism is based on the contribution of Vickrey (1961), Clarke (1971), and Groves (1973).

[3] Kos and Manea (2009) conducted a similar analysis in the context of bilateral trade setup.

Throughout this paper, we impose incentive compatibility and individual rationality on *direct mechanisms*, which maps each type profile to the probability of providing a public good and monetary transfers across agents. A direct mechanism satisfies *Bayesian incentive compatibility* (BIC) if all agents' announcing their true type constitutes a Bayesian Nash equilibrium of the direct mechanism. By the celebrated revelation principle, we focus on direct mechanisms without loss of generality so that we call a direct mechanism simply a mechanism. A mechanism satisfies *interim individual rationality* (IIR) if each agent guarantees a utility of zero (utility of non-participation), provided that all agents announce their types truthfully. We introduce two more requirements we sometimes impose on the mechanisms. A mechanism satisfies *decision efficiency* (EFF) if the public good is provided if and only if the total surplus from providing the public good is at least as high as the cost of the public good. A mechanism satisfies *ex post budget balance* (BB) if it satisfies budget balance at any state.

To state our results below, we introduce the following categories. By the *trivial* cases we mean that it is always efficient to provide a public good or it is always efficient not to provide a public good. We call any other case a *nontrivial case*. We obtain the following Bayesian implementation results in our discrete setup.

- Theorem 1: There exists a variant of the VCG mechanism that maximizes the ex ante budget surplus among all the mechanisms satisfying BIC, IIR, and EFF.
- Corollary 1: There exists a mechanism satisfying BIC, IIR, EFF, and BB if and only if a variant of the VCG mechanism results in ex ante budget surplus.
- Theorem 2: In any nontrivial case, as the population size gets large, any mechanism satisfying BIC, IIR, EFF results in the per capita ex ante budget deficit.
- Theorem 3: There exists a variant of the VCG mechanism that maximizes the ex ante budget surplus among all the mechanisms satisfying BIC and IIR.
- Theorem 4: In any nontrival case, as the population size gets large, the ex ante probability that the public good is provided converges to zero in any mechanism

satisfying BIC, IIR, and BB.[4]

Theorem 1 is the key result of this paper, which allows us to obtain Corollary 1 and Theorem 2. This is a powerful result because it can reduce our search for a mechanism satisfying BIC, IIR, and EFF to the *VCG' mechanism* (a variant of the VCG mechanism) which we appropriately adapt from the same mechanism of Kos and Manea (2009) in a bilateral trade environment.

Corollary 1 is considered a discrete type space counterpart of Theorem 2 of Krishna and Perry (2000) in a public good setup. The difficulty of obtaining this result lies in the fact that we cannot appeal to the optimality of the VCG mechanism established in Krishna and Perry (2000) and Williams (1999), who exploit the continuum type space assumption. We rather appeal to the optimality of the VCG' mechanism by our Theorem 1.

Next, Theorem 2 characterizes the implications of our Corollary 1 in the context of large economies. It shows that in all nontrivial cases, "any" mechanism satisfying BIC, IIR, and EFF results in *per capita* ex ante budget deficit when the population size of the economy gets large enough. Thus, for any mechanism satisfying BIC, IIR, and EFF in large economies, we must accept not only the ex ante budget deficit but also the *per capita* ex ante budget deficit. The basic logic for Theorem 2 goes as follows. Each agent of a higher type can lower his payment by announcing a lower type. The only incentive to not do so is that the agent is pivotal, i.e., his announcement will change the probability that the public good is provided. However, in large economies, the probability that an agent is pivotal converges to zero and thus it is prohibitively costly to induce all agents of higher types to tell the truth in large economies. For this result, however, we make an additional assumption, which says that the ex ante probability that the efficient decision rule provides the public good converges to some positive probability as the population size gets infinite. This assumption strikes us as being innocuous because if it is not satisfied, the provision of the public good will not be expected in large economies.

Then, we are left with the case where the ex ante probability that the efficient decision rule provides the public good converges to zero as the population size gets infinite. In this case, we show in our Proposition 1 that there exists a mechanism satisfying BIC, IIR,

[4]For this result, we assume that there are only two types.

EFF, and BB and as the size of the economy gets large, the probability that the public good is provided converges to zero in any such mechanism. This is consistent with the analysis in Section 4.2 of Ledyard and Palfrey (1994).

To understand how large the economy should be for Theorem 2, we also conduct a simulation analysis for this negative result. Our simulation result roughly suggests that there are no mechanisms satisfying BIC, IIR, EFF, and BB even for a relatively small size of the economy. Sometimes, even five agents are large enough to establish the impossibility result. However, we can sometimes promote the case for small economies. For example, if a society is going to decide whether to provide a public good or not, the social planner selects a small group of representative agents and design a mechanism as suggested in our paper. In that case, our mechanism will satisfy BIC, IIR, EFF, and BB.

Given the difficulty of imposing all the four properties, BIC, IIR, EFF, and BB on the mechanisms, we now drop decision efficiency (EFF) from the requirements. Theorem 3 is considered a generalization of Theorem 1 so that we can reduce our search for a mechanism satisfying BIC and IIR to the *tight mechanism* which we appropriately adapt from the same mechanism of Kos and Manea (2009) in a bilateral trade environment. Note that when we insist on decision efficiency, the tight mechanism is reduced to the VCG' mechanism. Using Theorem 3, we show that there are mechanisms satisfying BIC, IIR, and BB if and only if the tight mechanism results in ex ante budget surplus (Corollary 3).

Our Theorem 4 is concerned with the implications of large economies in a two-type environment. That is, in any nontrivial case, as the size of the economy gets large, the ex ante probability that the public good is provided converges to zero in any mechanism satisfying BIC, IIR, and BB. This theorem is considered a discrete type space counterpart of Theorem 2 of Mailath and Postlewaite (1990). Therefore, this might even suggest the necessity of mandatory payment of taxes as opposed to voluntary contribution, which is embodied by the individual rationality constraint. Our rationale for focusing on a two-type setup is based on the fact that the negative result is obtained in the continuum type space and by continuity, there is likely to be a similar negative result on a large finite type space, which is considered an approximation of the continuum type space. Of course, we benefit a lot from the two-type setup in terms of tractability.

Finally, we strengthen BIC and IIR into dominant strategy incentive compatibility (DSIC) and ex post individual rationality (EPIR), respectively.[5] One benefit of doing so is that we can completely drop any distributional assumption about types and allow for any degree of correlation. We obtain the following dominant strategy implementation result.

- Proposition 2 In any nontrivial case, there are no mechanisms satisfying DSIC, EPIR, EFF, and BB, regardless of the size of the economy.

This result is a discrete type space counterpart of Theorem 7 of Green and Laffont (1977) and therefore, also considered a stress test for the negative result of dominant strategy implementation in the public good provision problem.[6]

The rest of the paper is organized as follows. In Section 1.2, we introduce the general concepts and notation used throughout the paper. Section 1.3 identifies a necessary and sufficient condition for the existence of mechanisms satisfying BIC, IIR, EFF, and BB, investigates the implication of the results in large economies, and finally presents the simulation results. In Section 1.4, we drop EFF from the requirements and characterize a condition for the existence of mechanisms satisfying BIC, IIR, and BB and investigate its implication in large economies. In Section 1.5, we replace BIC and IIR with DSIC and EPIR, respectively so that we investigate the corresponding implications. Section 1.6 concludes. The Appendix contains all the proofs omitted from the main body of the paper.

1.2 Preliminaries

There are N agents and we denote by $\mathcal{N} = \{1, \ldots, N\}$ the set of agents. We assume $N \geq 2$ throughout the paper. A group of N agents must decide whether to undertake the public project and if undertaken, how to distribute the costs of the project among the

[5]The reader is referred to Chapter 4.3 of Börgers (2015) for the textbook treatment of the public good problem using DSIC and EPIR. Once again, a big difference from our paper is that Börgers' type space is assumed to be a closed interval in the real line.

[6]Green and Laffont (1977) need to include non-separable preferences as part of the domain in establishing their Theorem 7. The reader is referred to Green and Laffont (1977) for the exact nature of their rich environments. On the contrary, we do not need this richness at all.

members of the group. Each agent i has $M \geqslant 2$ possible types $\theta_i \in \Theta \equiv \{\theta^1, \cdots, \theta^M\}$ such that $0 \leq \theta^1 < \cdots < \theta^M$(i.e., θ^1 is allowed to be zero). We assume that every agent i of the same type θ^m attach the common value θ^m to the public project . We further assume that each agent's type is private information. Denote by $\Theta^N = \{\theta^1, \cdots, \theta^M\}^N$ the set of possible type profiles. The types are independently drawn from an identical distribution where $P(\theta^m)$ denotes the probability that θ^m is chosen. Therefore, there is a common prior P^N over Θ^N such that for each $\theta = (\theta_1, \ldots, \theta_N) \in \Theta^N$,

$$P^N(\theta) \equiv P(\theta_1) \times \cdots \times P(\theta_N).$$

The independence of types is essential for our results.[7] Preferences of each agent depend upon whether or not the public project is implemented and how much of the monetary payment is incurred by that agent. Agents evaluate lotteries over outcomes using expected utility. If the public project is built with probability $q \in [0, 1]$ and agent i makes a payment t_i to the planner, then i's preferences can be represented by

$$v^i = q\theta_i - t_i.$$

This formulation assumes that each agent's preferences are quasilinear in money and each individual is risk neutral.

A *direct* mechanism is defined as a triplet $\Gamma = (\Theta, x, (t_i))_{i \in \mathcal{N}}$ where $\Theta = \{\theta^1, \cdots, \theta^M\}$ is the set of actions available to agent i, i.e., each agent is asked to reveal his type; $x : \Theta^N \to [0, 1]$ is the *decision rule* which specifies the probability that the public good is provided; and $t_i : \Theta^N \to \mathbb{R}$ is the payment or subsidy to agent i and $t = (t_1, \ldots, t_N)$ is called the *transfer rule*. By the well known revelation principle, we lose nothing to focus on direct mechanisms. In what follows, we denote by (x, t) a direct mechanism or simply a mechanism.

Definition 1 *A mechanism (x, t) satisfies Bayesian incentive compatibility (BIC) if, for each $i \in \mathcal{N}$, $\theta_i, \theta_i^{'} \in \Theta$,*

$$\sum_{\theta_{-i} \in \Theta^{N-1}} P^{N-1}(\theta_{-i}) \left[\theta_i x(\theta_i, \theta_{-i}) - t_i(\theta_i, \theta_{-i}) - \theta_i x(\theta_i^{'}, \theta_{-i}) + t_i(\theta_i^{'}, \theta_{-i})\right] \geq 0,$$

where $P^{N-1}(\theta_{-i}) \equiv \times_{j \neq i} P(\theta_j)$.

[7] We conjecture that the identical distribution assumption is only needed for the ease of computation.

We introduce a stronger version of incentive compatibility.

Definition 2 *A mechanism (x, t) satisfies dominant strategy incentive compatibility (DSIC) if, for each $i \in \mathcal{N}$, $\theta \in \Theta^N$ and $\theta_i^{'} \in \Theta$,*

$$\theta_i x(\theta_i, \theta_{-i}) - t_i(\theta_i, \theta_{-i}) \geq \theta_i x(\theta_i^{'}, \theta_{-i}) - t_i(\theta_i^{'}, \theta_{-i}).$$

Dominant strategy incentive compatibility does not need to make any distributional assumption about how each agent's type is realized and because of this property, it is stronger than Bayesian incentive compatibility.

When there are N agents in the economy, providing the public good will incur a cost equal to $c(N)$ which is assumed to be an increasing function in N. Throughout the paper, we further assume that $\theta^1 < c(N)/N \leq \theta^M$. We do not discuss the case where $\theta^1 = c(N)/N$ because it is considered a trivial case in the sense that the public good should always be provided. This implies that the non-rivalry property of a pure public good does not hold here. This is consistent with the setup of Mailath and Postlewaite (1990).[8]

Definition 3 *A mechanism (x, t) satisfies decision efficiency (EFF) if, for each $\theta \in \Theta^N$,*

$$x(\theta) = \begin{cases} 1 & \text{if } \sum_{i \in \mathcal{N}} \theta_i \geqslant c(N) \\ 0 & \text{otherwise}. \end{cases}$$

In what follows, we denote by $x^*(\cdot)$ the efficient decision rule.

Definition 4 *A mechanism (x, t) satisfies the ex post budget balance (BB) if, for each $\theta \in \Theta^N$,*

$$\sum_{i \in \mathcal{N}} t_i(\theta) = c(N) \cdot x(\theta).$$

Remark: Since $x : \Theta^N \to [0, 1]$ is a stochastic decision rule, $t_i(\theta)$ here is interpreted as the expected transfer, i.e., $t_i(\theta) = x(\theta)\tilde{t}_i(\theta)$ where $\tilde{t}_i(\theta)$ is agent i's payment if public good is provided at profile θ. This change is inconsequential because the agents

[8] Hellwig (2003) points out that this assumption is crucial for the result. Indeed, he completely overturns the result of Mailath and Postlewaite (1990) by isolating the effect of changes in the number of participants, while keeping cost technologies fixed.

are assumed to be risk-neutral. Then the ex post budget balance constraint (BB) is the same as before, i.e., $\sum_{i \in \mathcal{N}} t_i(\theta) = c(N)x(\theta)$, but it has a different implication: if $\sum_{i \in \mathcal{N}} t_i(\theta) = c(N)x(\theta)$, then, for each $\theta \in \Theta^N$,

$$x(\theta) \left(\sum_{i \in \mathcal{N}} \tilde{t}_i(\theta) - c(N) \right) = 0,$$

which implies

$$\sum_{i \in \mathcal{N}} \tilde{t}_i(\theta) = c(N) \text{ when } x(\theta) > 0.$$

The literature often assumes that every agent must participate in the mechanism; otherwise, he obtains a utility of zero. See, for example, Börgers (2015) for the details.[9]

With this, we introduce the individual rationality constraint.

Definition 5 *A mechanism (x, t) satisfies the interim individual rationality (IIR) if, for each $i \in \mathcal{N}$ and $\theta_i \in \Theta$,*

$$\sum_{\theta_{-i} \in \Theta^{N-1}} P^{N-1}(\theta_{-i}) \left[\theta_i x(\theta_i, \theta_{-i}) - t_i(\theta_i, \theta_{-i}) \right] \geq 0.$$

We introduce a stronger version of individual rationality.

Definition 6 *A mechanism (x, t) satisfies ex post individual rationality (EPIR) if, for each $i \in \mathcal{N}$, $\theta_i \in \Theta$ and $\theta_{-i} \in \Theta^{N-1}$,*

$$\theta_i x(\theta_i, \theta_{-i}) - t_i(\theta_i, \theta_{-i}) \geq 0.$$

Note that ex post individual rationality implies interim individual rationality.

1.3 When the First-Best is Implementable

In this section, we will investigate the existence of mechanisms satisfying BIC, IIR, EFF, and BB.

[9] Saijo and Yamato (1999) assume instead that each agent can not be excluded from the consumption of the public good even if he decides not to participate in the mechanism. Although the individual rationality of Saijo and Yamato (1999) is a lot more demanding than EPIR, we nevertheless establish a few negative results. Thus, we rather stick to our weaker individual rationality. The reader is referred to Saijo and Yamato (1999) for the discussion of their individual rationality constraints. Yenmez (2013) considers a similar constraint in a one-to-one matching environment.

1.3.1 Preliminaries

We call a mechanism (x, t) the *first-best* solution if x is the efficient decision rule and (x, t) satisfies ex post budget balance. Recall that $x^* : \Theta^N \to [0, 1]$ denotes the efficient decision rule. In order for any mechanism (x^*, t) to satisfy ex post budget balance, we set $t_N(\theta) = x^*(\theta) \cdot c(N) - \sum_{i \neq N} t_i(\theta)$. As usual, the vector $\theta = (\theta_1, \theta_2, \cdots, \theta_N)$ denotes the types of all agents and the vector $\theta_{-i} = (\theta_1, \theta_2, \cdots, \theta_{i-1}, \theta_{i+1}, \cdots, \theta_N)$ the types of all agents other than i. The vector $(\theta_i', \theta_{-i}) = (\theta_1, \theta_2, \cdots, \theta_{i-1}, \theta_i', \theta_{i+1}, \cdots, \theta_N)$.

In what follows, we construct the transfer rule $t = (t_1, \ldots, t_N)$ such that the direct mechanism (x^*, t) satisfies BIC, IIR, EFF, and BB.

We consider the following modification of the Vickrey-Clarke-Groves mechanism (VCG'), which was originally proposed by Kos and Manea (2009) in a bilateral trade environment. Here we aim to adapt this to our public good environment.

Definition 7 *A mechanism (x^*, t') is called the VCG' mechanism if for each agent $i \in \mathcal{N}$ and each $\theta \in \Theta^N$,*

$$t_i'(\theta) = \begin{cases} \min\{\tilde{\theta}_i \in \Theta : \sum_{j \neq i} \theta_j + \tilde{\theta}_i \geqslant c(N)\} & \text{if } x^*(\theta) = 1 \\ 0 & \text{otherwise}. \end{cases} \tag{1}$$

In the VCG' mechanism, when the public good is provided, each agent's payment is equal to the lowest possible type which makes sure that the total valuation is higher than the cost.

It is easy to see that the VCG' mechanism satisfies EPIR. We move on to the incentive compatibility of the VCG' mechanism.

Claim 1 *The VCG' mechanism satisfies DSIC.*

Proof: Fix $\theta^m, \theta^n \in \Theta$ such that $\theta^m > \theta^n$ arbitrarily. We write down the DSIC constraints for these two types: for any $\theta_{-i} \in \Theta^{N-1}$,

$$IC_{\theta^m \to \theta^n} : \quad \theta^m x^*(\theta^m, \theta_{-i}) - t_i(\theta^m, \theta_{-i}) \geqslant \theta^m x^*(\theta^n, \theta_{-i}) - t_i(\theta^n, \theta_{-i}),$$

$$IC_{\theta^n \to \theta^m} : \quad \theta^n x^*(\theta^n, \theta_{-i}) - t_i(\theta^n, \theta_{-i}) \geqslant \theta^n x^*(\theta^m, \theta_{-i}) - t_i(\theta^m, \theta_{-i}),$$

where $IC_{\theta^m \to \theta^n}$ stands for the incentive constraint which prevents type θ^m from pretending to be type θ^n. We further rewrite the DSIC constraints: for any $\theta_{-i} \in \Theta^{N-1}$,

$$IC_{\theta^m \to \theta^n}: \quad t_i(\theta^m, \theta_{-i}) - t_i(\theta^n, \theta_{-i}) \leqslant \theta^m \left(x^*(\theta^m, \theta_{-i}) - x^*(\theta^n, \theta_{-i}) \right), \tag{2}$$

$$IC_{\theta^n \to \theta^m}: \quad t_i(\theta^m, \theta_{-i}) - t_i(\theta^n, \theta_{-i}) \geqslant \theta^n \left(x^*(\theta^m, \theta_{-i}) - x^*(\theta^n, \theta_{-i}) \right). \tag{3}$$

We consider the following three cases:

1. if $c(N) - \sum_{j \neq i} \theta_j > \theta^m > \theta^n$, then $x^*(\theta^m, \theta_{-i}) = x^*(\theta^n, \theta_{-i}) = 0$ and $t_i'(\theta^m, \theta_{-i}) = t_i'(\theta^n, \theta_{-i}) = 0$. In this case, for (2), LHS $=$ RHS $= 0$ and thus $IC_{\theta^m \to \theta^n}$ is satisfied; on the other hand, for (3), LHS $=$ RHS $= 0$ and hence $IC_{\theta^n \to \theta^m}$ is satisfied;

2. if $\theta^m > c(N) - \sum_{j \neq i} \theta_j > \theta^n$, then $x^*(\theta^m, \theta_{-i}) = 1$ and $x^*(\theta^n, \theta_{-i}) = 0$. By definition, $t_i'(\theta^m, \theta_{-i}) \in (\theta^n, \theta^m]$ and $t_i'(\theta^n, \theta_{-i}) = 0$. In this case, for (2), LHS $\leqslant \theta^m =$ RHS and thus $IC_{\theta^m \to \theta^n}$ is satisfied; on the other hand, for (3), LHS $\geqslant \theta^n =$ RHS and thus $IC_{\theta^n \to \theta^m}$ is satisfied;

3. if $\theta^m > \theta^n > c(N) - \sum_{j \neq i} \theta_j$, then $x^*(\theta^m, \theta_{-i}) = x^*(\theta^n, \theta_{-i}) = 1$ and $t_i'(\theta^m, \theta_{-i}) = t_i'(\theta^n, \theta_{-i}) \geqslant 0$. In this case, for (2), LHS $=$ RHS $= 0$ and thus $IC_{\theta^m \to \theta^n}$ is satisfied; on the other hand, for (3), LHS $=$ RHS $= 0$ and hence $IC_{\theta^n \to \theta^m}$ is satisfied.

Therefore, the VCG' mechanism satisfies DSIC. ■

We also re-express the transfer rule of the VCG' mechanism.

Claim 2 *The transfer rule in the VCG' mechanism can be rewritten as follows: for each $i \in \mathcal{N}$, $\theta^m \in \Theta$, and $\theta_{-i} \in \Theta^{N-1}$,*

$$t_i'(\theta^m, \theta_{-i}) = \sum_{l=1}^{m} \theta^l \left(x^*(\theta^l, \theta_{-i}) - x^*(\theta^{l-1}, \theta_{-i}) \right), \tag{4}$$

where with abuse of notation, we slightly expand the domain of x^* by including $\theta^0 < 0$ as a normalization such that $x^*(\theta^0, \theta_{-i}) = 0$ for every $\theta_{-i} \in \Theta^{N-1}$.

Proof: Observe that for every $\theta^l \in \Theta$, $x^*(\theta^l, \theta_{-i})$ is either 0 or 1, and $x^*(\theta^l, \theta_{-i}) \geqslant x^*(\theta^{l-1}, \theta_{-i})$ because $\sum_{j \neq i} \theta_j + \theta^l > \sum_{j \neq i} \theta_j + \theta^{l-1}$. So, for each $i \in \mathcal{N}$, $\theta^m \in \Theta$, and $\theta_{-i} \in \Theta^{N-1}$,

1. if $x^*(\theta^m, \theta_{-i}) = 0$, then $t_i'(\theta^m, \theta_{-i}) = 0$ according to (1). On the other hand, in this case, $x^*(\theta^l, \theta_{-i}) = 0$ for each $l < m$, and the right-hand side of (4) is equal to $\sum_{l=1}^{m} \theta^l(0-0) = 0$, which is the same as $t_i'(\theta^m, \theta_{-i})$.

2. if $x^*(\theta^m, \theta_{-i}) = 1$, then there exists $\bar{l} \leqslant m$ such that $x^*(\theta^{\bar{l}}, \theta_{-i}) = 1$ and $x^*(\theta^{\bar{l}-1}, \theta_{-i}) = 0$, or equivalently, $\sum_{j \neq i} \theta_j + \theta^m \geqslant \sum_{j \neq i} \theta_j + \theta^{\bar{l}} \geqslant c(N) > \sum_{j \neq i} \theta_j + \theta^{\bar{l}-1}$. Hence, $t_i'(\theta^m, \theta_{-i}) = \theta^{\bar{l}}$ according to (1). On the other hand, the right-hand side of (4) becomes
$$\sum_{l=\bar{l}}^{m} \theta^l \left(x^*(\theta^l, \theta_{-i}) - x^*(\theta^{l-1}, \theta_{-i}) \right) = \theta^{\bar{l}}(1-0) + \sum_{l=\bar{l}+1}^{m} \theta^l(1-1) = \theta^{\bar{l}},$$
which is the same as $t_i'(\theta^m, \theta_{-i})$ in this case.

■

1.3.2 The Existence of Mechanisms satisfying BIC, IIR, EFF, and BB

We will show that the VCG' mechanism maximizes the ex ante budget surplus among all mechanisms satisfying BIC, IIR, and EFF. To establish this, we need to compute the ex ante budget surplus of any decision efficient mechanism (x^*, t). This result reduces our search for an appropriate mechanism to the class of the VCG' mechanisms.

To simplify the notation, for each agent $i \in \mathcal{N}$ in a mechanism (x, t), we denote by $\bar{x}_i(\theta_i)$ the expected probability that the public good is provided and by $\bar{t}_i(\theta_i)$ agent i's expected payment "when he is of type θ_i," respectively. That is,
$$\bar{x}_i(\theta_i) \equiv \sum_{\theta_{-i} \in \Theta^{N-1}} P^{N-1}(\theta_{-i}) x(\theta_i, \theta_{-i}),$$
and
$$\bar{t}_i(\theta_i) \equiv \sum_{\theta_{-i} \in \Theta^{N-1}} P^{N-1}(\theta_{-i}) t_i(\theta_i, \theta_{-i}).$$
In particular, for the VCG' mechanism, we have
$$\begin{aligned} \bar{x}_i^*(\theta_i) &= \sum_{\theta_{-i} \in \Theta^{N-1}} P^{N-1}(\theta_{-i}) x^*(\theta_i, \theta_{-i}); \\ \bar{t}_i'(\theta^m) &= \sum_{\theta_{-i} \in \Theta^{N-1}} P^{N-1}(\theta_{-i}) \sum_{l=1}^{m} \theta^l \left(x^*(\theta^l, \theta_{-i}) - x^*(\theta^{l-1}, \theta_{-i}) \right) = \sum_{l=1}^{m} \theta^l \left(\bar{x}_i^*(\theta^l) - \bar{x}_i^*(\theta^{l-1}) \right). \end{aligned}$$

Now, we can compute the ex ante budget surplus $\Pi_{ea}(x^*)$ of any decision efficient mechanism (x^*, t):

$$\begin{aligned}
\Pi_{ea}(x^*) &\equiv \sum_{\theta \in \Theta^N} P^N(\theta) \left(\sum_{i \in \mathcal{N}} t_i(\theta) - c(N) x^*(\theta) \right) \\
&= \sum_{i \in \mathcal{N}} \sum_{\theta \in \Theta^N} P^N(\theta) t_i(\theta) - c(N) \sum_{\theta \in \Theta^N} P^N(\theta) x^*(\theta) \\
&= \sum_{i \in \mathcal{N}} \sum_{\theta_i \in \Theta} P(\theta_i) \left(\sum_{\theta_{-i} \in \Theta^{N-1}} P^{N-1}(\theta_{-i}) t_i(\theta_i, \theta_{-i}) \right) \\
&\quad - c(N) \sum_{\theta_i \in \Theta} P(\theta_i) \left(\sum_{\theta_{-i} \in \Theta^{N-1}} P^{N-1}(\theta_{-i}) x^*(\theta) \right) \\
&= \sum_{i \in \mathcal{N}} \sum_{\theta_i \in \Theta} P(\theta_i) \bar{t}_i(\theta_i) - c(N) \sum_{\theta_i \in \Theta} P(\theta_i) \bar{x}_i^*(\theta_i).
\end{aligned}$$

So, $\bar{t}_i(\theta_i)$ for each agent i and $\theta_i \in \Theta$ must be as large as possible in order to achieve the maximum ex ante expected budget surplus. Our objective here is to find their maximum values among all mechanisms satisfying BIC and IIR.

We write down the lowest type's IIR and the downward adjacent BIC constraints for each agent i in any decision efficient mechanism (x^*, t):

$$\begin{aligned}
IR_{\theta^1}: &\quad \theta^1 \bar{x}_i^*(\theta^1) - \bar{t}_i(\theta^1) \geqslant 0; \\
IC_{\theta^m \to \theta^{m-1}}: &\quad \theta^m \bar{x}_i^*(\theta^m) - \bar{t}_i(\theta^m) \geqslant \theta^m \bar{x}_i^*(\theta^{m-1}) - \bar{t}_i(\theta^{m-1}).
\end{aligned}$$

We further rewrite the BIC constraints:

$$IC_{\theta^m \to \theta^{m-1}}: \quad \bar{t}_i(\theta^m) - \bar{t}_i(\theta^{m-1}) \leq \theta^m \left(\bar{x}_i^*(\theta^m) - \bar{x}_i^*(\theta^{m-1}) \right).$$

We are ready to state the key result of this paper.

Theorem 1 *The VCG' mechanism (x^*, t') maximizes the ex ante budget surplus among all mechanisms satisfying BIC, IIR, and EFF.*[10] *Moreover, the ex ante budget surplus of the VCG' mechanism $\Pi'_{ea}(x^*)$ is given as follows:*

$$\Pi'_{ea}(x^*) = \sum_{m=1}^{M-1} \bar{x}_i^*(\theta^m) \left(N\theta^m \sum_{l=m}^{M} P(\theta^l) - N\theta^{m+1} \sum_{l=m+1}^{M} P(\theta^l) - c(N) P(\theta^m) \right) + \bar{x}_i^*(\theta^M) \left(N\theta^M - c(N) \right) P(\theta^M). \tag{5}$$

[10] Since the VCG' mechanism satisfies DSIC and EPIR, then it also satisfies BIC and IIR.

Proof: From IR_{θ^1}, we know that for each agent $i \in \mathcal{N}$,

$$\bar{t}_i(\theta^1) \leqslant \theta^1 \bar{x}_i^*(\theta^1) = \bar{t}_i'(\theta^1).$$

If $\theta^m > \theta^1$, or equivalently, $m > 1$, then adding IR_{θ^1} and $IC_{\theta^l \to \theta^{l-1}}$ for every $2 \leqslant l \leqslant m$, we obtain that for each agent $i \in \mathcal{N}$,

$$\bar{t}_i(\theta^m) \leqslant \sum_{l=1}^{m} \theta^l \left(\bar{x}_i^*(\theta^l) - \bar{x}_i^*(\theta^{l-1})\right) = \bar{t}_i'(\theta^m).$$

Therefore, each type has the largest expected payment in the VCG' mechanism, and thus the VCG' mechanism maximizes the ex ante expected budget surplus. Thus, it only remains to compute Π_{ea}':

$$\begin{aligned}
\Pi_{ea}'(x^*) &= \sum_{i\in\mathcal{N}} \sum_{\theta_i\in\Theta} P(\theta_i)\bar{t}_i'(\theta_i) - c(N) \sum_{\theta_i\in\Theta} P(\theta_i)\bar{x}_i^*(\theta_i) \\
&= \sum_{i\in\mathcal{N}} \sum_{m=1}^{M} P(\theta^m)\bar{t}_i'(\theta^m) - c(N) \sum_{m=1}^{M} P(\theta^m)\bar{x}_i^*(\theta^m) \\
&= N \sum_{m=1}^{M} P(\theta^m) \sum_{l=1}^{m} \theta^l \left(\bar{x}_i^*(\theta^l) - \bar{x}_i^*(\theta^{l-1})\right) - c(N) \sum_{m=1}^{M} P(\theta^m)\bar{x}_i^*(\theta^m) \\
&\quad \big(\text{recall the definition of } \bar{t}_i'(\theta^m)\big) \\
&= N \sum_{m=1}^{M} \theta^m \left(\bar{x}_i^*(\theta^m) - \bar{x}_i^*(\theta^{m-1})\right) \sum_{l=m}^{M} P(\theta^l) - c(N) \sum_{m=1}^{M} P(\theta^m)\bar{x}_i^*(\theta^m) \\
&= \sum_{m=1}^{M-1} \bar{x}_i^*(\theta^m) \left(N\theta^m \sum_{l=m}^{M} P(\theta^l) - N\theta^{m+1} \sum_{l=m+1}^{M} P(\theta^l) - c(N)P(\theta^m)\right) \\
&\quad + \bar{x}_i^*(\theta^M) \left(N\theta^M - c(N)\right) P(\theta^M).
\end{aligned}$$

We thus obtain the desired expression for $\Pi_{ea}'(x^*)$ as in (5). ■

The main implication of Theorem 1 is that if there exists a mechanism satisfying BB together with BIC, IIR and EFF, then the VCG' mechanism must achieve a nonnegative ex ante budget surplus. On the other hand, if the VCG' mechanism generates a nonnegative ex ante budget surplus, we may redistribute the ex ante surplus in such a way that we can construct a mechanism that satisfies BB as well as BIC, IIR, and EFF. We formally state the result below.

Corollary 1 *There exists a mechanism* (x^*, t) *satisfying BIC, IIR, EFF, and BB if and only if* $\Pi'_{ea}(x^*) \geqslant 0$.[11]

Proof: We first prove the necessity of $\Pi'_{ea}(x^*) \geqslant 0$. Suppose that (x^*, t) satisfies BIC, IIR, EFF, and BB. Then (x^*, t) has zero ex ante expected budget surplus. By Theorem 1, we obtain $\Pi'_{ea}(x^*) \geqslant 0$.

We now prove the sufficiency. Consider the mechanism (x^*, t) where

$$\begin{aligned} t_1(\theta) &= (t'_1(\theta) - \Pi'_{ea}(x^*)) + \left(c(N)x^*(\theta) - \sum_{i\in\mathcal{N}} t'_i(\theta) + \Pi'_{ea}(x^*)\right) \\ &\quad - \left(c(N)\bar{x}^*_1(\theta_1) - \sum_{i\in\mathcal{N}} \bar{t}'_i(\theta_1) + \Pi'_{ea}(x^*)\right); \\ t_2(\theta) &= t'_2(\theta) + \left(c(N)\bar{x}^*_1(\theta_1) - \sum_{i\in\mathcal{N}} \bar{t}'_i(\theta_1) + \Pi'_{ea}(x^*)\right); \\ t_i(\theta) &= t'_i(\theta) \text{ for any } i \in \mathcal{N}\backslash\{1,2\}. \end{aligned}$$

Then, the ex post budget balance (BB) is satisfied because for all $\theta \in \Theta^N$,

$$\sum_{i=1}^{N} t_i(\theta) = \sum_{i=1}^{N} t'_i(\theta) - \Pi'_{ea}(x^*) + \left(c(N)x^*(\theta) - \sum_{i=1}^{N} t'_i(\theta) + \Pi'_{ea}(x^*)\right) = c(N)x^*(\theta).$$

Besides, the interim expected payment of each agent $i \in \mathcal{N}$ is obtained as follows.

1. For $i = 1$, $\bar{t}_1(\theta_1) = \bar{t}'_1(\theta_1) - \Pi'_{ea}(x^*) \leqslant \bar{t}'_1(\theta_1)$ because $\Pi'_{ea}(x^*) \geq 0$;

[11] This result is considered a discrete type space counterpart of Theorem 2 of Krishna and Perry (2000) in the context of the public good provision problem.

2. For $i = 2$,

$$\begin{aligned}
\bar{t}_2(\theta_2) &= \bar{t}'_2(\theta_2) + \sum_{\theta_{-2} \in \Theta^{N-1}} P^{N-1}(\theta_{-2}) \left(c(N)\bar{x}^*_1(\theta_1) - \sum_{i \in \mathcal{N}} \bar{t}'_i(\theta_1) + \Pi'_{ea}(x^*) \right) \\
&= \bar{t}'_2(\theta_2) + \sum_{\theta_1 \in \Theta} P(\theta_1) \left(c(N)\bar{x}^*_1(\theta_1) - \sum_{i \in \mathcal{N}} \bar{t}'_i(\theta_1) \right) + \Pi'_{ea}(x^*) \\
&\quad \left(\because \ c(N)\bar{x}^*_1(\theta_1) - \sum_{i \in \mathcal{N}} \bar{t}'_i(\theta_1) + \Pi'_{ea}(x^*) \ \text{ only depends on } \theta_1 \right) \\
&= \bar{t}'_2(\theta_2) + \sum_{\theta \in \Theta^N} P^N(\theta) \left(c(N)x^*(\theta) - \sum_{i \in \mathcal{N}} t'_i(\theta) \right) + \Pi'_{ea}(x^*) \\
&\quad (\because \ \text{types are independently distributed}) \\
&= \bar{t}'_2(\theta_2) - \Pi'_{ea}(x^*) + \Pi'_{ea}(x^*) = \bar{t}'_2(\theta_2);
\end{aligned}$$

3. For $i \in \mathcal{N} \backslash \{1, 2\}$, $\bar{t}_i(\theta_i) = \bar{t}'_i(\theta_i)$.

Hence, the interim expected transfers of all agents in the mechanism (x^*, t) are the same as those in the VCG' mechanism (x^*, t'), except agent 1. In particular, agent 1's interim expected transfer in mechanism (x^*, t) differs from that in (x^*, t') by a negative constant $-\Pi'_{ea}(x^*) \leqslant 0$. Therefore, (x^*, t) also satisfies BIC and IIR. ■

1.3.3 Efficient Mechanisms in Large Economies

Now, let us drop BB and investigate the implication of mechanisms satisfying BIC, IIR, and EFF in large economies. Let $x^*[N]$ denote the efficient decision rule in an economy with N agents. In the theorem below, we shall show that in any nontrivial case, the VCG' mechanism results in the per capita ex ante budget deficit in large economies. For this result, we assume $\lim_{N \to \infty} \sum_{\theta \in \Theta^N} P^N(\theta) x^*[N](\theta) > 0$, which means that the ex ante probability that the efficient decision rule provides the public good converges to some positive probability, as the population size N goes to infinity. This assumption strikes us as being innocuous because if it is not satisfied, the public good will not be expected to be provided in large economies.

Theorem 2 *Assume* $\lim_{N\to\infty}\sum_{\theta\in\Theta^N} P^N(\theta)x^*[N](\theta) > 0$. *Then, as the size of the economy gets large (i.e., $N \to \infty$), any mechanism satisfying BIC, IIR, and EFF results in the per capita ex ante budget deficit, i.e.,* $\lim_{N\to\infty}\Pi'_{ea}(x^*[N])/N < 0$.

Proof: We have assumed that $\theta^1 < \lim_{N\to\infty} c(N)/N \leq \theta^M$. We take the expression for $\Pi'_{ea}(x^*[N])$ from Theorem 1:

$$\Pi'_{ea}(x^*[N]) = \sum_{m=1}^{M-1} \bar{x}_i^*[N](\theta^m)\left(N\theta^m \sum_{l=m}^{M} P(\theta^l) - N\theta^{m+1}\sum_{l=m+1}^{M} P(\theta^l) - c(N)P(\theta^m)\right)$$
$$+ \bar{x}_i^*[N](\theta^M)\left(N\theta^M - c(N)\right)P(\theta^M),$$

where $\bar{x}_i^*[N](\theta^m) = \sum_{\theta_{-i}\in\Theta^{N-1}} P^{N-1}(\theta_{-i})x^*(\theta^m, \theta_{-i})$ for each $\theta^m \in \Theta$. We use the following lemma whose proof is in the Appendix.

Lemma 1 $\lim\limits_{N\to\infty} \bar{x}_i^*[N](\theta^m) = \lim\limits_{N\to\infty} \bar{x}_i^*[N](\theta^n)$ *for any* $\theta^m, \theta^n \in \Theta$ *and* $i \in \mathcal{N}$.

This lemma says that the probability that any agent can be pivotal is approximately zero in large economies. Therefore, for N large enough,

$$\begin{aligned}
\Pi'_{ea}(x^*[N]) &\approx \bar{x}_i^*[N](\theta^M)\left\{\sum_{m=1}^{M-1}\left(N\theta^m\sum_{l=m}^{M}P(\theta^l) - N\theta^{m+1}\sum_{l=m+1}^{M}P(\theta^l)\right) + N\theta^M P(\theta^M) - c(N)\right\} \\
&\approx \bar{x}_i^*[N](\theta^M)\left\{\sum_{m=2}^{M} N\theta^m\left(\sum_{l=m}^{M}P(\theta^l) - \sum_{l=m}^{M}P(\theta^l)\right) + N\theta^1\sum_{l=1}^{M}P(\theta^l) - c(N)\right\} \\
&\approx \bar{x}_i^*[N](\theta^M)\left(N\theta^1 - c(N)\right).
\end{aligned}$$

We next use the following lemma whose proof is also in the Appendix.

Lemma 2 $\lim\limits_{N\to\infty}\sum\limits_{\theta\in\Theta^N} P^N(\theta)x^*[N](\theta) = \lim\limits_{N\to\infty}\bar{x}_i^*[N](\theta^M)$.

This lemma says that the ex ante expected probability of public good provision is approximately the same as the interim expected probability of public good provision in large economies. Hence,

$$\begin{aligned}
\lim_{N\to\infty}\frac{\Pi'_{ea}(x^*[N])}{N} &= \lim_{N\to\infty}\bar{x}_i^*[N](\theta^M)\left(\theta^1 - \frac{c(N)}{N}\right) \\
&= \lim_{N\to\infty}\left(\sum_{\theta\in\Theta^N}P^N(\theta)x^*[N](\theta)\right)\left(\theta^1 - \frac{c(N)}{N}\right) < 0,
\end{aligned}$$

because $\lim_{N\to\infty}\sum_{\theta\in\Theta^N} P^N(\theta)x^*[N](\theta) > 0$ and $\theta^1 < \lim_{N\to\infty} c(N)/N$. This completes the proof of Theorem 2. ■

So, as the economy gets large, if the ex ante probability that the efficient decision rule x^* provides the public good is strictly positive, the VCG' mechanism results in the per capita ex ante budget deficit. Combining Theorem 2 and Corollary 1, we conclude that there are no mechanisms satisfying BIC, IIR, EFF, and BB in large economies. We formally state this result below.

Corollary 2 *Assume* $\lim_{N\to\infty}\sum_{\theta\in\Theta^N} P^N(\theta)x^*[N](\theta) > 0$. *Then, as the size of the economy gets large (i.e.,* $N\to\infty$*), there are no mechanisms satisfing BIC, IIR, EFF, and BB.*

Therefore, the existence of positive results in large economies, if any, only lies in the trivial cases: (i) $\lim_{N\to\infty}\sum_{\theta\in\Theta^N} P^N(\theta)x^*[N](\theta) = 0$, i.e., the ex ante expected probability that the efficient decision rule x^* provides the public good converges to zero as the size of the economy gets large; or (ii) $\theta^1 \geq c(N)/N$, i.e., the probability that providing the public good is efficient is one.

Proposition 1 *Suppose that either (i)* $\lim_{N\to\infty}\sum_{\theta\in\Theta^N} P^N(\theta)x^*[N](\theta) = 0$[12] *or (ii)* $\theta^1 \geq c(N)/N$ *holds. Then, as the size of the economy gets large (i.e.,* $N\to\infty$*), there exists a mechanism satisfying BIC, IIR, EFF, and BB in both cases.*

Proof: Recall that in the proof of Theorem 2, we obtain

$$\lim_{N\to\infty}\frac{\Pi'_{ea}(x^*[N])}{N} = \lim_{N\to\infty}\left(\sum_{\theta\in\Theta^N} P^N(\theta)x^*[N](\theta)\right)\left(\theta^1 - \frac{c(N)}{N}\right).$$

Suppose that (i) $\lim_{N\to\infty}\sum_{\theta\in\Theta^N} P^N(\theta)x^*[N](\theta) = 0$. Then, $\lim_{N\to\infty}\Pi'_{ea}(x^*[N])/N = 0$ and by Corollary 1, there exists a mechanism (x^*, t) satisfying BIC, IIR, EFF, and BB.

[12]In this case, as the size of the economy gets large (i.e., $N\to\infty$), the ex ante probability that the public good is provided converges to zero. This is consistent with the analysis of Section 4.2 of Ledyard and Palfrey (1994).

Suppose that (ii) $\theta^1 \geq c(N)/N$. Then, $\lim_{N\to\infty} \Pi'_{ea}(x^*[N])/N \geq 0$ and by Corollary 1, there exists a mechanism (x^*, t) satisfying BIC, IIR, EFF, and BB. For example, the following mechanism (x^*, t) where for any $\theta \in \Theta^N$,

$$x^*(\theta) = 1,$$

and

$$t_i(\theta) = c(N)/N \text{ for every } i \in \mathcal{N},$$

satisfies BIC, IIR, EFF, and BB. ■

1.3.4 Simulation Results

In the previous subsection, we have shown that when N is sufficiently large, there exists no mechanism satisfying BIC, IIR, EFF, and BB. Now we will run simulations to find out when a mechanism satisfying BIC, IIR, EFF, and BB starts to disappear. For the sake of simplicity, we confine ourselves to a two type environment in this subsection. That is, $\Theta = \{\theta^1, \theta^2\}$. Assume $P(\theta^1) = \varepsilon \in (0, 1)$ and $\theta^1 < c(N)/N \leq \theta^2$.

In this case, there must exist $k \in (0, N]$ such that $c(N) = k\theta^2 + (N-k)\theta^1$ is satisfied for any $c(N) \in (N\theta^1, N\theta^2]$. We obtain $k = (c(N) - N\theta^1)/(\theta^2 - \theta^1)$. Note that k need not be an integer and $k/N = (c(N)/N - \theta^1)/(\theta^2 - \theta^1) > 0$ because $c(N)/N > \theta^1$. Hence, k is an increasing function of N and in what follows, we denote k by $k(N)$.

Define $k^*(N) \equiv \lceil k(N) \rceil$, the least integer greater than or equal to $k(N)$ for every N. Then we have

$$(k^*(N) - 1)\theta^2 + (N - k^*(N) + 1)\theta^1 < c(N) \leq k^*(N)\theta^2 + (N - k^*(N))\theta^1$$

for every N and the efficient decision rule x^* can be rewritten as follows:

$$x(\theta) = \begin{cases} 1 & \text{if at least } k^*(N) \text{ agents are of type } \theta^2 \\ 0 & \text{otherwise.} \end{cases}$$

The interim expected probability of public good provision is computed below: for each agent $i \in \mathcal{N}$,

$$\bar{x}_i^*(\theta^1) \;=\; P^{N-1}(\theta_{-i})x^*(\theta^1, \theta_{-i}) = \sum_{k=k^*(N)}^{N-1} p^{N-1}(k);$$
$$\bar{x}_i^*(\theta^2) \;=\; P^{N-1}(\theta_{-i})x^*(\theta^2, \theta_{-i}) = \sum_{k=k^*(N)-1}^{N-1} p^{N-1}(k),$$

where $p^{N-1}(k)$ is the probability that k out of $(N-1)$ agents are of type θ^2, i.e.,

$$p^{N-1}(k) \equiv \varepsilon^{N-1-k} \cdot (1-\varepsilon)^k \begin{pmatrix} N-1 \\ k \end{pmatrix}.$$

We show the following property:

Lemma 3

$$\Pi'_{ea}(x^*) < (\theta^2 - \theta^1)\bar{\Pi}_{ea}(x^*)$$

where

$$\bar{\Pi}_{ea}(x^*) \equiv -(k^*(N)-1) \sum_{k=k^*(N)}^{N-1} p^{N-1}(k) + (1-\varepsilon)(N - k^*(N)+1)p^{N-1}(k^*(N)-1).$$

Proof: The proof is relegated to the Appendix. ■

Therefore, if $\bar{\Pi}_{ea}(x^*) > 0$, then $\Pi'_{ea}(x^*) \geqslant 0$ can be satisfied and by Corollary 1, we are able to construct a transfer rule t such that the direct mechanism (x^*, t) satisfies BIC, IIR, EFF, and BB.

In what follows, we run simulations to find out when $\bar{\Pi}_{ea}(x^*) > 0$ is no longer satisfied. We first fix $k(N) = (1-\varepsilon)N$ and $k^*(N) = \lceil (1-\varepsilon)N \rceil$.[13] Recall that $(1-\varepsilon)$ is the probability that a given agent is of type θ^2. So, if θ^2 is more likely to occur than θ^1 for each agent, i.e., $1-\varepsilon > 1/2$, then it is efficient to provide the public good if more than $N/2$ agents are of type θ^2.

There are two inputs in the simulation: $N \geqslant 2$ and $\varepsilon \in (0,1)$. The output is $\bar{\Pi}_{ea}(x^*)$. The simulation results are shown in Table 1 below and the positive outputs are highlighted in red color.

[13]In this case, the VCG' mechanism results in per capita ex ante budget deficit as the size of economy gets large. The details of this assertion are in the Appendix.

We are particularly interested in finding the cutoff N_0 such that $\bar{\Pi}_{ea}(x^*) > 0$ for all $N \leqslant N_0$ and $\bar{\Pi}_{ea}(x^*) < 0$ for all $N > N_0$. For example, we consider when $\varepsilon = 0.5$, $N_0 = 4$. This implies that no mechanisms satisfy BIC, IIR, EFF, and BB in an economy with more than four agents. This is an extremely small economy. However, N_0 reaches its local maximum at the two extremes of ε where either $\varepsilon \to 0$ or $\varepsilon \to 1$. For example, when $\varepsilon = 0.9$, $N_0 = 18$. Nevertheless, these are still relatively small economies. Therefore, we conclude that the positive results exist only in relatively small economies and the negative results in large economies are quite prevalent.

Table 1: Simulation Results for $\bar{\Pi}_{ea}(x^*)$

	$\varepsilon = 0.1$	$\varepsilon = 0.2$	$\varepsilon = 0.3$	$\varepsilon = 0.4$	$\varepsilon = 0.5$	$\varepsilon = 0.6$	$\varepsilon = 0.7$	$\varepsilon = 0.8$	$\varepsilon = 0.9$
$N = 2$	0.810	0.640	0.490	0.360	0.500	0.480	0.420	0.320	0.180
$N = 3$	0.729	0.512	0.343	0.216	0.250	0.224	0.441	0.384	0.243
$N = 4$	0.656	0.410	-0.069	0.086	0.063	0.166	0.181	0.410	0.292
$N = 5$	0.590	-0.573	-0.144	-0.328	-0.063	0.028	0.146	0.410	0.328
$N = 6$	0.531	-0.655	-0.168	-0.389	-0.375	-0.082	0.068	0.147	0.354
$N = 7$	0.478	-0.682	-1.000	-0.373	-0.406	-0.289	-0.025	0.127	0.372
$N = 8$	0.430	-0.671	-0.980	-0.983	-0.816	-0.289	-0.134	0.091	0.383
$N = 9$	0.387	-0.638	-0.909	-0.908	-0.770	-0.549	-0.274	0.040	0.387
$N = 10$	-2.402	-2.329	-2.030	-1.661	-1.262	-0.850	-0.251	-0.020	0.387
$N = 11$	-2.441	-2.282	-1.932	-1.541	-1.146	-0.765	-0.411	-0.101	0.124
$N = 12$	-2.448	-2.190	-1.783	-1.365	-1.710	-1.112	-0.598	-0.175	0.119
$N = 13$	-2.429	-2.068	-1.604	-2.250	-1.534	-0.947	-0.803	-0.260	0.111
$N = 14$	-2.389	-1.927	-2.967	-2.027	-2.162	-1.331	-0.682	-0.353	0.099
$N = 15$	-2.333	-4.128	-2.750	-3.019	-1.929	-1.744	-0.907	-0.454	0.083
$N = 16$	-2.265	-3.976	-2.500	-2.761	-2.616	-1.515	-1.151	-0.405	0.064
$N = 17$	-2.187	-3.786	-4.067	-2.455	-2.331	-1.964	-0.946	-0.516	0.042
$N = 18$	-2.101	-3.569	-3.790	-3.554	-3.073	-1.669	-1.204	-0.636	0.017
$N = 19$	-2.011	-3.335	-3.477	-3.209	-2.738	-2.148	-1.481	-0.762	-0.009
$N = 20$	-6.375	-5.954	-5.221	-4.395	-3.531	-2.651	-1.202	-0.893	-0.038
$N = 21$	-6.281	-5.710	-4.890	-4.020	-3.150	-2.301	-1.490	-0.740	-0.104
$N = 22$	-6.164	-5.434	-4.520	-3.603	-3.991	-2.835	-1.795	-0.880	-0.136
$N = 23$	-6.026	-5.133	-4.129	-4.881	-3.566	-2.427	-2.113	-1.027	-0.170
$N = 24$	-5.870	-4.817	-6.041	-4.430	-4.452	-2.987	-1.764	-1.180	-0.207
$N = 25$	-5.701	-7.797	-5.620	-5.783	-3.985	-3.567	-2.096	-1.337	-0.246

1.4 When the First Best is Not Implementable

Thus far, we conclude, especially by Corollary 2, that there are no mechanisms satisfying BIC, IIR, EFF, and BB in large economies. The simulation results in the previous section even suggest that there are no mechanisms satisfying BIC, IIR, EFF, and BB in a relatively small economy. In this section, we drop decision efficiency (EFF) from the requirements and mainly consider the case where $\Pi'_{ea}(x^*) < 0$, which together with our Corollary 1 implies that no mechanism (x^*, t) satisfies BIC, IIR, EFF, and BB.

1.4.1 Preliminaries

First, we define the expected social welfare of any mechanism (x, t):

$$SW(x) = \sum_{\theta \in \Theta^N} P^N(\theta) x(\theta) \left(\sum_{i \in \mathcal{N}} \theta_i - c(N) \right).$$

Clearly, the decision efficient rule x^* maximizes the expected social welfare. However, in this section, we include the case that no mechanism (x^*, t) satisfies BIC, IIR and BB.

Recall that $\bar{x}_i(\theta_i)$ is the interim expected probability of public good provision and that $\bar{t}_i(\theta_i)$ is the expected payment of agent i of type θ_i. We characterize the mechanisms satisfying BIC below. We say that a decision rule x is *implementable* if there exists a transfer rule $t : \Theta^N \to \mathbb{R}^N$ such that the mechanism (x, t) satisfies BIC. We first characterize the implementability in terms of monotonicity of a decision rule.

Lemma 4 *A decision rule x is implementable if and only if, for each $i \in \mathcal{N}$ and $\theta^m, \theta^n \in \Theta$ with $m > n$, $\bar{x}_i(\theta^m) \geq \bar{x}_i(\theta^n)$.*

Proof: Now we can write down the BIC constraints for any mechanism (x, t): for any $m \neq n$,

$$IC_{\theta^m \to \theta^n} : \quad \theta^m \bar{x}_i(\theta^m) - \bar{t}_i(\theta^m) \geqslant \theta^m \bar{x}_i(\theta^n) - \bar{t}_i(\theta^n);$$

$$IC_{\theta^n \to \theta^m} : \quad \theta^n \bar{x}_i(\theta^n) - \bar{t}_i(\theta^n) \geqslant \theta^n \bar{x}_i(\theta^m) - \bar{t}_i(\theta^m).$$

We further rewrite the BIC constraints: for any $m \neq n$,

$$IC_{\theta^m \to \theta^n} : \quad \bar{t}_i(\theta^m) - \bar{t}_i(\theta^n) \leq \theta^m \left(\bar{x}_i(\theta^m) - \bar{x}_i(\theta^n) \right); \tag{6}$$

$$IC_{\theta^n \to \theta^m} : \quad \bar{t}_i(\theta^m) - \bar{t}_i(\theta^n) \geq \theta^n \left(\bar{x}_i(\theta^m) - \bar{x}_i(\theta^n) \right). \tag{7}$$

Note that we say that a decision rule x is implementable if there exists a transfer rule t such that (x,t) satisfies BIC. So, from (6) and (7), we know that x is implementable if and only if $(\theta^m - \theta^n)(\bar{x}_i(\theta^m) - \bar{x}_i(\theta^n)) \geqslant 0$, i.e., $\bar{x}_i(\theta^m) \geqslant \bar{x}_i(\theta^n)$ for any $m > n$. ■

Second, we compute the ex ante budget surplus for the mechanism (x,t):

$$\begin{aligned}
\Pi_{ea}(x^*) &\equiv \sum_{\theta\in\Theta^N} P^N(\theta)\left(\sum_{i\in\mathcal{N}} t_i(\theta) - c(N)x(\theta)\right) \\
&= \sum_{i\in\mathcal{N}}\sum_{\theta\in\Theta^N} P^N(\theta)t_i(\theta) - c(N)\sum_{\theta\in\Theta^N} P^N(\theta)x(\theta) \\
&= \sum_{i\in\mathcal{N}}\sum_{\theta_i\in\Theta} P(\theta_i)\left(\sum_{\theta_{-i}\in\Theta^{N-1}} P^{N-1}(\theta_{-i})t_i(\theta_i,\theta_{-i})\right) \\
&\quad -c(N)\sum_{\theta_i\in\Theta} P(\theta_i)\left(\sum_{\theta_{-i}\in\Theta^{N-1}} P^{N-1}(\theta_{-i})x(\theta)\right) \\
&= \sum_{i\in\mathcal{N}}\sum_{\theta_i\in\Theta} P(\theta_i)\bar{t}_i(\theta_i) - c(N)\sum_{\theta_i\in\Theta} P(\theta_i)\bar{x}_i(\theta_i).
\end{aligned}$$

So, $\bar{t}_i(\theta_i)$ for each agent i and $\theta_i \in \Theta$ must be as large as possible in order to achieve the maximum ex ante expected budget surplus.

Finally, let us introduce the *tight mechanism* (x, t^T), which was originally proposed by Kos and Manea (2009) in a bilateral trade environment. We adapt this to our public good environment.

Definition 8 *A mechanism (x, t^T) is called the tight mechanism if for each agent $i \in \mathcal{N}$, $\theta^m \in \Theta$ and $\theta_{-i} \in \Theta^{N-1}$,*

$$t_i^T(\theta^m, \theta_{-i}) = \sum_{l=1}^{m} \theta^l \left(x(\theta^l, \theta_{-i}) - x(\theta^{l-1}, \theta_{-i})\right),$$

where with abuse of notation, we slightly expand the domain of x by including $\theta^0 < 0$ as a normalization such that $x(\theta^0, \theta_{-i}) \equiv 0$ for every $\theta_{-i} \in \Theta^{N-1}$.

Remark: Note that if $x = x^*$, then the tight mechanism indeed reduces to the VCG' mechanism we discussed in the previous section.

In the tight mechanism, an agent's payment is equal to his marginal contribution to the public good. The interim expected payment for agent i of type θ^m in the tight mechanism is

$$\bar{t}^T(\theta^m) = \sum_{\theta_{-i}\in\Theta^{N-1}} P^{N-1}(\theta_{-i}) \sum_{l=1}^{m} \theta^l \left(x(\theta^l,\theta_{-i}) - x(\theta^{l-1},\theta_{-i})\right) = \sum_{l=1}^{m} \theta^l \left(\bar{x}_i(\theta^l) - \bar{x}_i(\theta^{l-1})\right),$$

where $\bar{x}_i(\theta_i) \equiv \sum_{\theta_{-i}\in\Theta^{N-1}} P^{N-1}(\theta_{-i})x(\theta_i,\theta_{-i})$ for every $\theta_i \in \Theta$. We shall show that the tight mechanism satisfies BIC and IIR.

Claim 3 *The tight mechanism satisfies BIC and IIR.*

Proof: Fix $\theta^m, \theta^n \in \Theta$ such that $\theta^m > \theta^n$ arbitrarily. Recall that the BIC constraints for a mechanism (x,t) are given as follows:

$$IC_{\theta^m\to\theta^n}: \quad \bar{t}_i(\theta^m) - \bar{t}_i(\theta^n) \le \theta^m \left(\bar{x}_i(\theta^m) - \bar{x}_i(\theta^n)\right);$$
$$IC_{\theta^n\to\theta^m}: \quad \bar{t}_i(\theta^m) - \bar{t}_i(\theta^n) \ge \theta^n \left(\bar{x}_i(\theta^m) - \bar{x}_i(\theta^n)\right).$$

Note that in the tight mechanism (x, t^T),

$$\begin{aligned}
\bar{t}_i^T(\theta^m) - \bar{t}_i^T(\theta^n) &= \sum_{l=1}^{m} \theta^l \left(\bar{x}_i(\theta^l) - \bar{x}_i(\theta^{l-1})\right) - \sum_{l=1}^{n} \theta^l \left(\bar{x}_i(\theta^l) - \bar{x}_i(\theta^{l-1})\right) \\
&= \sum_{l=n+1}^{m} \theta^l \left(\bar{x}_i(\theta^l) - \bar{x}_i(\theta^{l-1})\right) \\
&\le \sum_{l=n+1}^{m} \theta^m \left(\bar{x}_i(\theta^l) - \bar{x}_i(\theta^{l-1})\right) \quad (\because \theta^l \le \theta^m \text{ for any } l \le m) \\
&= \theta^m \sum_{l=n+1}^{m} \left(\bar{x}_i(\theta^l) - \bar{x}_i(\theta^{l-1})\right) \\
&= \theta^m \left(\bar{x}_i(\theta^m) - \bar{x}_i(\theta^n)\right).
\end{aligned}$$

Hence, the tight mechanism (x, t^T) satisfies $IC_{\theta^m\to\theta^n}$. In particular, when $n = m-1$,

$$\bar{t}_i^T(\theta^m) - \bar{t}_i^T(\theta^{m-1}) = \theta^m \left(\bar{x}_i(\theta^m) - \bar{x}_i(\theta^{m-1})\right).$$

In other words, each agent's downward adjacent BIC constraint is binding. Also,

$$\begin{aligned}
\bar{t}_i^T(\theta^m) - \bar{t}_i^T(\theta^n) &= \sum_{l=n+1}^{m} \theta^l \left(\bar{x}_i(\theta^l) - \bar{x}_i(\theta^{l-1})\right) \\
&> \sum_{l=n+1}^{m} \theta^n \left(\bar{x}_i(\theta^l) - \bar{x}_i(\theta^{l-1})\right) \quad (\because \theta^l > \theta^n \text{ for any } l > n) \\
&= \theta^n \sum_{l=n+1}^{m} \left(\bar{x}_i(\theta^l) - \bar{x}_i(\theta^{l-1})\right) \\
&= \theta^n \left(\bar{x}_i(\theta^m) - \bar{x}_i(\theta^n)\right).
\end{aligned}$$

Hence, the tight mechanism (x, t^T) satisfies $IC_{\theta^n \to \theta^m}$.

Besides, we can write down the IIR constraints of any mechanism (x, t) as follows: for any $\theta^m \in \Theta$,

$$IR_{\theta^m} : \theta^m \bar{x}_i(\theta^m) - \bar{t}_i(\theta^m) \geq 0 \Rightarrow \bar{t}_i(\theta^m) \leq \theta^m \bar{x}_i(\theta^m).$$

In the tight mechanism (x, t^T), if $m > 1$, we have

$$\begin{aligned} \bar{t}_i^T(\theta^m) &= \sum_{l=1}^{m} \theta^l \left(\bar{x}_i(\theta^l) - \bar{x}_i(\theta^{l-1})\right) \\ &= \theta^m \bar{x}_i(\theta^m) + \sum_{l=1}^{m-1} (\theta^l - \theta^{l+1}) \bar{x}_i(\theta^l) \\ &< \theta^m \bar{x}_i(\theta^m). \ (\because \theta^l < \theta^{l+1} \text{ for any } l) \end{aligned}$$

On the other hand, if $m = 1$,

$$\bar{t}_i^T(\theta^m) = \bar{t}_i^T(\theta^1) = \theta^1 \left(\bar{x}_i(\theta^1) - \bar{x}_i(\theta^0)\right) = \theta^1 \bar{x}_i(\theta^1).$$

In other words, each agent's lowest type's IIR constraint is binding. Therefore, the tight mechanism (x, t^T) satisfies IIR. ■

Therefore, the tight mechanism (x, t^T) has the following property: each agent's downward adjacent incentive compatibility constraints as well as his lowest type's individual rationality constraints are binding.

1.4.2 Characterizations of Mechanisms Satisfying BIC, IIR, and BB

Recall that a decision rule x is implementable if there exists a transfer rule t such that the mechanism (x, t) satisfies BIC. We now establish the following theorem. This reduces our search for mechanisms to the class of the tight mechanisms.

Theorem 3 *Let x be an implementable decision rule. Then, the tight mechanism (x, t^T) maximizes the ex ante budget surplus among all mechanisms satisfying BIC and IIR. Moreover, we obtain the ex ante budget surplus of the tight mechanism as follows:*

$$\begin{aligned} \Pi_{ea}^T(x) = \sum_{m=1}^{M-1} \bar{x}_i(\theta^m) \left(N\theta^m \sum_{l=m}^{M} P(\theta^l) - N\theta^{m+1} \sum_{l=m+1}^{M} P(\theta^l) - c(N)P(\theta^m) \right) \\ + \bar{x}_i(\theta^M) \left(N\theta^M - c(N)\right) P(\theta^M). \end{aligned} \quad (8)$$

Proof: The proof is completed verbatim in the proof of Theorem 1, except that x^* is replaced by x. The complete proof is in the Appendix. ■

The following corollary is a generalization of Corollary 1, which includes the case that no mechanism (x^*, t) satisfies BIC, IIR and BB.

Corollary 3 *Let x be an implementable decision rule. Then, there exists a transfer rule $t : \Theta^N \to \mathbb{R}^N$ such that the mechanism (x, t) satisfies BIC, IIR, and BB if and only if $\Pi^T_{ea}(x) \geqslant 0$.*

Proof: The proof is completed verbatim in the proof of Corollary 1, except that x^* is replaced by x and the VCG' mechanism is replaced by the tight mechanism. ■

1.4.3 Mechanisms Satisfying BIC, IIR, and BB in Large Economies

We will investigate the class of mechanisms satisfying BIC, IIR, and BB in large economies. To simplify our argument, we restrict ourselves to the two type space in this subsection. That is, $\Theta = \{\theta^1, \theta^2\}$. After imposing the anonymity assumption which will be discussed later, we can partition all the possible type profiles based on the number of agents with type θ^2. This is certainly for its tractability. If each agent has more than two possible types, however, we no longer have such a nice book-keeping property so that we face a technical difficulty to keep track of all the possible type profiles. Our rationale for focusing on a two-type setup is reflected in the following consideration: (i) the negative result is obtained in the continuum type space; and (ii) moreover, by continuity, there is likely to be a similar negative result on a large finite type space, which is considered an approximation of the continuum type space.

Assume $\theta^1 < \lim_{N\to\infty} c(N)/N \leq \theta^2$ and $P(\theta^1) = \varepsilon$ where $0 < \varepsilon < 1$. Then there must exist $k(N)$ which is an increasing function of N such that $c(N) = k(N)\theta^2 + (N - k(N))\theta^1$ is satisfied. We obtain

$$\frac{k(N)}{N} = \frac{c(N)/N - \theta^1}{\theta^2 - \theta^1}.$$

Since $\theta^1 < \lim_{N\to\infty} c(N)/N \leq \theta^2$, $0 < \lim_{N\to\infty} k(N)/N \leq 1$ must be satisfied. Define $k^*(N) \equiv \lceil k(N) \rceil$, the least integer greater than or equal to $k(N)$ for every N.

Denote by $k(\theta)$ the number of θ^2-type agents in profile $\theta \in \Theta^N$. In particular, if $k(\theta) < k^*(N)$, then $\sum_{i=1}^N \theta_i < c(N)$. We impose the following assumption:

Assumption 1 *Every decision rule x satisfies the following property: $k(\theta) < k^*(N) \Rightarrow x(\theta) = 0$.*

In other words, the public good shall not be provided whenever the total surplus generated by the public good is lower than the cost of providing it. We consider this as a mild assumption because it is satisfied in a welfare-maximizing mechanism which is indeed considered by Theorem 2 of Mailath and Postlewaite (1990). We also impose the following anonymity condition on decision rules.

Assumption 2 *Every decision rule x satisfies the following property: for two profiles $\theta, \theta' \in \Theta^N$, $x(\theta) = x(\theta')$ whenever $k(\theta) = k(\theta')$.*

In other words, the probability that the public good is provided must be the same in any two type profiles if they have the same number of θ^2-type agents as well as the same number of θ^1-type agents. This strikes us as being quite natural in large economies.

By Assumption 2, we can rewrite the decision rule x as $x(k)$, a function of the number of θ^2-type agents which we denote by k. Let $p^N(k)$ be the probability that k out of N agents are of type θ^2. We are now ready to state the result.

Theorem 4 *Suppose that there are only two types, $\Theta = \{\theta^1, \theta^2\}$ and Assumption 1 and 2 hold. Let $\{x[N]\}_N$ be a sequence of decision rules such that for each population size N, there exists a transfer rule t for which the mechanism $(x[N], t)$ satisfies BIC, IIR, and BB in the N-agent economy.*[14] *Then, as the size of the economy gets large (i.e., $N \to \infty$), the ex ante probability that the public good is provided converges to zero, i.e., $\sum_{k=1}^N p^N(k)x(k) \to 0$.*[15]

[14]The existence of such a sequence is automatically guaranteed because we consider the mechanism (x, t) such that the public good never be provided and no transfers are made. Such a mechanism trivially satisfies BIC, IIR, and BB and it works for any number of agents.

[15]Theorem 4 is considered a discrete type space counterpart of Theorem 2 of Mailath and Postlewaite (1990).

Proof: The proof is in the Appendix. ■

The implication of this result is stronger than that of Corollary 2 in the sense that even if we drop EFF and consider mechanisms satisfying BIC, IIR, and BB, we end up with a negative result in large economies. This leads us to the need of dropping IIR so that we seek for more positive results in large economies. However, we leave this for future work.

1.5 Dominant Strategy Incentive Compatibility and Ex Post Individual Rationality

Now, let us replace the Bayesian incentive compatibility (BIC) and interim individual rationality (IIR) constraints with dominant strategy incentive compatibility (DSIC) and ex post individual rationality (EPIR) constraints, respectively. We then investigate the existence of mechanisms satisfying DSIC, EPIR, EFF, and BB. In this case, we could make any distributional assumption about the type space and in particular, we could allow for any correlation among types.

Suppose there are $N \geqslant 2$ agents in the economy and each agent has $M \geqslant 2$ types $0 \leq \theta^1 < \cdots < \theta^M$. We also assume that the cost of providing the public good $c(N)$ is an increasing function in N and $N\theta^1 < c(N) \leqslant (N-1)\theta^M$. Recall the decision efficient rule $x^*(\cdot)$: for any $\theta \in \Theta^N$,

$$x^*(\theta) = \begin{cases} 1 & \text{if } \sum_{i \in \mathcal{N}} \theta_i \geq c(N) \\ 0 & \text{otherwise.} \end{cases}$$

In order to satisfy BB, we assume that for every type profile $\theta \in \Theta^N$ and decision rule x,

$$t_N(\theta) = x(\theta) \cdot c(N) - \sum_{i \neq N} t_i(\theta).$$

Proposition 2 *Assume that $N\theta^1 < c(N) \leqslant (N-1)\theta^M$. Then, there are no mechanisms satisfying DSIC, EPIR, EFF, and BB.*

Remark: This is considered as a dominant strategy counterpart of our Corollary 2, which shows that there are no mechanisms satisfying BIC, IIR, EFF, and BB in large economies. Although DSIC and EPIR each are stronger than BIC and IIR, respectively, we establish here the negative result even if we make no assumption over types and do not appeal to the power of large economies. Theorem 7 of Green and Laffont (1977) shows a similar negative result in a rich environment with a continuum type space where the agents might have non-quasilinear preferences.

Proof: Consider the profile $\bar{\theta}$ where $\bar{\theta}_i = \theta^M$ for every $i \in \mathcal{N}$. Since $c(N) \leqslant (N-1)\theta^M$, then $x^*(\bar{\theta}) = x^*(\theta^1, \bar{\theta}_{-i}) = 1$. By DSIC, for each agent $i \in \mathcal{N}$, we have

$$\theta^M \cdot 1 - t_i(\bar{\theta}) \geqslant \theta^M \cdot 1 - t_i(\theta^1, \bar{\theta}_{-i}) \Rightarrow t_i(\bar{\theta}) \leqslant t_i(\theta^1, \bar{\theta}_{-i}),$$

and

$$\theta^1 \cdot 1 - t_i(\theta^1, \bar{\theta}_{-i}) \geqslant \theta^1 \cdot 1 - t_i(\bar{\theta}) \Rightarrow t_i(\theta^1, \bar{\theta}_{-i}) \leqslant t_i(\bar{\theta}).$$

Combining the above two inequalities, we obtain that for each agent $i \in \mathcal{N}$, $t_i(\bar{\theta}) = t_i(\theta^1, \bar{\theta}_{-i})$. In particular, for agent N, we have

$$\begin{aligned}
t_N(\bar{\theta}) =& t_N(\theta^1, \bar{\theta}_{-N}) \\
\Rightarrow c(N) - \sum_{i \neq N} t_i(\bar{\theta}) =& c(N) - \sum_{i \neq N} t_i(\theta^1, \bar{\theta}_{-N}) \ (\because BB) \\
\Rightarrow \sum_{i \neq N} t_i(\bar{\theta}) =& \sum_{i \neq N} t_i(\theta^1, \bar{\theta}_{-N}) \\
\Rightarrow \sum_{i \neq N} t_i(\theta^1, \bar{\theta}_{-i}) =& \sum_{i \neq N} t_i(\theta^1, \bar{\theta}_{-N}). \ (\because t_i(\bar{\theta}) = t_i(\theta^1, \bar{\theta}_{-i}) \text{ for each agent } i.)
\end{aligned} \tag{9}$$

Note that on the left-hand side of (9), $t_i(\theta^1, \bar{\theta}_{-i})$ denotes the payment of agent $i \neq N$ if he unilaterally deviates from the profile $\bar{\theta}$ and reports θ^1; on the right-hand side of (9), $t_i(\theta^1, \bar{\theta}_{-N})$ denotes the payment of agent $i \neq N$ if agent N unilaterally deviates from the profile $\bar{\theta}$ and reports θ^1.

For each agent $i \in \mathcal{N}$ under the profile $(\theta^1, \bar{\theta}_{-i})$, EPIR requires $\theta^1 \cdot 1 - t_i(\theta^1, \bar{\theta}_{-i}) \geqslant 0$, which is equivalent to $t_i(\theta^1, \bar{\theta}_{-i}) \leq \theta^1$. So, the left-hand side of (9) must be smaller than or equal to $(N-1)\theta^1$. Also, for agent N, we have

$$\begin{aligned}
\theta^1 \cdot 1 - t_N(\theta^1, \bar{\theta}_{-N}) \geqslant 0 \\
\Rightarrow \theta^1 \cdot 1 - (c(N) - \sum_{i \neq N} t_i(\theta^1, \bar{\theta}_{-N})) \geqslant 0 \ (\because BB) \\
\Rightarrow \sum_{i \neq N} t_i(\theta^1, \bar{\theta}_{-N}) \geqslant c(N) - \theta^1.
\end{aligned}$$

Using (9), we summarize some of the inequalities thus obtained below:

$$(N-1)\theta^1 \geq \sum_{i \neq N} t_i(\theta^1, \bar{\theta}_{-i}) \underbrace{=}_{(9)} \sum_{i \neq N} t_i(\theta^1, \bar{\theta}_{-N}) \geqslant c(N) - \theta^1.$$

We thus obtain $N\theta^1 \geqslant c(N)$, which contradicts the assumption that $c(N) > N\theta^1$ for any N. ■

Hence, in all nontrivial cases, we have no hope in finding mechanisms satisfying DSIC, EPIR, EFF, and BB simultaneously, even if we focus on small finite agent economies. Thus, we are left with the trivial cases: (i) $c(N) = N\theta^1$ or (ii) $c(N) = N\theta^M$.

Proposition 3 *Suppose that either (i) $c(N) = N\theta^1$ or (ii) $c(N) = N\theta^M$ holds. In both cases, there exists a mechanism satisfying DSIC, EPIR, EFF, and BB.*

Proof: Suppose that (i) $c(N) = N\theta^1$. Then for any $\theta \in \Theta^N$,

$$x^*(\theta) = 1,$$

and

$$t_i(\theta) = c(N)/N \text{ for any } i \in \mathcal{N}.$$

Suppose that (ii) $c(N) = N\theta^M$. Then for any $\theta \in \Theta^N$,

$$x^*(\theta) = \begin{cases} 1 & \text{if all agents are of type } \theta^M \\ 0 & \text{otherwise,} \end{cases}$$

and for any $i \in \mathcal{N}$,

$$t_i^*(\theta) = \begin{cases} c(N)/N & \text{if } x^*(\theta) = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, these two mechanisms satisfy DSIC, EPIR, EFF, and BB in each case, respectively. ■

1.6 Concluding Remark

This paper characterizes mechanisms satisfying BIC, IIR, BB, and/or EFF for public good production and cost decision in a finite-type environment with risk-neutral, quasi-linear preferences, and fixed-size projects. The main contribution of this paper is to undercover the structure of public good provision mechanisms in a discrete environment (Theorems 1 and 3) and make a coherent comparison with many papers in the literature which deal with the continuum type space. In our discrete setup, we restore many known results of the classic public good provision problem within the standard model.

Overall, this paper shows that even in a finite type environment, "positive results" exist only in very small economies. Recall our simulation result in Section 3.4 in which even five agents are sometime large enough to obtain an impossibility result. This is largely consistent with the literature which usually assumes a continuum type space. In particular, we show that a mechanism satisfying BIC, IIR, EFF, and BB exists only in very small economies and as the size of the economy gets large, the probability that the public good is provided converges to zero in any mechanism satisfying BIC, IIR, and BB. We thus establish the robustness of the overly negative implications in the classical public good provision problem.

Our results in large economies might even suggest the necessity of mandatory payment of taxes as opposed to voluntary contribution. Of course, if we completely ignore this individual rationality constraint, some agent might be forced to make an extremely large payment, which we consider unreasonable. Therefore, we seem to need some restrictions on the size of transfers. What would be nontrivial is that it is not clear at all how we relax the individual rationality constraints or, in other words, how much of mandatory payment should be allowed. This is an interesting question we leave for future research.

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

2.1 Introduction

This paper investigates efficient, voluntary bilateral trades in an interdependent values environment. By "bilateral trade" we mean a simple trading problem in which two individuals, one of whom has a single indivisible object to sell to the other, attempt to agree on exchange of the object for money. So, in this setup, the seller has the full property right for the object to be sold. Efficiency adopted in this paper is an ex post notion, which requires that (i) there be a trade of the good if and only if the buyer's valuation for the good is at least as high as the seller's valuation (decision efficiency) and (ii) whatever the buyer pays is always exactly what the seller receives (budget balance). This paper is mainly concerned with the following normative question: when can an efficient, voluntary trade be implementable in this bilateral trade problem? By the well-known revelation principle, we say that efficient, voluntary trades are possible if there exists a direct revelation mechanism that satisfies Bayesian incentive compatibility (BIC), decision efficiency (EFF), interim individual rationality (IIR), and ex post budget balance (BB). In the case of private values (i.e., each player is certain of the value of the object at the timing of trade), the celebrated impossibility result of Myerson and Satterthwaite (1983) shows that there are generally no mechanisms satisfying BIC, IIR, EFF, and BB in a bilateral trade setting. On the contrary, Cramton, Gibbons, and Klemperer (1987) show that under the equal-share ownership, there is a mechanism satisfying BIC, IIR, EFF, and BB. Hence, the equal-share partnership is dissolved efficiently.

In many practical instances, however, the assumption of private values is violated. This motivates us to investigate when efficient, voluntary bilateral trades are possible in interdependent values environments, which capture a class of situations in which the payoff of an agent depends not only on his own type, but also on the types (or informational signals) of the other agents. Such interdependence is natural in many trading situations. For instance, we consider a situation in which a seller has private information about the quality of the good which influences the valuations of both the seller and a potential buyer. This type of interdependence is the very situation this paper considers. Once we turn to interdependent values environments, however, we are well aware of bad news. We know from Fiesler, Kittsteiner, and Moldovanu (henceforth, FKM, 2003) and Gresik (1991) that Myerson and Satterthwaite's impossibility result is extended to interdependent values environments. FKM (2003, Theorem 4) also show that the efficient partnership dissolution of Cramton, Gibbons, and Klemperer (1987) cannot be extended to interdependent values environments.

To overcome this negative message in interdependent values environments, we seek for more positive results by looking at *two-stage* generalized revelation mechanisms (Mezzetti (2004)): in the first stage, agents are asked to report their type and the allocation of the good is determined on the type reports; after agents observe their allocation payoff, they are asked to report their realized allocation payoff in the second stage; and finally, the monetary transfers are finalized on the reports of both stages. In his Proposition 1, Mezzetti (2003) establishes the generalized revelation principle, which says that it entails no loss of generality to focus on two-stage generalized revelation mechanisms we briefly described above. By this generalized revelation principle, a two-stage generalized revelation mechanism is simply called a two-stage mechanism in this paper.

The assumption behind the use of two-stage mechanisms can be justified. For example, in the context of a labor market, employers learn the quality of the workers after employing them and after both the employer and the worker find out that the the worker is qualified for the job, the worker's contract is upgraded. We find this type of contracts in a tenure-track contract in academic institutions and consider this as a particular type of two-stage mechanisms. This example seems to suggest that a long-term relationship is essential for two-stage mechanisms to be viable. However, we argue this is not even necessary because of the advent of new technologies. As an example, we take up a *smart contract* based on the blockchain technology as a commitment device that prevents agents from reneging the contract terms. As Matsushima and Noda (2020) argue, the participants in the mechanism can replace a long-term relationship by the use of smart contracts so that a two-stage mechanism can be implemented without a trusted third party or long-term relationship. Nevertheless, the power of two-stage mechanisms can sometimes be compromised when it is difficult to justify that an agent who obtains the good can experience its quality. To see this, consider a situation in which the object to be traded is some art work and an agent's payoff from obtaining this art work depends on how the other people appreciate it. In this case, the agent will not be able to experience the quality of the object by consuming it. Hence, the power of two-stage mechanisms is sometimes dubious. In any case, we stress that our question here is mainly theoretical. If no two-stage mechanisms implement an efficient, voluntary trade, it is almost impossible to imagine that any mechanism used in a more realistic setup can implement it. In this sense, we are concerned with pushing the boundary between what is implementable and what is not by expanding our scope into two-stage mechanisms.

Considering two-stage mechanisms, we modify the notion of incentive compatibility. Following Mezzetti (2004), we say that 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. Here, the main question of our paper is rephrased: "when does there exist a two-stage mechanism satisfying BIC, IIR, EFF, and BB in a bilateral trade model with interdependent values?" In a general mechanism design problem, Mezzetti (2004) proposes the *generalized two-stage Groves* mechanism and shows that it always satisfies BIC, EFF, and BB. When we are concerned with efficient trades, the standard *one-stage* Groves mechanism is shown to be a "canonical" mechanism (See Krishna and Perry (2000) and Williams (1999) for the case of private values and FKM (2003) for the case of interdependent values). What we mean by "canonical" is that if we are to investigate the existence of the standard one-stage mechanisms satisfying BIC and EFF, we lose nothing to restrict our search to the family of the Groves mechanisms.

This paper considers a bilateral trade model with the following features: (i) each agent's type space constitutes a nonempty closed, bounded interval over the real line; (ii) each agent's type is chosen independently across agents; (iii) each agent's valuation depends on not only his own type but also the type of other agent (i.e., interdependent values); (iv) each agent's valuation for the object is strictly increasing in both his own type and the opponent's type; (v) utilities are quasilinear and so, utilities consist of the sum of a payoff from an outcome decision and a monetary transfer; and (vi) the single crossing property is satisfied. This condition is imposed in FKM (2003). It means that each agent's type must have a greater effect on his own valuation than on that of the other agent.

In Section 2.3, we confine our attention to a stylized model in which each agent's type is chosen from the uniform distribution over $[0, 1]$ and each agent i's valuation for the object is represented by a linear function, i.e., $\tilde{u}_i(\theta_i, \theta_j) = \theta_i + \gamma_i \theta_j$, where γ_i denotes the degree of interdependence of preferences for agent i. In this context, the single crossing property requires that $\gamma_i < 1$ for each agent i. We find it natural to start our investigation from the generalized two-stage Groves mechanism. We show that the generalized two-stage Groves mechanism never satisfies IIR (Proposition 4). Throughout the paper, we revisit this example multiple times to illustrate the implications of our analysis.

In Section 2.4, we establish the main result of this paper in a general environment. This section consists of several subsections. In Subsection 2.4.1, we introduce an additional property imposed on two-stage mechanisms. The property says that if trade does not occur, no payments are made. We call this property the "no-trade-then-no-payments" (henceforth, NTNP) property. In the example in Section 2.4.1, we confirm that the generalized two-stage Groves mechanism violates the NTNP property (Claim 6). We impose another additional monotonicity property on two-stage mechanisms. We say that a two-stage mechanism is *monotone* if the buyer's payment is nondecreasing in his own type announcement conditional upon trade occurring. In the example of Section 2.3, we confirm that the generalized two-stage Groves mechanism is indeed monotone (Claim 7). This suggests that monotonicity is a mild condition. In Subsection 2.4.2, we propose a two-stage NTNP, monotone mechanism which is used for our main result. Subsection 2.4.3 introduces Condition α which is needed for our main result. Subsection 2.4.4 states Theorem 5 as our main result. Theorem 5 of this paper says that if our Condition α is satisfied, the two-stage NTNP, monotone mechanism proposed in Subsection 2.4.2 satisfies BIC, EFF, BB, and IIR. Thus, the generalized two-stage Groves mechanism turns out to be "not" canonical because the generalized two-stage Groves mechanism does not implement an efficient, voluntary trade, whereas our proposed two-stage mechanism implements it. What distinguishes our proposed two-stage mechanism from the generalized two-stage Groves one is the NTNP property.

Section 2.5 assesses the restrictiveness of our Condition α using the example in Section 2.3. We argue that our Condition α is very weak because it is satisfied as long as the buyer's degree of interdependence of preferences (γ_2) is not too high relative to the seller's counterpart γ_1. By a set of simulation results, we conclude that our Condition α is satisfied for a large number of cases.

In Section 2.6, we compare our results with the results of Galavotti, Muto, and Oyama (henceforth GMO, 2011), who consider the problem of partnership dissolution of Cramton, Gibbons, and Klemperer (1987) in an interdependent values environment. GMO (2011) show in their Theorem 4 that when GMO's Assumption 5.1 is satisfied, for any ownership structure, there exists a two-stage mechanism satisfying BIC, IIR, EFF, and BB.[16] To make our comparison meaningful, we focus on our bilateral trade setup, i.e., there are only two agents and the seller has the full property right over the good. We first show in our Lemma 12 that our Condition α is weaker than GMO's Assumption 5.1. Second, we show in Lemma 13 that in the example in Section 3, GMO's Assumption 5.1 is satisfied if and only if $\gamma_1 = \gamma_2$, i.e., the seller's degree of interdependence of preferences is exactly identical to the buyer's counterpart. This suggests that GMO's Assumption 5.1 is generically violated in our bilateral trade setup. Of course, the advantage of GMO (2011) lies in rather handling any ownership structure, which exhibits a contrast with this paper's focus on a particular ownership structure in which the seller has the full property right over the good.

The rest of the paper is organized as follows. In Section 2.2, we introduce the general notation and basic concepts for the paper and go over some key important results in the literature to benchmark our paper. Section 2.3 specializes in a highly stylized but well studied model of bilateral trade with interdependent values. In Section 2.4, we introduce our Condition α and discuss our main result. Section 2.5 assesses the restrictiveness of our Condition α. In Section 2.6, we compare the results of our paper with those of GMO

[16] To be precise, their result is stronger than this because GMO (2011) strengthen IIR into ex post individual rationality (EPIR). See Section 2.6 for the definition of EPIR.

(2011). In the Appendix, we provide all the proofs of the results omitted from the main text of the paper.

2.2 Preliminaries

A seller (agent 1) has one indivisible object for sale and there is one potential buyer (agent 2). Each agent $i \in \{1, 2\}$ has his type θ_i about the value of the object. The set of possible types for agent i is denoted by Θ_i and we assume that $\Theta_i = [\underline{\theta}_i, \bar{\theta}_i]$ is a closed, bounded interval over $\mathbb{R}$ with $\underline{\theta}_i < \bar{\theta}_i$. We use the notation convention that $\Theta = \Theta_1 \times \Theta_2$ and $\Theta_{-i} = \Theta_j$ where $j \neq i$ with a generic element θ_{-i}. Types are independently distributed between agents. For each agent $i \in \{1, 2\}$, denote by f_i and F_i the probability density function and cumulative distribution function of θ_i, respectively. We further assume that $f_i(\theta_i) > 0$ for all $\theta_i \in (\underline{\theta}_i, \bar{\theta}_i)$ and $i \in \{1, 2\}$.

Let $q \in Q = [0, 1]$ be the probability that the good is sold to the buyer, or trading probability for short. Preferences of each agent $i \in \{1, 2\}$ are given by $U_i : Q \times \Theta \times \mathbb{R} \to \mathbb{R}$, which depends on the trading probability q, the type profile θ and his monetary transfer p_i:

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

where $u_i(q, \theta)$ is agent i's allocation payoff and $\tilde{u}_i(\theta)$ is agent i's valuation for the object in state $\theta \in \Theta$. We assume that for all $i, j \in \{1, 2\}$ with $j \neq i$, $\tilde{u}_i(\theta_i, \theta_j)$ is differentiable in both θ_i and θ_j and $\tilde{u}_{i,i} \equiv \partial\tilde{u}_i(\theta_i, \theta_j)/\partial\theta_i > 0$ and $\tilde{u}_{i,j} \equiv \partial\tilde{u}_i(\theta_i, \theta_j)/\partial\theta_j > 0$ (i.e., strictly increasing in both agents' types).

We further assume the following single crossing condition:

$$\tilde{u}_{i,i} > \tilde{u}_{j,i}, \ \ \forall i, j \in \{1, 2\} \text{ with } i \neq j.$$

When the agents' types are independently distributed, as we assume, Dasgupta and Maskin (2000, footnote 13) argue that in the auction setups, the single crossing property is necessary for the existence of mechanisms satisfying efficiency. This is one of the reasons why we impose the single crossing property. Another reason for this imposition is that we need to rely on Theorem 5 (shown below) of Fieseler, Kittsteiner, and Moldovanu (FKM, 2003) who impose the single crossing condition on their environment. We denote by $\Pi_i = \{\tilde{u}_i(\theta) |\ \theta \in \Theta\}$ the range of agent i's allocation payoff. We assume that for any realization of the type profile $\theta \in \Theta$, if agent i receives the object, he observes his realized allocation payoff $\tilde{u}_i(\theta)$ before final transfers are made.

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 each agent's type constitutes a Bayesian Nash equilibrium, which is known as Bayesian incentive compatibility (BIC). In the case of private-value environments, Myerson and Satterthwaite (1983) show that efficiency and voluntary participation are not achieved in an incentive compatible manner. Focusing on the standard one-stage direct mechanisms, Fieseler, Kittsteiner, and Moldovanu (2003) establish the following counterpart of the Myerson and Satterthwaite impossibility result in an interdependent values environment.

Lemma 5 *[Theorem 5 in FKM (2003)] There exists a one-stage mechanism satisfying Bayesian incentive compatibility (BIC), interim individual rationality (IIR), ex post efficiency (EFF), and ex post budget balance (BB) if and only if either one of the following two conditions hold:*

1. *there is a price p such that* $\mathbb{E}_{\theta_1}\left[\tilde{u}_2(\theta_1, \underline{\theta}_2)\right] \geqslant p \geqslant \mathbb{E}_{\theta_2}\left[\tilde{u}_1(\bar{\theta}_1, \theta_2)\right]$.
2. $\tilde{u}_2(\theta) \leq \tilde{u}_1(\theta)$ *for all* $\theta \in \Theta$.[17]

Remark: This first condition means that there is a price p such that all types of buyer and seller agree to trade at this price. The second condition means that it is always efficient

[17] To be precise, FKM (2003) only require ex ante budget surplus rather than ex post budget balance (BB), which we assume. However, Borgers (2015) in Proposition 3.6 and Borgers and Norman (2009) in Proposition 2 show that in the case of independent beliefs, as in our paper, ex ante budget surplus implies ex post budget balance (BB).

not to trade. In this second case, a mechanism always dictating no trades and transfers satisfies BIC, IIR, EFF, and BB trivially. It is important to note that Gresik (1991) already derived a different condition for the existence of a one-stage mechanism satisfying all the four properties (Theorem 3 of Gresik (1991)).

In the example in Section 2.3, which is a stylized, representative example, we know by Lemma 5 that there are no one-stage mechanisms satisfying BIC, IIR, EFF, and BB. Taking this negative result seriously and looking for more positive results, we then 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
- $\tau = (\tau[1], \tau[2])$ where $\tau[i] : M^1 \times M_i^2 \to \mathbb{R}^2$ is the transfer rule specifying the monetary transfer for both agents when agent i receives the good at the beginning of the second stage.

In words, in the first stage, after observing his own type, each agent sends a message from M_i^1 and then the good is allocated according to the decision rule δ; in the second stage, after agent i who receives the good (either the seller or the buyer) 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 $r_i = (r_i^1, r_i^2)$ agent i's strategy such that $r_i^1 : \Theta_i \to M_i^1$ is his strategy in the first stage and $r_i^2 : Q \times \Theta_i \times \Pi_i \to M_i^2$ is his strategy in the second stage.

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 $t = (t[1], t[2])$ such that $t[i] : \Theta \times \Pi_i \to \mathbb{R}^2$ is given by the composite function $t[i](\theta; u_i) = \tau[i](r^1(\theta), r^2(\delta(r^1(\theta)), \theta, u_i))$. Since each agent i's allocation payoff $\tilde{u}_i(\theta_i, \theta_{-i})$ depends on the whole type profile, then the second-stage reports in the generalized revelation mechanism indeed provide extra information about the type profile, while there is a loss of generality in assuming that the designer only uses the standard "one-stage" revelation mechanisms.

Following Mezzetti (2003), we adopt perfect Bayesian equilibrium as a solution concept and appeal to the following generalized revelation principle, the counterpart of revelation principle in one-stage mechanisms.[18]

Lemma 6 (The Generalized Revelation Principle in Mezzetti (2003)) *For any perfect Bayesian equilibrium outcome of any two-stage mechanism (M^1, M^2, δ, τ), there exist a generalized revelation mechanism (Θ, Π, x, t) and a perfect Bayesian equilibrium such that, for each agent, reporting his true allocation payoff in the second stage and reporting his true type in the first stage constitute the equilibrium strategy.*

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

Definition 9 *A two-stage mechanism (Θ, Π, x, t) satisfies Bayesian incentive compatibility (BIC) if truthtelling in both stages constitutes an equilibrium strategy of each agent in a perfect Bayesian equilibrium; that is, for each agent i and each type profile $(\theta_i, \theta_{-i}), (\theta_i^r, \theta_{-i}^r) \in \Theta_i \times \Theta_{-i}$, the equilibrium second-stage report is $u_i^r = u_i(x(\theta_i^r, \theta_{-i}^r), \theta_i, \theta_{-i})$ and the equilibrium first-stage report is $\theta_i^r = \theta_i$.*

BIC implies that, given the first-stage report, each agent reports his realized allocation payoff truthfully in the second stage. BIC further implies that, on the equilibrium path, each agent reports his true type in the first stage and for any type profile $(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2$, $u_i(x(\theta_1, \theta_2), \theta_1, \theta_2)$ is agent i's true allocation payoff.

We also 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

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

non-participation. To be specific,

$$U_1^O(\theta_1) = \int_{\Theta_2} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) \text{ for all } \theta_1 \in \Theta_1$$

and

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

We introduce the following individual rationality constraint:

Definition 10 *A two-stage mechanism* (Θ, Π, x, t) *satisfies interim individual rationality (IIR) if, for all* $\theta_1 \in \Theta_1$*,*

$$\int_{\Theta_2} \left(u_1(x(\theta_1, \theta_2), \theta_1, \theta_2) + t_1(\theta_1, \theta_2; u_1, u_2)\right) dF(\theta_2) \geq U_1^O(\theta_1),$$

and for all $\theta_2 \in \Theta_2$*,*

$$\int_{\Theta_1} \left(u_2(x(\theta_1, \theta_2), \theta_1, \theta_2) + t_2(\theta_1, \theta_2; u_1, u_2)\right) dF(\theta_1) \geq U_2^O(\theta_2),$$

where $u_1 = u_1(x(\theta_1, \theta_2), \theta_1, \theta_2)$ *and* $u_2 = u_2(x(\theta_1, \theta_2), \theta_1, \theta_2)$*.*

Note that this paper's formulation of IIR is the same as the one used by FKM (2003) and Gresik (1991). Next, we require that trade occur if and only if there are gains from trade from ex post point of view.

Definition 11 *A two-stage mechanism* (Θ, Π, x, t) *satisfies decision efficiency (EFF) if, for all* $(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2$*,*

$$x(\theta_1, \theta_2) \in \arg\max_{x \in Q} \left(u_1(x, \theta_1, \theta_2) + u_2(x, \theta_1, \theta_2)\right).$$

In what follows, we denote by x^* the efficient decision rule. We further require that what the seller receives be exactly the same as what the buyer pays.

Definition 12 *A two-stage mechanism* (Θ, Π, x, t) *satisfies ex post budget balance (BB) if, for all* $(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2$*,*

$$t_1(\theta_1, \theta_2; u_1, u_2) + t_2(\theta_1, \theta_2; u_1, u_2) = 0,$$

where $u_1 = u_1(x(\theta_1, \theta_2), \theta_1, \theta_2)$ *and* $u_2 = u_2(x(\theta_1, \theta_2), \theta_1, \theta_2)$*.*

Mezzetti (2004) proposes the following generalized two-stage Groves mechanism and shows that it always satisfies BIC, BB and EFF.

Definition 13 *A two-stage mechanism* (Θ, Π, x^*, t^G) *is called the generalized two-stage Groves mechanism if, for each agent* $i \in \{1, 2\}$*, type report* $(\theta_i^r, \theta_{-i}^r) \in \Theta_i \times \Theta_{-i}$ *and payoff report* $(u_i^r, u_{-i}^r) \in \Pi_i \times \Pi_{-i}$*,*

$$t_i^G(\theta_i^r, \theta_{-i}^r; u_i^r, u_{-i}^r) = u_{-i}^r - h_i(\theta_i^r, \theta_{-i}^r)$$

where

$$\begin{aligned} 2h_i(\theta_i^r, \theta_{-i}^r) &= \sum_{j=1}^{2} u_j\left(x^*(\theta^r), \theta^r\right) - \mathbb{E}_{\theta_{-i}}\left(\sum_{j=1}^{2} u_j\left(x^*(\theta_i^r, \theta_{-i}), \theta_i^r, \theta_{-i}\right)\right) \\ &\quad + \mathbb{E}_{\theta_{-(i+1)}}\left(\sum_{j=1}^{2} u_j\left(x^*(\theta_{i+1}^r, \theta_{-(i+1)}), \theta_{i+1}^r, \theta_{-(i+1)}\right)\right) \end{aligned}$$

with $\mathbb{E}_{\theta_{-i}}$ *being the expectation operator over* θ_{-i} *and* $\mathbb{E}_{\theta_{-3}} = \mathbb{E}_{\theta_{-1}}$*.*

Although the result below is already proved by Mezzetti (2004), we find it instructive to go through its proof to appreciate how the generalized two-stage Groves mechanism works in our bilateral trade setup.

Lemma 7 (Proposition 2 in Mezzetti (2004)) *The generalized two-stage Groves mechanism always satisfies BIC, EFF, and BB.*

Proof: The transfer rule is constructed in such a way that the generalized two-stage Groves mechanism always satisfies BIC and BB. Note that agent i's transfer is independent of his payoff report u_i^r so that he has no incentive to deviate in the second stage. Suppose agent i of type θ_i misreports θ_i^r whereas his opponent always reports the true type θ_{-i} in the first stage. Assume further that both agents report the allocation payoff truthfully in the second stage, i.e., $u_i^r = u_i(x^*(\theta_i^r, \theta_{-i}), \theta_i, \theta_{-i})$ and $u_{-i}^r = u_{-i}(x^*(\theta_i^r, \theta_{-i}), \theta_i, \theta_{-i})$.

Then, agent i's expected utility is

$$
\begin{aligned}
& \mathbb{E}_{\theta_{-i}}\left[u_i(x^*(\theta_i^r,\theta_{-i}),\theta_i,\theta_{-i}) + t_i^G(\theta_i^r,\theta_{-i}; u_i(x^*(\theta_i^r,\theta_{-i}),\theta_i,\theta_{-i}), u_{-i}(x^*(\theta_i^r,\theta_{-i}),\theta_i,\theta_{-i}))\right] \\
= \; & \mathbb{E}_{\theta_{-i}}\left[u_i(x^*(\theta_i^r,\theta_{-i}),\theta_i,\theta_{-i}) + u_{-i}(x^*(\theta_i^r,\theta_{-i}),\theta_i,\theta_{-i}) - h_i(\theta_i^r,\theta_{-i})\right] \\
= \; & \mathbb{E}_{\theta_{-i}}\left(\sum_{j=1}^{2} u_i(x^*(\theta_i^r,\theta_{-i}),\theta_i,\theta_{-i})\right) - \mathbb{E}_{\theta_{-i}}\left(h_i(\theta_i^r,\theta_{-i})\right) \\
= \; & \mathbb{E}_{\theta_{-i}}\left(\sum_{j=1}^{2} u_i(x^*(\theta_i^r,\theta_{-i}),\theta_i,\theta_{-i})\right) - \frac{1}{2}\mathbb{E}\left(\sum_{j=1}^{2} u_j(x^*(\theta),\theta)\right) \\
\leq \; & \mathbb{E}_{\theta_{-i}}\left(\sum_{j=1}^{2} u_i(x^*(\theta_i,\theta_{-i}),\theta_i,\theta_{-i})\right) - \frac{1}{2}\mathbb{E}\left(\sum_{j=1}^{2} u_j(x^*(\theta),\theta)\right),
\end{aligned}
$$

where $\mathbb{E}$ denotes the expectation over (θ_i,θ_{-i}), and the last inequality follows because, by definition, $x^*(\theta_i,\theta_{-i}) \in \arg\max_{x\in Q}\sum_{j=1}^{2} u_j(x,\theta_i,\theta_{-i})$ and the second term is a constant. Hence, agent i achieves the highest expected utility by truth-telling so that BIC is satisfied.

Furthermore, on the equilibrium path where each agent i reports his true type θ_i and true allocation payoff $u_i = u_i(x^*(\theta_i,\theta_{-i}),\theta_i,\theta_{-i})$, the total transfer is computed as follows: for each $(\theta_1,\theta_2)\in\Theta$,

$$
\begin{aligned}
& t_1^G(\theta_1,\theta_2; u_1(x^*(\theta_1,\theta_2),\theta_1,\theta_2), u_2(x^*(\theta_1,\theta_2),\theta_1,\theta_2)) \\
& + t_2^G(\theta_1,\theta_2; u_1(x^*(\theta_1,\theta_2),\theta_1,\theta_2), u_2(x^*(\theta_1,\theta_2),\theta_1,\theta_2)) \\
= \; & u_2(x^*(\theta_1,\theta_2),\theta_1,\theta_2) - h_1(\theta_1,\theta_2) + u_1(x^*(\theta_1,\theta_2),\theta_1,\theta_2) - h_2(\theta_1,\theta_2) \\
= \; & \sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2) \\
& -\frac{1}{2}\left[\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2) - \mathbb{E}_{\theta_2}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right) + \mathbb{E}_{\theta_1}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right)\right] \\
& -\frac{1}{2}\left[\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2) - \mathbb{E}_{\theta_1}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right) + \mathbb{E}_{\theta_2}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right)\right] \\
= \; & \sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2) - \sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2) \\
= \; & 0.
\end{aligned}
$$

Hence, BB is satisfied. This completes the proof. ■

However, it is not clear whether the generalized two-stage Groves mechanism also satisfies IIR or not. We investigate this issue by means of an example in the next section.

2.3 An Example

In this section, we show by means of an example that the generalized two-stage Groves mechanism with lump-sum transfers always fails IIR.

Both agents' types are uniformly distributed on the unit interval $[0, 1]$ and for each type profile $(\theta_1, \theta_2) \in [0, 1]^2$, their valuation functions are $\tilde{u}_1(\theta_1, \theta_2) = \theta_1 + \gamma_1\theta_2$ and $\tilde{u}_2(\theta_1, \theta_2) = \theta_2 + \gamma_2\theta_1$ where $\gamma_1, \gamma_2 > 0$. Then,

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

implying that the efficient decision rule depends on the values of γ_1 and γ_2. We need to satisfy the single crossing condition, which implies that $\gamma_1 < 1$ and $\gamma_2 < 1$. Then, we are left with two cases to consider: (i) $0 < \gamma_2 \leq \gamma_1 < 1$ and (ii) $0 < \gamma_1 < \gamma_2 < 1$.

In this example, the first condition in Lemma 5 amounts to $\gamma_2 - \gamma_1 \geqslant 2$, which contradicts the assumption $\gamma_i \in (0, 1)$ for each $i \in \{1, 2\}$. Similarly, the second condition in Lemma 5 implies $\tilde{u}_2(\underline{\theta}_1, \bar{\theta}_2) < \tilde{u}_1(\underline{\theta}_1, \bar{\theta}_2)$, which amounts to $1 \geq \gamma_1$. It also contradicts the assumption that $\gamma_1 \in (0, 1)$. Therefore, there are no one-stage mechanisms satisfying BIC, IIR, EFF, and BB in this example.

In Proposition 4 below, we show that the generalized two-stage Groves mechanism violates IIR in both cases (i) and (ii). For this result, we even allow for lump-sum transfers in addition to the original transfers in the generalized two-stage Groves mechanism. There are two reasons why we emphasize the addition of lump-sum transfers to it. First, this simply makes our result stronger. Second, this has a close connection to Theorem 2 of Krishna and Perry (2000), which roughly shows that in a private values environment, the search for a (one-stage) mechanism satisfying all the properties is reduced to the the class of the generalized (one-stage) Groves mechanisms (or what Krishna and Perry (2000) call the VCG mechanisms) with lump-sum transfers. As we argue in our Theorem 1, there is a loss of generality to restrict our search to the generalized two-stage Groves mechanisms with lump-transfers.

Proposition 4 *The generalized two-stage Groves mechanism* (Θ, Π, x^*, t^G) *with lump-sum transfers violates IIR in both cases.*

Proof: Recall that on the equilibrium path in which both agents' reports are truthful in both stages, agent i of type θ_i receives the following expected utility:

$$\begin{aligned} U_i^G(\theta_i) &= \mathbb{E}_{\theta_{-i}}\left[u_i(x^*(\theta_i,\theta_{-i}),\theta_i,\theta_{-i}) + t_i^G(x^*(\theta_i,\theta_{-i}),u_i,u_{-i})\right] \\ &= \mathbb{E}_{\theta_{-i}}\left(\sum_{j=1}^{2} u_j(x^*(\theta_i,\theta_{-i}),\theta_i,\theta_{-i})\right) - \frac{1}{2}\mathbb{E}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right), \end{aligned}$$

where $u_i = u_i(x^*(\theta_i,\theta_{-i}),\theta_i,\theta_{-i})$, $u_{-i} = u_{-i}(x^*(\theta_i,\theta_{-i}),\theta_i,\theta_{-i})$, $\mathbb{E}_{\theta_{-i}}$ denotes the expectation over θ_{-i}, and $\mathbb{E}$ denotes the expectation over (θ_i,θ_{-i}). Then we can derive the worst-off type θ_i^w of each agent i from participating in the generalized two-stage Groves mechanism:

$$\begin{aligned} \theta_i^w &\in \arg\min_{\theta_i\in\Theta_i}\left[U_i^G(\theta_i) - U_i^O(\theta_i)\right] \\ &= \arg\min_{\theta_i\in\Theta_i}\left[\mathbb{E}_{\theta_{-i}}\left(\sum_{j=1}^{2} u_j(x^*(\theta_i,\theta_{-i}),\theta_i,\theta_{-i})\right) - \frac{1}{2}\mathbb{E}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right) - U_i^O(\theta_i)\right]. \end{aligned}$$

Since the second term is a constant and hence independent of θ_i, it is equivalent to say

$$\theta_i^w \in \arg\min_{\theta_i\in\Theta_i}\left[\mathbb{E}_{\theta_{-i}}\left(\sum_{j=1}^{2} u_j(x^*(\theta_i,\theta_{-i}),\theta_i,\theta_{-i})\right) - U_i^O(\theta_i)\right].$$

Let $L_i \equiv U_i^O(\theta_i^w) - U_i^G(\theta_i^w)$ be the expected loss for agent i's worst-off type. By Proposition 3 of Mezzetti (2003), we know that the generalized two-stage Groves mechanism with lump-sum transfers satisfies IIR without violating BIC, EFF and BB if and only if $L_1 + L_2 \leq 0$. So, it remains to verify whether $L_1 + L_2 \leq 0$ is satisfied in this example. There are two cases we consider: (i) $0 < \gamma_2 \leq \gamma_1 < 1$ and (ii) $0 < \gamma_1 < \gamma_2 < 1$.

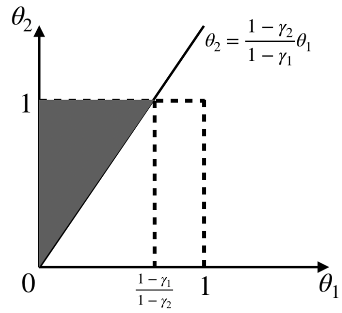
Case (i): $0 < \gamma_2 \leq \gamma_1 < 1$

Since $\tilde{u}_2(\theta_1,\theta_2) - \tilde{u}_1(\theta_1,\theta_2) = (1-\gamma_1)\theta_2 - (1-\gamma_2)\theta_1$ for each $(\theta_1,\theta_2) \in \Theta$, then we have that $\tilde{u}_2(\theta_1,\theta_2) > \tilde{u}_1(\theta_1,\theta_2)$ if and only if $\theta_2 > (1-\gamma_2)\theta_1/(1-\gamma_1)$. Hence, the efficient decision rule dictates that, for each $(\theta_1,\theta_2) \in \Theta$,

$$x^*(\theta_1,\theta_2) = \begin{cases} 1 & \text{if } \theta_2 > (1-\gamma_2)\theta_1/(1-\gamma_1) \\ 0 & \text{otherwise.} \end{cases}$$

The following figure illustrates the decision at different type profiles in this case; in particular, the shaded region represents $\Theta^* = \{(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2 : x^*(\theta_1, \theta_2) = 1\}$, which exhausts all the type profiles in which trade occurs.

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



Claim 4 $L_1 + L_2 > 0$ *when* $0 < \gamma_2 \leq \gamma_1 < 1$.

Proof: The proof is in the Appendix. ■

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

Similar to the previous case, for each $(\theta_1, \theta_2) \in \Theta$, we have that $\tilde{u}_2(\theta_1, \theta_2) > \tilde{u}_1(\theta_1, \theta_2)$ if and only if $\theta_2 > (1 - \gamma_2)\theta_1/(1 - \gamma_1)$. Hence, the efficient decision rule dictates that, for each $(\theta_1, \theta_2) \in \Theta$,

$$x^*(\theta_1, \theta_2) = \begin{cases} 1 & \text{if } \theta_2 > (1 - \gamma_2)\theta_1/(1 - \gamma_1) \\ 0 & \text{otherwise.} \end{cases}$$

Figure 2 below illustrates the decision at different type profiles in this case; in particular, the shaded region represents $\Theta^* = \{(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2 : x^*(\theta_1, \theta_2) = 1\}$, which describes the set of possible type profiles for which it is efficient to trade.

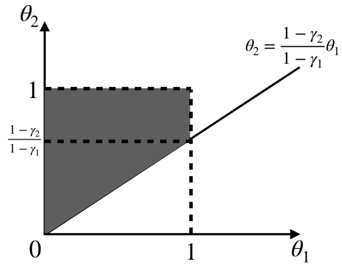
Claim 5 $L_1 + L_2 > 0$ *when* $0 < \gamma_1 < \gamma_2 < 1$.

Proof: The proof is in the Appendix. ■

By Claims 4 and 5, the generalized two-stage Groves mechanism fails IIR. ■

This example will become important for illustrating many of our results, and we shall revisit it in multiple times in due course.

Figure 2: when $0 < \gamma_1 < \gamma_2 < 1$



2.4 The Main Result

This section is organized as follows. In Subsection 2.4.1, we propose two properties on the class of two stage mechanisms. The first property is called the "no-trade-and-then-no-payment (NTNP) property, which means that when trade does not occur, no agents either receive subsidies or make payments. The second property requires that a two-stage mechanism be "monotone" in the sense that, conditional on the trade occurring, the buyer's payment is nondecreasing in his own type announcement. Subsection 2.4.2 proposes a NTNP, monotone two-stage mechanism which is used for our main result (Theorem 1). Subsection 2.4.3 introduces Condition α. In the example in Section 2.3, Condition α loosely says that the buyer's degree of interdependence of preferences is not too high relative to the seller's counterpart. In Subsection 2.4.4, we show in our Theorem 1 that when Condition α holds, our proposed NTNP, monotone two-stage mechanism satisfies BIC, IIR, EFF, and BB.

2.4.1 A Class of Two-Stage Mechanisms

Since the generalized two-stage Groves mechanism always fails IIR in the example of Section 2.3, we propose a new class of two-stage mechanisms which satisfy all the desired properties including IIR. To do so, we first impose the following property on two-stage mechanisms.

Definition 14 (NTNP) *A two-stage mechanism* (Θ, Π, x, t) *satisfies the "no-trade-then-no-payments" (NTNP) property if, for any type realization* $(\theta_1, \theta_2) \in \Theta$*,*

$$x^*(\theta_1, \theta_2) = 0 \Rightarrow t_1(\theta_1, \theta_2; u_1, u_2) = t_2(\theta_1, \theta_2; u_1, u_2) = 0,$$

where $u_1 = u_1(x^*(\theta_1,\theta_2),\theta_1,\theta_2) = \tilde{u}_1(\theta_1,\theta_2)$ *and* $u_2 = u_2(x^*(\theta_1,\theta_2),\theta_1,\theta_2) = 0.$

We observe that the NTNP property has a bite only if the reported type profile is the true type profile and the second stage report is also the agents' true allocation payoffs. Therefore, it imposes no restrictions on monetary transfers if the agents deviate from reporting their true allocation payoffs in the second stage. In other words, large amounts of penalties can be imposed off the equilibrium path.

In what follows, we call a two-stage mechanism satisfying this property a two-stage NTNP mechanism. We first confirm that in the example of Section 2.3, the generalized two-stage Groves mechanism violates this property.

Claim 6 *In the example of Section 2.3, the generalized two-stage Groves mechanism* (Θ,Π,x^*,t^G) *always violates NTNP.*

Remark: In the generalized two-stage Groves mechanism, even if trade does not occur in some state, the buyer might receive some positive subsidy from the seller. This is the reason why NTNP is violated.

Proof: The proof is in the Appendix. ■

This result suggests that the NTNP property is a defining one that is distinguished from the generalized two-stage Groves mechanism. To propose another property we impose on two-stage mechanisms, we first establish the following useful lemma:

Lemma 8 *Suppose the single crossing condition holds. Then, there exists a unique cutoff point* $\theta_2^* \in (\underline{\theta}_2,\bar{\theta}_2]$ *such that for all* $\theta_2 \in \Theta_2$,

$$\int_{\Theta_1} x^*(\theta_1,\theta_2)dF_1(\theta_1) \begin{cases} < 1 & \text{if } \theta_2 < \theta_2^* \\ = 1 & \text{if } \theta_2 \geq \theta_2^*. \end{cases}$$

Proof: There are two cases we need to consider. The first case is that $\int_{\Theta_1} x^*(\theta_1,\theta_2)dF_1(\theta_1) < 1$ for all $\theta_2 < \bar{\theta}_2$. The second case is that there exists $\theta_2^* \in (\underline{\theta}_2,\bar{\theta}_2)$ such that $\int_{\Theta_1} x^*(\theta_1,\theta_2)dF_1(\theta_1) < 1$ for all $\theta_2 < \theta_2^*$ and $\int_{\Theta_1} x^*(\theta_1,\theta_2)dF_1(\theta_1) = 1$ for any $\theta_2 \geq \theta_2^*$. When $\theta_2^* = \bar{\theta}_2$, the event $\{(\theta_1,\theta_2) \in \Theta|\ \theta_2 \geq \theta_2^*\}$ is of measure zero in Θ. Therefore, if $\theta_2^* = \bar{\theta}_2$, the expression $\int_{\Theta_1} x^*(\theta_1,\bar{\theta}_2)dF_1(\theta_1) = 1$ does not affect the calculation of interim expected payoffs of any agent at all so that this requirement is inconsequential. Therefore, the first case can be handled as a special case of the second case by setting $\theta_2^* = \bar{\theta}_2$.

Thus, we assume that $\theta_2^* \in (\underline{\theta}_2, \bar{\theta}_2)$. Suppose on the contrary that there exists some $\tilde{\theta}_2 < \hat{\theta}_2$ such that

$$\int_{\Theta_1} x^*(\theta_1, \tilde{\theta}_2) dF_1(\theta_1) = 1,$$

and

$$\int_{\Theta_1} x^*(\theta_1, \hat{\theta}_2) dF_1(\theta_1) < 1.$$

Note that $\int_{\Theta_1} x^*(\theta_1, \tilde{\theta}_2) dF_1(\theta_1) = 1$ implies $\tilde{u}_2(\theta_1, \tilde{\theta}_2) > \tilde{u}_1(\theta_1, \tilde{\theta}_2)$ for all $\theta_1 \in \Theta_1$. By the single crossing condition, for any $\theta_1 \in \Theta_1$, $\tilde{u}_2(\theta_1, \theta_2)$ must grow faster than $\tilde{u}_1(\theta_1, \theta_2)$ as θ_2 increases; since $\hat{\theta}_2 > \tilde{\theta}_2$ and $\tilde{u}_2(\theta_1, \tilde{\theta}_2) > \tilde{u}_1(\theta_1, \tilde{\theta}_2)$ for all $\theta_1 \in \Theta_1$, it follows that

$$\tilde{u}_2(\theta_1, \hat{\theta}_2) > \tilde{u}_1(\theta_1, \hat{\theta}_2)$$

for all $\theta_1 \in \Theta_1$, or equivalently,

$$\int_{\Theta_1} x^*(\theta_1, \hat{\theta}_2) dF_1(\theta_1) = 1,$$

contradicting our hypothesis. This completes the proof. ■

To have a better understanding of Lemma 8, we also provide two figures for illustration. The following figures illustrate the allocation decision at different type profiles in general. The shaded region represents $\Theta^* = \{(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2 : x^*(\theta_1, \theta_2) = 1\}$, which describes the set of possible type profiles for which it is efficient to trade. In the left figure, we have $\int_{\Theta_1} x^*(\theta_1, \theta_2) dF_1(\theta_1) < 1$ for all $\theta_2 < \bar{\theta}_2$. In the right figure, it is always efficient to trade when θ_2 is greater than the cutoff type θ_2^*.

We introduce the following monotonicity property on the class of two-stage mechanisms.

Definition 15 *Let $\theta_2^* \in (\underline{\theta}_2, \bar{\theta}_2]$ be the unique cutoff point identified in Lemma 8. A two-stage mechanism (Θ, Π, x^*, t^M) is monotone if, for any $\theta_1^r \in \Theta_1$, any $\theta_2^r, \hat{\theta}_2^r \in \Theta_2$, and any $(u_1^r, u_2^r), (\hat{u}_1^r, \hat{u}_2^r) \in \Pi_1 \times \Pi_2$, whenever $\hat{\theta}_2^r > \theta_2^r$ and $x^*(\theta_1^r, \hat{\theta}_2^r) = x^*(\theta_1^r, \theta_2^r) = 1$, then*

$$\begin{cases} t_2^M(\theta_1^r, \hat{\theta}_2^r, \hat{u}_1^r, \hat{u}_2^r) < t_2^M(\theta_1^r, \theta_2^r, u_1^r, u_2^r) & \text{if } \hat{\theta}_2^r < \theta_2^* \\ t_2^M(\theta_1^r, \hat{\theta}_2^r, \hat{u}_1^r, \hat{u}_2^r) = t_2^M(\theta_1^r, \theta_2^r, u_1^r, u_2^r) & \text{if } \theta_2^r \geq \theta_2^* \end{cases}$$

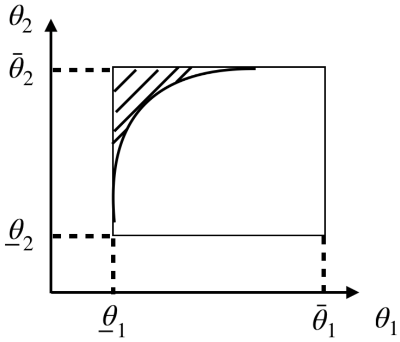


Figure 3:

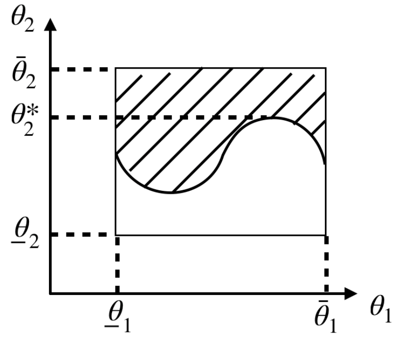


Figure 4

In words, a monotone two-stage mechanism has the property that, conditional on the trade occurring, the buyer's payment is strictly increasing in his own type if his type is smaller than θ_2^* and it is constant if his type is at least as high as θ_2^*.

We will show that in the example in Section 2.3, the generalized two-stage Groves mechanism is monotone.

Claim 7 *In the example in Section 2.3, the generalized two-stage Groves mechanism* (Θ, Π, x^*, t^G) *is monotone.*

Proof: The proof is in the Appendix. ■

This suggests that monotonicity is a mild requirement imposed on two-stage mechanisms. On the contrary, as we already argued, the NTNP is rather a stringent requirement.

2.4.2 The Proposed Two-Stage Mechanism

In this subsection, we propose a two-stage NTNP, monotone mechanism we use for our main result in the next subsection.

Recall that the efficient decision rule dictates that, for each $(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2$,

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

We propose the following two-stage mechanism (Θ, Π, x^*, t^M) which satisfies BIC, IIR, EFF, and BB. By construction, the proposed two-stage mechanism satisfies EFF. For each type report $(\theta_1^r, \theta_2^r) \in \Theta_1 \times \Theta_2$ and each payoff report $(u_1^r, u_2^r) \in \Pi_1 \times \Pi_2$,

$$t_1(\theta_1^r, \theta_2^r; u_1^r, u_2^r) = \begin{cases} \tilde{u}_2(\theta_1^r, \theta_2^r) & \text{if } \theta_2^r < \theta_2^*, x^*(\theta_1^r, \theta_2^r) = 1, \text{ and } u_2^r = u_2(x^*(\theta_1^r, \theta_2^r), \theta_1^r, \theta_2^r) \\ -g(\theta_1^r) & \text{if } \theta_2^r \geq \theta_2^*, x^*(\theta_1^r, \theta_2^r) = 1, \text{ and } u_2^r = u_2(x^*(\theta_1^r, \theta_2^r), \theta_1^r, \theta_2^r) \\ -\psi & \text{if } x^*(\theta_1^r, \theta_2^r) = 1 \text{ and } u_2^r \neq u_2(x^*(\theta_1^r, \theta_2^r), \theta_1^r, \theta_2^r) \\ 0 & \text{if } x^*(\theta_1^r, \theta_2^r) = 0, \end{cases}$$

and

$$t_2(\theta_1^r, \theta_2^r; u_1^r, u_2^r) = \begin{cases} -\tilde{u}_2(\theta_1^r, \theta_2^r) & \text{if } \theta_2^r < \theta_2^* \text{ and } x^*(\theta_1^r, \theta_2^r) = 1 \\ g(\theta_1^r) & \text{if } \theta_2^r \geq \theta_2^* \text{ and } x^*(\theta_1^r, \theta_2^r) = 1 \\ 0 & \text{if } x^*(\theta_1^r, \theta_2^r) = 0 \text{ and } u_1^r = u_1(x^*(\theta_1^r, \theta_2^r), \theta_1^r, \theta_2^r) \\ -\psi & \text{if } x^*(\theta_1^r, \theta_2^r) = 0 \text{ and } u_1^r \neq u_1(x^*(\theta_1^r, \theta_2^r), \theta_1^r, \theta_2^r), \end{cases}$$

where ψ is a strictly positive constant (which is determined later), $\theta_2^* \in (\underline{\theta}_2, \bar{\theta}_2]$ is the cutoff point identified in Lemma 8, and

$$g(\theta_1^r) = \begin{cases} -\tilde{u}_2(\theta_1^r, \theta_2^*) & \text{if } \theta_2^* = \bar{\theta}_2 \\ G(\theta_1^r) / (1 - F_2(\theta_2^*)) & \text{if } \theta_2^* \in (\underline{\theta}_2, \bar{\theta}_2) \end{cases}$$

with

$$\begin{aligned} G(\theta_1^r) \quad = \quad & \int_{\Theta_2^*(\theta_1^r) \setminus \Theta_2^{**}} \tilde{u}_2(\theta_1^r, \theta_2) dF_2(\theta_2) - \int_{\Theta_2^*(\theta_1^r)} \tilde{u}_1(\theta_1^r, \theta_2) dF_2(\theta_2) \\ & - \int_{\Theta_1} \int_{\Theta_2^*(\theta_1) \setminus \Theta_2^{**}} (\tilde{u}_2(\theta_1, \theta_2) - \tilde{u}_1(\theta_1, \theta_2)) \, dF_2(\theta_2) dF_1(\theta_1) \\ & - \int_{\Theta_1} \int_{\Theta_2^{**}} (\tilde{u}_2(\theta_1, \theta_2^*) - \tilde{u}_1(\theta_1, \theta_2)) \, dF_2(\theta_2) dF_1(\theta_1), \end{aligned} \tag{10}$$

where for each $\theta_1^r \in \Theta_1$,

$$\Theta_2^*(\theta_1^r) = \begin{cases} \{\bar{\theta}_2\} & \text{if } \{\theta_2 \in \Theta_2 : x^*(\theta_1^r, \theta_2) = 1\} = \emptyset \\ \{\theta_2 \in \Theta_2 : x^*(\theta_1^r, \theta_2) = 1\} & \text{otherwise,} \end{cases}$$

and $\Theta_2^{**} = [\theta_2^*, \bar{\theta}_2]$.

In this mechanism, if each agent i reports his true type θ_i and true allocation payoff $u_i = u_i(x^*(\theta_i, \theta_{-i}), \theta_i, \theta_{-i})$, then the following three properties are confirmed.

1. when $x^*(\theta_1, \theta_2) = 0$, $t_1(\theta_1, \theta_2; u_1, u_2) = t_2(\theta_1, \theta_2; u_1, u_2) = 0$, i.e., when no trade occurs, there are no monetary transfers. Hence, NTNP is satisfied;

2. when $x^*(\theta_1, \theta_2) = 1$ and $\theta_2 < \theta_2^*$, $t_1(\theta_1, \theta_2; u_1, u_2) = -t_2(\theta_1, \theta_2; u_1, u_2) = \tilde{u}_2(\theta_1, \theta_2)$, implying that the buyer's payment is strictly increasing in his type; and

3. when $x^*(\theta_1, \theta_2) = 1$ and $\theta_2 \geq \theta_2^*$, $t_1(\theta_1, \theta_2; u_1, u_2) = -t_2(\theta_1, \theta_2; u_1, u_2) = -g(\theta_1)$ which is independent of the buyer's type.

By construction, the proposed two-stage mechanism is monotone. It also satisfies BB by construction. By contrast, we need break the budget off the equilibrium. If it is efficient not to trade and the seller's payoff report in the second stage is inconsistent with the type reports in the first stage, then the buyer is punished with a penalty ψ. Similarly, if it is efficient to trade and the buyer's payoff report in the second stage is inconsistent with the type reports in the first stage, then the seller is punished with a penalty ψ.

Remark: Our proposed two-stage mechanism can be considered a generalization of the "shoot-the-liar" mechanism in Mezzetti (2007) in an auction setup. We highlight the following two aspects. First, in the "shoot-the-liar" mechanism, the seller plays a 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 our mechanism, however, the seller has private information which should be elicited within the mechanism and he is asked to make monetary transfers based on the reports. Second, in the "shoot-the-liar" mechanism, the seller always extracts the full surplus. On the contrary, the payment rule in our mechanism varies over the buyer's types. Even though the payment rule below the cutoff θ_2^* shares the same spirit as the "shoot-the-liar" mechanism, the payment above the cutoff θ_2^* is different from it in the sense that the buyer is left with positive expected surplus.[19]

[19] We will elaborate on the second difference between ours and the "shoot-the-liar" mechanism in Section 4.4.

2.4.3 A Sufficient Condition for A Possibility Result

To state our main result, we introduce the following condition.

Definition 16 *An environment satisfies Condition α if,*

$$\int_{\Theta_1}\int_{\Theta_2^*(\theta_1)\setminus\Theta_2^{**}} \left(\tilde{u}_2(\theta_1,\theta_2)-\tilde{u}_1(\theta_1,\theta_2)\right) dF_2(\theta_2)dF_1(\theta_1)$$
$$+\int_{\Theta_1}\int_{\Theta_2^{**}} \left(\tilde{u}_2(\theta_1,\theta_2^*)-\tilde{u}_1(\theta_1,\theta_2)\right) dF_2(\theta_2)dF_1(\theta_1) \geqslant 0, \quad (11)$$

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

$$\Theta_2^*(\theta_1)=\begin{cases}\{\bar{\theta}_2\} & \text{if } \{\theta_2\in\Theta_2 : x^*(\theta_1,\theta_2)=1\}=\emptyset\\ \{\theta_2\in\Theta_2 : x^*(\theta_1,\theta_2)=1\} & \text{otherwise,}\end{cases}$$

$\Theta_2^{**}=[\theta_2^*,\bar{\theta}_2]$, *and* $\theta_2^*\in(\underline{\theta}_2,\bar{\theta}_2]$ *is the cutoff point identified in Lemma 8.*

Remark: If $\Theta_2^*(\theta_1)=\{\bar{\theta}_2\}$ for some $\theta_1\in\Theta_1$, then $\Theta_2^*(\theta_1)\setminus\Theta_2^{**}$ is an empty set. In this case, any integration over $\Theta_2^*(\theta_1)\setminus\Theta_2^{**}$ is always zero. Since the first term in the left-hand-side of inequality (11) corresponds to the ex ante gains from trade over $[\underline{\theta}_1,\bar{\theta}_1]\times[\underline{\theta}_2,\theta_2^*]$, it is always nonnegative. If $\theta_2^*=\bar{\theta}_2$, the second term in inequality (11) is zero by definition. Therefore, Condition α is automatically satisfied when $\theta_2^*=\bar{\theta}_2$.

To further illustrate this condition, we first consider Case (i) $0<\gamma_2\leqslant\gamma_1<1$ in the example of Section 2.3. In this case, we have $\theta_2^*=\bar{\theta}_2=1$. Then, we obtain

$$\Theta_2^*(\theta_1)=\begin{cases}\left[\frac{1-\gamma_2}{1-\gamma_1}\theta_1,1\right] & \text{if } 0<\theta_1<(1-\gamma_1)/(1-\gamma_2)\\ \{1\} & \text{if } (1-\gamma_1)/(1-\gamma_2)\leqslant\theta_1<1.\end{cases}$$

We next consider Case (ii) $0<\gamma_1<\gamma_2<1$. In this case, we obtain $\theta_2^*=(1-\gamma_2)/(1-\gamma_1)<1=\bar{\theta}_2$ and $\Theta_2^*(\theta_1)=[(1-\gamma_2)\theta_1/(1-\gamma_1),1]$ for any $\theta_1\in[0,1]$.

We describe the logic behind why Condition α is needed for the proposed two-stage mechanism to satisfy all the desired properties. First, we let the buyer pay an amount equal to his "reported" valuation when his type report is below the cutoff θ_2^*. Next, we solve the appropriate payment function above the cutoff which satisfies BIC and IIR. It turns out that we can find an upper bound and lower bound on the buyer's payment function above the cutoff θ_2^*. Specifically, the upper bound comes from the seller's IIR constraints and the lower bound comes from the buyer's IIR constraints. Condition α plays a role of ensuring that the upper and lower bound are compatible with each other in the two-stage mechanism constructed in Subsection 2.4.2.

As we mentioned in the above remark, Condition α is automatically satisfied in Case (i) $0 < \gamma_2 \leq \gamma_1 < 1$ in the example of Section 2.3. To check when Condition α is satisfied even in Case (ii) of the example of Section 2.3, we are going to use the following result.

Lemma 9 *In the example of Section 2.3, our Condition α is reduced to*

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

Proof: The proof is in the Appendix. ■

The lemma below shows that Condition α sometimes holds in Case (ii) of the example in Section 2.3.

Lemma 10 *Suppose that in the example in Section 2.3, both agents' valuation functions are $\tilde{u}_1(\theta_1, \theta_2) = \theta_1 + \theta_2/3$ and $\tilde{u}_2(\theta_1, \theta_2) = \theta_2 + \theta_1/2$. That is, $0 < 1/3 = \gamma_1 < 1/2 = \gamma_2 < 1$. Then, Condition α holds.*

Proof: Plugging $\gamma_1 = 1/3$ and $\gamma_2 = 1/2$ into the inequality in Lemma 9, we obtain

$$\frac{1}{6}\frac{(1/2)^2}{2/3} + \frac{1/2}{2/3} - \frac{1}{2}\left(\frac{1/2}{2/3}\right)^2 + \frac{1}{2}(1/2 - 1/3 - 1) \quad = \quad \frac{11}{96} > 0.$$

Thus, Condition α is satisfied. ■

We can also show in the lemma below that Condition α is sometimes violated in Case (ii) of the example of Section 2.3.

Lemma 11 *Suppose that in the example in Section 2.3, both agents' valuation functions are $\tilde{u}_1(\theta_1, \theta_2) = \theta_1 + \theta_1/2$ and $\tilde{u}_2(\theta_1, \theta_2) = \theta_2 + 4\theta_1/5$. That is, $0 < 1/2 = \gamma_1 < 4/5 = \gamma_2 < 1$. Then, Condition α fails.*

Proof: Plugging $\gamma_1 = 1/2$ and $\gamma_2 = 4/5$ into the inequality in Lemma 9 so that we obtain

$$\frac{1}{6}\frac{0.2^2}{0.5} + \frac{0.2}{0.5} - \frac{1}{2}\left(\frac{0.2}{0.5}\right)^2 + \frac{1}{2}(0.8 - 0.5 - 1) \quad = \quad -\frac{1}{60} < 0.$$

Thus, Condition α is violated. ■

2.4.4 The Theorem

Using the two-stage NTNP, monotone mechanism proposed in Subsection 2.4.2, we are able to establish the main result of the paper.

Theorem 5 *Suppose that an environment satisfies Condition α. Then, there exists a two-stage NTNP, monotone mechanism (Θ, Π, x^*, t^M) satisfying BIC, IIR, EFF, and BB.*

Proof: We make use of the two-stage mechanism constructed in Subsection 2.4.2.

Since the seller's transfer t_1^M is independent of his payoff report u_1^r and the buyer's transfer t_2^M is independent of u_2^r, then each agent has no incentive to deviate from the truth-telling in their payoff report in the second stage. Given this, it remains to verify that the truth-telling in the first stage constitutes part of a perfect Bayesian equilibrium (Steps 1 and 2) and that IIR is satisfied for both agents (Step 3). The proof is completed by the following three steps.

Step 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.

Proof: The proof is in the Appendix. ■

Step 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.

Proof: The proof is in the Appendix. ■

In Steps 1 and 2, we show that the constructed two-stage mechanism (Θ, Π, x^*, t^M) satisfies BIC.

Step 3: The two-stage mechanism (Θ, Π, x^*, t^M) also satisfies IIR.

Proof: The proof is in the Appendix. ■

Taking into account that both EFF and BB are already built in the mechanism, we complete the proof of Theorem 5. ■

We record the implications of Theorem 5 as well as the properties of the proposed two-stage mechanism in the context of the example in Section 2.3.

1. When $0 < \gamma_2 \leq \gamma_1 < 1$, we have $\theta_2^* = 1 = \bar{\theta}_2$. In this case, we get $g(\theta_1^r) = -\tilde{u}_2(\theta_1^r, \bar{\theta}_2)$. Recall $t_2(\theta_1^r, \theta_2^r; u_1^r, u_2^r) = -\tilde{u}_2(\theta_1^r, \theta_2^r)$ when $\theta_2^r < \theta_2^*$ and $x^*(\theta_1^r, \theta_2^r) = 1$. If both agents report truthfully in both stages, the buyer always pays an amount equal to his true valuation to the seller. In other words, the seller extracts the full surplus in this case.

2. When $0 < 1/3 = \gamma_1 < 1/2 = \gamma_2 < 1$, we have $\theta_2^* = 3/4$. In this case, we set $g(\theta_1^r) = 3(\theta_1^r)^2/4 - 5\theta_1^r/2$. If both agents report truthfully in both stages and the buyer's true type is $\theta_2 > \theta_2^*$, the buyer's ex post utility becomes
$$\tilde{u}_2(\theta_1, \theta_2) + g(\theta_1) = \theta_2 + \frac{1}{2}\theta_1 + \frac{3}{4}(\theta_1)^2 - \frac{5}{2}\theta_1 = \theta_2 - \frac{4}{3} + \frac{3}{4}\left(\theta_1 - \frac{4}{3}\right)^2.$$
To further illustrate the properties of the proposed two-stage mechanism when $\gamma_1 = 1/3$ and $\gamma_2 = 1/2$, we consider the following subcases:

   (a) when $\theta_1 = 0$, we have $g(\theta_1) = 0$. This means that the buyer receives the good without making any payment. Hence, the buyer receives the full surplus.

   (b) when $\theta_1 = 1$, we have that $\tilde{u}_2(\theta_1, \theta_2) + g(\theta_1) = \theta_2 - 5/4 < 0$, implying that the buyer's ex post utility is always negative because $\theta_2 \leq 1$. Thus, the ex post individual rationality (EPIR) is violated. Nonetheless, since our Condition α holds, the proposed two-stage mechanism satisfies IIR (as opposed to EPIR) together with BIC, EFF, and BB. This exhibits a contrast with the analysis of GMO (2011) which maintains EPIR throughout.

   Moreover, from Step 3 in the proof of Theorem 5 (see the Appendix (Section 8.8) for the details), we know that if $\theta_2 \geqslant \theta_2^*$, the expected utility of the buyer of type θ_2 after participation is
$$\int_{\Theta_1} \tilde{u}_2(\theta_1, \theta_2) dF_1(\theta_1) + \int_{\Theta_1} g(\theta_1) dF_1(\theta_1) = \int_{\Theta_1} \tilde{u}_2(\theta_1, \theta_2) dF_1(\theta_1) - \int_{\Theta_1} \tilde{u}_2(\theta_1, \theta_2^*) dF_1(\theta_1) \geqslant 0,$$
where the weak inequality follows because $\tilde{u}_2(\cdot)$ is strictly increasing in θ_2. Therefore, if $\theta_2 \leqslant \theta_2^*$, the buyer of type θ_2 is always left with zero expected surplus; if $\theta_2 > \theta_2^*$, the buyer receives a positive expected surplus.

These features we described above distinguishes our proposed two-stage mechanism from the "shoot-the-liar" mechanism proposed by Mezzetti (2007) in which the seller always extracts the full surplus. By contrast, GMO (2011, Section 5) apply the "shoot-the-liar mechanism" without modifications to their partnership dissolution problem.

We know that Condition α is a necessary and sufficient condition for our proposed two-stage mechanism to satisfy BIC, IIR, EFF, and BB. However, this does not exclude a possibility that there might be some other two-stage NTNP monotone mechanism satisfying BIC, IIR, EFF, and BB, even when Condition α is violated. Although it would be interesting to explore such a possibility, we leave this for future work.

2.5 Simulation

To assess the permissiveness and restrictiveness of Condition α, we provide a set of simulation results based on the example in Section 2.3. Both agents' types are uniformly distributed on the unit interval $[0,1]$ and for each type profile $(\theta_1, \theta_2) \in [0,1]^2$, their valuation functions are $\tilde{u}_1(\theta_1,\theta_2) = \theta_1 + \gamma_1\theta_2$ and $\tilde{u}_2(\theta_1,\theta_2) = \theta_2 + \gamma_2\theta_1$ where $\gamma_1 \in \{0.01, 0.02, \cdots, 0.98\}$ and $\gamma_2 \in \{\gamma_1 + 0.01, \gamma_1 + 0.02, \cdots, 0.99\}$ for each γ_1. As we discuss in the previous section, Condition α is always satisfied when $0 < \gamma_2 \leq \gamma_1 < 1$, which is called Case (i) in the example of Section 2.3. Then, by our Theorem 5, we know that there exists a two-stage NTNP, monotone mechanism satisfying BIC, EFF, BB, and IIR. Thus, what remains to investigate is the extent to which there exists a two-stage NTNP, monotone mechanism satisfying all the desired properties in Case (ii) $0 < \gamma_1 < \gamma_2 < 1$. In the simulation, we select finitely many values of γ_1 and γ_2 satisfying this inequality.

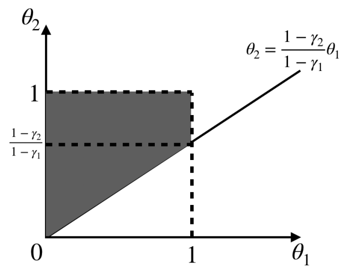
We note that $\tilde{u}_2(\theta_1,\theta_2) > \tilde{u}_1(\theta_1,\theta_2)$ if and only if $\theta_2 > (1-\gamma_2)\theta_1/(1-\gamma_1)$. Since we assume $\gamma_2 > \gamma_1$, the slope of the efficient frontier is $(1-\gamma_2)/(1-\gamma_1) < 1$. The efficient decision rule dictates that, for each $(\theta_1, \theta_2) \in [0,1]^2$,

$$x^*(\theta_1,\theta_2) = \begin{cases} 1 & \text{if } \theta_2 > (1-\gamma_2)\theta_1/(1-\gamma_1) \\ 0 & \text{otherwise.} \end{cases}$$

The shaded region in Figure 2 (which is reproduced below) represents $\Theta^* = \{(\theta_1, \theta_2) \in$

$\Theta_1 \times \Theta_2 : x^*(\theta_1, \theta_2) = 1\}$, which describes the set of possible type profiles for which it is efficient to trade.

Figure 2: when $0 < \gamma_1 < \gamma_2 < 1$



Recall that Lemma 9 allows us to translate our Condition α into the following inequality:

$$\frac{1}{6}\frac{(1-\gamma_2)^2}{1-\gamma_1} + \frac{1-\gamma_2}{1-\gamma_1} - \frac{1}{2}\left(\frac{1-\gamma_2}{1-\gamma_1}\right)^2 + \frac{1}{2}(\gamma_2 - \gamma_1 - 1) \geq 0. \tag{12}$$

Observe that Condition α becomes an inequality about γ_1 and γ_2. Then, for each pair (γ_1, γ_2) satisfying $0 < \gamma_1 < \gamma_2 < 1$, we check whether or not inequality (12) is satisfied. Here is a summary of the simulation results. There are two possible scenarios:

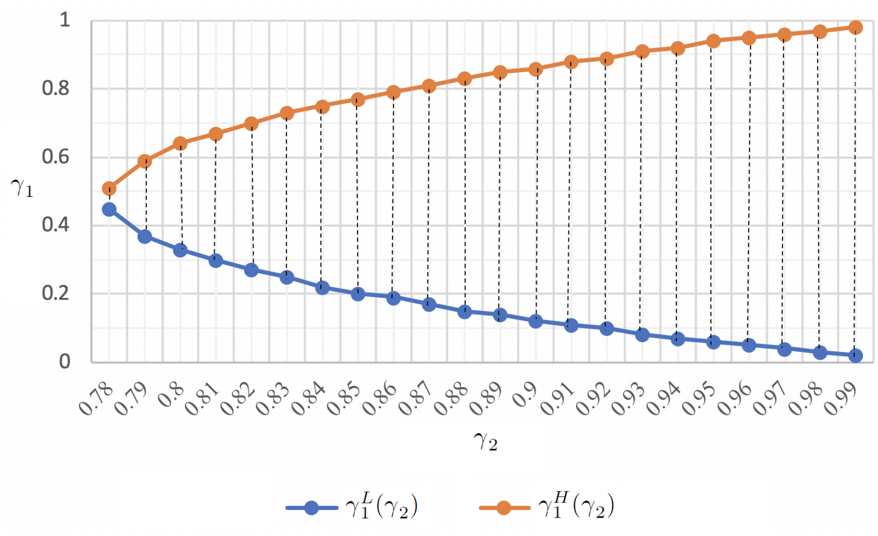
1. If $\gamma_2 \leq 0.77$, inequality (12) is always satisfied for all $\gamma_1, \gamma_2 \in (0, 1)$ satisfying $\gamma_1 < \gamma_2$;

2. For each $\gamma_2 > 0.77$, there exist $\gamma_1^L(\gamma_2), \gamma_1^H(\gamma_2) \in (0, 1)$ such that $\gamma_1^L(\gamma_2) < \gamma_1^H(\gamma_2)$ and inequality (12) is violated whenever $\gamma_1^L(\gamma_2) < \gamma_1 < \gamma_1^H(\gamma_2)$ and it is satisfied otherwise.

We illustrate the second scenario in Figure 6 below. For each $\gamma_2 > 0.77$, there are a corresponding point on the upper curve indicating $\gamma_1^H(\gamma_2)$ and another corresponding point on the lower curve indicating $\gamma_1^L(\gamma_2)$. Then, if $\gamma_1^L(\gamma_2) < \gamma_1 < \gamma_1^H(\gamma_2)$, inequality (12) is violated. The region where inequality (12) is violated is represented by the dotted region in Figure 6. The region outside the dotted region dictates the case in which inequality (12) is satisfied.

In Figure 7, we track all possible pairs of $(\gamma_1, \gamma_2) \in (0, 1)^2$ satisfying inequality (12). In particular, the upper triangle in $[0, 1]^2$, i.e., the region where $\gamma_2 > \gamma_1$ corresponds to

Figure 6: When $\gamma_2 > 0.77$



Case (ii) of the example in Section 2.3. The lightly shaded region describes all pairs of (γ_1, γ_2) within this upper triangle for which our Condition α is satisfied.

On the other hand, the lower triangle in the unit square, i.e., the region where $\gamma_2 < \gamma_1$ corresponds to Case (i) of the example in Section 2.3. Then, by our Theorem 5, we can always find a two-stage NTNP, monotone mechanism satisfying BIC, IIR, EFF, and BB within this region. Hence, the heavily shaded region describes all pairs of (γ_1, γ_2) within the lower triangle for which our Condition α is satisfied.

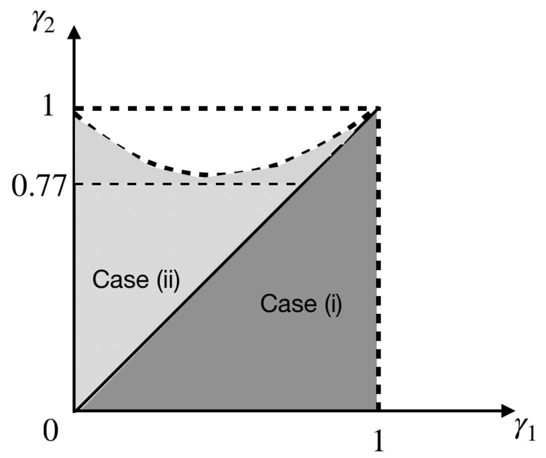
Therefore, the lightly and heavily shaded regions together indicate the set of (γ_1, γ_2) for which our Condition α is satisfied. Since the whole shaded (regardless of whether lightly or heavily) region spans quite a large part of the unit square, we conclude that our Condition α can be satisfied in many cases in the example of Section 2.3.

We can also verify that if Condition α is violated, then the two-stage mechanism we propose in Section 2.4.2 violates the seller's IIR constraint.[20] We make this point by the following claim:

Claim 8 *If* $\gamma_1 = 1/2$ *and* $\gamma_2 = 4/5$ *in the example of Section 2.3, the seller's IIR constraint is violated in our two-stage NTNP, monotone mechanism constructed in Subsection*

[20] By the very proof of Theorem 5, Condition α has a bite exactly when the seller's IIR constraint has a bite. In other words, if inequality (11) in Condition α is violated, it is the seller's IIR constraint that is violated.

Figure 7: Summary of Simulation



2.4.2.

Proof: The proof is in the Appendix. ■

By this claim, we loosely say that our Condition α is violated when the degree of interdependence of preferences of the buyer is too high relative to that of the seller.

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

In this section, we will discuss the relation between this paper and Galavotti, Muto, and Oyama (2011) (hereafter, GMO). GMO (2011) consider the problem of partnership dissolution in a model with interdependent values where there are one asset, and n risk-neutral agents indexed by $i \in \{1, \ldots, n\}$ where $n \geq 2$. Each agent i owns a share α_i of the asset such that $0 \leq \alpha_i \leq 1$ and $\sum_{i=1}^{n} \alpha_i = 1$. In private values environments, Cramton, Gibbons, and Klemperer (1987) show that the equal-share ownership $(\alpha_1, \ldots, \alpha_n) = (1/n, \ldots, 1/n)$ allows us to construct a mechanism satisfying BIC, EFF, IIR, and BB, which exhibits a contrast with this paper's extreme ownership structure where the seller has the full property right over the good. However, FKM (2003) show that this positive result of Cramton, Gibbons, and Klemperer (1987) cannot be extended to a model with interdependent values. This explains why GMO (2011, Section 5) also resort to the use of two-stage mechanisms in order to obtain more positive results.

To make the comparison between GMO (2011) and our paper, we assume that there are only two agents, i.e., $n = 2$. By an *ownership structure* (α_1, α_2) where each $\alpha_i \in [0, 1]$ and $\alpha_1 + \alpha_2 = 1$, we mean that agent 1 (the seller) has the property right over α_1 fraction of the asset and agent 2 (the buyer) has the property right over α_2 fraction of the asset. To discuss the contribution of GMO (2001), we first strengthen our IIR constraint into its ex post counterpart.

Definition 17 *Let (α_1, α_2) be an ownership structure. A two-stage mechanism (Θ, Π, x, t) satisfies ex post individual rationality (EPIR) if, for all $(\theta_1, \theta_2) \in \Theta$ and $(u_1, u_2) \in \Pi$,*

$$u_1(x(\theta_1, \theta_2), \theta_1, \theta_2) + t_1(\theta_1, \theta_2; u_1, u_2) \geq \alpha_1 \tilde{u}_1(\theta_1, \theta_2),$$

and

$$u_2(x(\theta_1, \theta_2), \theta_1, \theta_2) + t_2(\theta_1, \theta_2; u_1, u_2) \geq \alpha_2 \tilde{u}_2(\theta_1, \theta_2),$$

where $u_1 = u_1(x^*(\theta_1, \theta_2), \theta_1, \theta_2)$ and $u_2 = u_2(x^*(\theta_1, \theta_2), \theta_1, \theta_2)$.

GMO (2011) provide the following sufficient condition (called Assumption 5.1 on p.14) under which the "shoot-the-liar" mechanism satisfies BIC, EPIR, EFF, and BB for any ownership structure. We formally state GMO's Assumption 5.1.

GMO's Assumption 5.1: There exist $M_1, M_2 \geq 0$ such that for all $i \in \{1, \ldots, n\}$, all $\theta_i, \hat{\theta}_i \in \Theta_i$ with $\hat{\theta}_i \neq \theta_i$,

$$\begin{aligned} & \mathbb{E}_{\theta_{-i}} \left[\mathbb{1}_{\{\theta_{-i} | i = m(\hat{\theta}_i, \theta_{-i})\}}(\theta_{-i}) \left(\tilde{u}_i(\bar{\theta}_i, \theta_{-i}) - \tilde{u}_i(\hat{\theta}_i, \theta_{-i}) \right) \right] \\ \leqslant \; & M_1 \sum_{j \neq i} \mathbb{E}_{\theta_{-i}} \left[\mathbb{1}_{\{\theta_{-i} | j = m(\hat{\theta}_i, \theta_{-i}), \tilde{u}_j(\theta_i, \theta_{-i}) \neq \tilde{u}_j(\hat{\theta}_i, \theta_{-i})\}}(\theta_{-i}) \right], \end{aligned} \tag{13}$$

and

$$\begin{aligned} & \sum_{j \neq i} \mathbb{E}_{\theta_{-i}} \left[\mathbb{1}_{\{\theta_{-i} | j = m(\hat{\theta}_i, \theta_{-i}), \tilde{u}_j(\theta_i, \theta_{-i}) = \tilde{u}_j(\hat{\theta}_i, \theta_{-i})\}}(\theta_{-i}) \right] \\ \leqslant \; & M_2 \sum_{j \neq i} \mathbb{E}_{\theta_{-i}} \left[\mathbb{1}_{\{\theta_{-i} | j = m(\hat{\theta}_i, \theta_{-i}), \tilde{u}_j(\theta_i, \theta_{-i}) \neq \tilde{u}_j(\hat{\theta}_i, \theta_{-i})\}}(\theta_{-i}) \right], \end{aligned} \tag{14}$$

where $\mathbb{1}_X(x)$ is the index function such that $\mathbb{1}_X(x) = 1$ if $x \in X$ and 0 if $x \notin X$, and $m(\theta) = \max(\arg\max_j \tilde{u}_j(\theta))$.

In our bilateral trade setup, we always have $(\alpha_1, \alpha_2) = (1, 0)$, i.e., the seller has the property right over the good, while the buyer has no property right over it. We know from our Lemma 8 that there are generally two cases: (i) $\theta_2^* = \bar{\theta}_2$ and (ii) $\theta_2^* \in (\underline{\theta}_2, \bar{\theta}_2)$ where θ_2^* is the cutoff point identified in Lemma 8. In Case (i) $\theta_2^* = \bar{\theta}_2$, which corresponds to the case that $\int_{\Theta_1} x^*(\theta_1, \theta_2) dF_1(\theta_1) < 1$ for all $\theta_2 < \bar{\theta}_2$, we can use our proposed two-stage mechanism and show that it satisfies BIC, IIR, EFF, and BB. As in GMO (2011), we can strengthen IIR into EPIR for this result.

In what follows, we will focus on the bilateral trade model and then compare GMO's Assumption 5.1 with our Condition α. We obtain the following claim:

Lemma 12 *The relation between Assumption 5.1 in GMO (2011) and our Condition α is summarized as follows:*

1. *Inequality (13) in GMO's Assumption 5.1 implies our Condition α;*
2. *Inequality (14) in GMO's Assumption 5.1 is automatically satisfied under the bilateral trade model in our paper.*

Proof: The proof is in the Appendix. ■

Intuitively, inequality (13) requires that each agent's deviation be detected by the other agent with strictly positive probability, Case (i) $\theta_2^* = \bar{\theta}_2$, i.e., $\int_{\Theta_1} x^*(\theta_1, \theta_2) dF_1(\theta_1) < 1$ for all $\theta_2 < \bar{\theta}_2$ requires that only the buyer's deviation be detected by the seller with strictly positive probability. Therefore, the condition that $\int_{\Theta_1} x^*(\theta_1, \theta_2) dF_1(\theta_1) < 1$ for all $\theta_2 < \bar{\theta}_2$ is weaker than inequality (13).[21]

To further illustrate the stringent nature of inequality (13) relative to our Condition α, we revisit the example in Section 2.3. We obtain the following lemma:

Lemma 13 *In the example in Section 2.3, GMO's Assumption 5.1 is satisfied if and only if $\gamma_1 = \gamma_2$.*

[21]The logic behind our first general case is that even if the seller's deviation is not detected by the buyer, this is not a profitable deviation because in this case, the seller keeps the good without receiving any monetary transfer.

Proof: The proof is in the Appendix. ■

This suggests that GMO's Assumption 5.1 is generically violated in the bilateral trade model.

2.7 Conclusion

This paper characterizes when efficient, voluntary bilateral trades are incentive compatible in an environment with interdependent values. 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). We show by means of an example that the generalized two-stage Groves mechanism never satisfies IIR. If our Condition α is satisfied in a general environment, we next show that there exists a two-stage mechanism satisfying BIC, IIR, EFF, and BB. In the context of the example in Section 2.3, our Condition α roughly says that the buyer's degree of interdependence of preferences is not too high relative to the seller's counterpart. In Section 2.5, we also argue by the same example that our Condition α can be satisfied for a large number of cases. The property that distinguishes our proposed two-stage mechanism from the generalized two-stage Groves mechanism is the "no-trade-then-no-payments" (NTNP) property, which means that if trade does not occur, no payments are made. Indeed, the generalized two-stage Groves mechanism does not satisfy the NTNP property. By expanding our scope into two-stage mechanisms, we consider our paper as the first attempt to further pushing the boundary between when efficient, voluntary bilateral trades are implementable and when they are not.

3 The Interplay of Interdependence and Correlation in Bilateral Trade

3.1 Introduction

This paper revisits the bilateral trade problem by focusing on a model with a discrete type space in which valuations are interdependent as well as signals are correlated. The primary objective of this paper is to investigate when there are mechanisms satisfying Bayesian incentive compatibility (BIC), interim individually rationality (IIR), decision efficiency (EFF), and ex post budget balance (BB) in such an economy.

Myerson and Satterthwaite (1983) established the following celebrated impossibility result: in a bilateral trade model with private values (i.e., each agent is certain of the value of the object at the timing of trade) and independently distributed continuous types, there are generally no mechanisms satisfying BIC, IIR, EFF, and BB. This paper sheds light on three features of the setup of Myerson and Satterthwaite (1983): (i) private values; (ii) independently distributed signals; and (iii) continuous types.

When extending a model of private values to that of interdependent values but keeping (ii) and (iii), Fieseler, Kittsteiner, and Moldovanu (2003) and Gresik (1991b) restore the Myerson and Satterthwaite impossibility theorem. This leads us to relax either (ii), (iii), or both, if we are to obtain a possibility result in an interdependent values environment. Indeed, this paper relaxes both (ii) and (iii) to obtain positive results.

In contrast to the agents' continuum type space used in many papers in the literature, Matsuo (1989) keeps (i) and (ii) but studies a two-type bilateral trade environment and characterize the set of valuations and independent beliefs that guarantees the existence of mechanisms satisfying the aforementioned four properties. Roughly speaking, he shows that such a desirable mechanism exists if and only if the probability that a mutually beneficial trade exists is small. Like Matsuo (1989), Gresik (1991a) also studies a two-type bilateral trade model and maintain (i) but unlike Matsuo (1989), he allows for correlated signals. In such an environment, Gresik (1991a) also characterizes the set of correlated beliefs and valuations that guarantee the existence of desirable mechanisms. With correlated beliefs, such a desirable mechanism can exist even when the probability that a mutually beneficial trade exists is large. The objective of this paper is to generalize the analysis of Matsuo (1989) and Gresik (1991a) to a general finite type space with interdependent valuations as well as correlated signals.

We first identify the set of interdependent valuations and correlated beliefs that constitutes a necessary condition for the existence of mechanisms satisfying BIC, IIR, EFF, and BB in a bilateral trade model with a finite discrete type space. Second, we show that the identified necessary condition becomes also sufficient when we consider a two-type model. Third, we apply our characterization to a two-type model with linear valuation functions and symmetric joint distributions. We call the extent to which the buyer's valuation depends on the seller's type "the buyer's degree of interdependence", and the extent to which the seller's valuation depends on the buyer's type "seller's degree of interdependence". We find that

1. For any fixed degree of correlation, the upper bound of the seller's degree of interdependence allowing desirable mechanisms is increasing in the buyer's degree of interdependence;

2. If we fix buyer's degree of interdependence, the upper bound of the seller's degree of interdependence allowing desirable mechanisms is increasing in the degree of correlation.

Finally, we drop IIR as part of the requirement and establish the general existence of mechanisms satisfying BIC, EFF, and BB in a two-type model. We obtain this result using Farkas lemma. This suggests that IIR is the most stringent constraint for the existence of mechanisms satisfying BIC, IIR, EFF, and BB.

The rest of the paper is organized as follows. In Section 3.2, we introduce the general notation and basic concepts for the paper. Section 3.3 deals with a general finite type space and identifies a necessary condition for the existence of mechanisms satisfying BIC, IIR, EFF, and BB. In Section 3.4, we confine our attention to a two-type model. This section constitutes two subsections. In Subsection 3.4.1, we show that the identified necessary condition in the previous section also becomes sufficient. In Subsection 3.4.2, we conduct the comparative statics for how the interdependence of valuations and correlated signals interact with each other in order to guarantee the existence of desirable mechanisms. In Section 3.5, we still consider a two-type model as in Section 3.4 but drop IIR as part of the requirement for desirable mechanisms. We establish the general existence of mechanisms satisfying BIC, IIR, and BB by using Farkas lemma. Section 3.6 concludes the paper by suggesting a few open questions which remain unresolved. In the Appendix, we provide all the proofs of the results omitted from the main text of the paper.

3.2 Preliminaries

There are one buyer, one seller, and one indivisible good owned by the seller. They decide whether to trade or not and how much each agent pays or receives. Each agent has some private information (which we call type) concerning the value of the good. In this paper, we assume that each agent has m types where $m \geq 2$. The buyer's type space is $B = \{B_1, B_2, \cdots, B_m\}$ where $B_1 < B_2 < \cdots < B_m$; similarly, the seller's type space is $S = \{S_1, S_2, \cdots, S_m\}$ where $S_1 < S_2 < \cdots < S_m$.

We assume that the agents' types are drawn from the joint probability function $g(\cdot, \cdot)$ so that $g(B_i, S_j) > 0$ for each $(B_i, S_j) \in B \times S$ and $\sum_{i=1}^{m} \sum_{j=1}^{m} g(B_i, S_j) = 1$. We define the following notation for conditional probabilities: for each $B_i \in B$ and each $S_j \in S$,

$$g_b(B_i|S_j) = \frac{g(B_i, S_j)}{\sum_{k=1}^{m} g(B_k, S_j)};$$

$$g_s(S_j|B_i) = \frac{g(B_i, S_j)}{\sum_{l=1}^{m} g(B_i, S_l)}.$$

We assume that each agent's valuation for the good depends not only on his own types, but also on the other's types. That is, given a type profile (B_i, S_j), $v_b(B_i, S_j)$ denotes the buyer's valuation of the good and $v_s(B_i, S_j)$ denotes the valuation of the good. We assume that $v_b(B_i, S_j)$ is strictly increasing in B_i and nondecreasing in S_j; similarly, $v_s(B_i, S_j)$ is strictly increasing in S_j and nondecreasing in B_i.

Definition 18 *The valuation functions satisfy the single crossing condition if,*

1. *For each S_j, if there exists some B_i such that $v_b(B_i, S_j) > v_s(B_i, S_j)$, then $v_b(B_k, S_j) > v_s(B_k, S_j)$ for any $k \in \{i+1, \cdots, m\}$.*

2. *For each B_i, if there exists some S_j such that $v_s(B_i, S_j) > v_b(B_i, S_j)$, then $v_s(B_i, S_l) > v_s(B_i, S_l)$ for any $l \in \{j+1, \cdots, m\}$.*

Roughly speaking, this condition means that if the buyer of some type has a higher valuation than the seller, the buyer of a higher type continues to have a higher valuation than the seller. Similar comment applies to the seller. Throughout the paper, we impose the single crossing condition.

If trade occurs with probability $q \in [0,1]$ and the buyers' monetary payment t_b and the seller's t_s, and the buyer is of type B_i and the seller is of type S_j, then the buyer's state-dependent preferences $u_b : [0,1] \times \mathbb{R}^2 \times B \times S \to \mathbb{R}$ can be represented by

$$u_b(q, t_b, t_s; B_i, S_j) = q \cdot v_b(B_i, S_j) - t_b,$$

and the seller's state-dependent preference $u_s : [0,1] \times \mathbb{R}^2 \to \mathbb{R}$ is represented by

$$u_s(q, t_b, t_s; B_i, S_j) = (1-q) v_s(B_i, S_j) - t_s.$$

A *direct mechanism* is defined as a triplet $\Gamma = ((B \times S), x, t)$ where $B = \{B_1, B_2, \cdots, B_m\}$ and $S = \{S_1, S_2, \cdots, S_m\}$ are the set of actions available to the buyer and seller, respectively; $x : B \times S \to [0,1]$ is the *decision rule* which specifies the probability that trade occurs; and $t : B \times S \to \mathbb{R}^2$ is the *transfer rule* which describes the monetary payments for both agents. We further impose the following property on the class of direct mechanisms:

Definition 19 *A mechanism (x, t) satisfies the no-trade-then-no-payments (NTNP) property if, for any $(B_i, S_j) \in B \times S$,*

$$x(B_i, S_j) = 0 \Rightarrow t_b(B_i, S_j) = t_s(B_i, S_j) = 0.$$

This property says that if trade does not occur, no monetary transfers are made.[22] In what follows, we call such a mechanism satisfying this property a no-trade-then-no-payments (NTNP) mechanism. In what follows, we call a direct mechanism satisfying NTNP property simply a mechanism and we denote it by (x, t).

[22] Theorem 1 of Gresik (1991) also imposes this property.

Definition 20 *A mechanism* (x, t) *satisfies Bayesian incentive compatibility (BIC) if, for all* $i, j \in \{1, \ldots, m\}$ *with* $i \neq k$*,*

$$\sum_{j=1}^{m} g_s(S_j|B_i)\left[x(B_i, S_j)v_b(B_i, S_j) - t_b(B_i, S_j)\right] \geqslant \sum_{j=1}^{m} g_s(S_j|B_i)\left[x(B_k, S_j)v_b(B_i, S_j) - t_b(B_k, S_j)\right],$$

and for all $j, l \in \{1, \ldots, m\}$ *with* $j \neq l$*,*

$$\begin{aligned} & \sum_{i=1}^{m} g_b(B_i|S_j)\left[(1 - x(B_i, S_j))\, v_s(B_i, S_j) - t_s(B_i, S_j)\right] \\ \geqslant & \sum_{i=1}^{m} g_b(B_i|S_j)\left[(1 - x(B_i, S_l))\, v_s(B_i, S_j) - t_s(B_i, S_k)\right]. \end{aligned}$$

We also assume that each agent has the option of not participating in the mechanism (x, t). In particular, the seller's outside option utility is $\sum_{i=1}^{m} g_b(B_i|S_j)v_s(B_i, S_j)$ for all $S_j \in S$ and the buyer's outside option utility is zero for all $B_i \in B$.

Definition 21 *A mechanism* (x, t) *satisfies the interim individual rationality (IIR) if, for each* $B_i \in B$*,*

$$\sum_{j=1}^{m} g_s(S_j|B_i)\left[x(B_i, S_j)v_b(B_i, S_j) - t_b(B_i, S_j)\right] \geqslant 0,$$

and for each $S_j \in S$*,*

$$\sum_{i=1}^{m} g_b(B_i|S_j)\left[(1 - x(B_i, S_j))\, v_s(B_i, S_j) - t_s(B_i, S_j)\right] \geqslant \sum_{i=1}^{m} g_b(B_i|S_j)v_s(B_i, S_j).$$

Definition 22 *A mechanism* (x, t) *satisfies decision efficiency (EFF) if, for each* $(B_i, S_j) \in B \times S$*,*

$$x(B_i, S_j) = \begin{cases} 1 & \text{if } v_b(B_i, S_j) > v_s(B_i, S_j) \\ 0 & \text{if } v_b(B_i, S_j) < v_s(B_i, S_j) \end{cases}$$

We denote by $x^*(\cdot)$ the efficient decision rule. We say that the buyer has no *redundant types* if, for any $B_i \neq B_k$, there exists some j such that

$$x^*(B_i, S_j) \neq x^*(B_k, S_j).$$

Similarly, the seller has no *redundant types* if, for any $j \neq l$, there exists some i such that

$$x^*(B_i, S_j) \neq x^*(B_i, S_l).$$

Throughout the paper, we assume that every agent has no redundant types.

We further require that what the seller receives be exactly the same as what the buyer pays. This is ex post budget balance.

Definition 23 *A mechanism (x, t) satisfies ex post budget balance (BB) if, for each $(B_i, S_j) \in B \times S$,*

$$t_b(B_i, S_j) + t_s(B_i, S_j) = 0.$$

3.3 Finite Type Spaces

In this section, we let $m \geqslant 2$, i.e., each agent has at least two types. First, we will introduce one generalized efficient decision rule and explain the intuition behind. Second, we will provide a necessary condition under which there exists a mechanism satisfying BIC, IIR, EFF, and BB under the generalized efficient decision rule.

3.3.1 The Efficient Decision Rule

Consider the following efficient decision rule:

$$x^*(B_i, S_j) = \begin{cases} 1 & \text{if } i \geqslant j \\ 0 & \text{otherwise} \end{cases}$$

The following table illustrates the above efficient decision rule:

Table 2

$x^*(\cdot)$	S_m	S_{m-1}	S_{m-2}	$\cdots$	S_1
B_m	1	1	1	$\cdots$	1
B_{m-1}	0	1	1	$\cdots$	1
B_{m-2}	0	0	1	$\cdots$	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
B_1	0	0	0	$\cdots$	1

Indeed, in the above table, every entry in the upper triangle is equal to 1 while all the other entries are zero. We show below that this table depicts the unique efficient decision rule considered in our setup.

Proposition 5 *There is a unique efficient decision rule described in Table 2.*

Proof: The proof consists of three steps. In Step 1, we show that $x^*(B_i, S_j)$ must be nondecreasing in B_i. In Step 2, we show that $x^*(B_i, S_j)$ must be nonincreasing in S_j. Finally, in Step 3, we induce the unique efficient decision rule in Table 2.

Step 1: $x^*(B_i, S_j)$ is nondecreasing in B_i.

Proof: Fix S_j as the seller's type. Suppose not. Then there exist $i > k$ such that $x^*(B_i, S_j) = 0$ and $x^*(B_k, S_j) = 1$. Notice that $x^*(B_k, S_j) = 1$ implies $v_b(B_k, S_j) > v_s(B_k, S_j)$. According to the single crossing condition, if the buyer's type increases to B_i, the buyer's valuation must grow faster; as a result,

$$v_b(B_i, S_j) > v_s(B_i, S_j)$$

must be true, implying $x^*(B_i, S_j) = 1$, which contradicts our hypothesis that $x^*(B_i, S_j) = 0$. ■

Step 2: $x^*(B_i, S_j)$ is nonincreasing in S_j.

Proof: Fix the buyer's type B_i. Suppose not. Then there exist $j > l$ such that $x^*(B_i, S_j) = 1$ and $x^*(B_i, S_l) = 0$. Notice that $x^*(B_i, S_l) = 0$ implies $v_s(B_i, S_l) > v_b(B_i, S_l)$. According to the single crossing condition, if the seller's type increases to S_j, the seller's valuation must grow faster; as a result,

$$v_s(B_i, S_j) > v_b(B_i, S_j)$$

must be true, implying $x^*(B_i, S_j) = 0$, which contradicts our hypothesis $x^*(B_i, S_j) = 1$. ■

Step 3: The efficient decision rule is uniquely determined as the one described in Table 2.

Proof: From the previous steps, we know that at profile (B_m, S_1), $x^*(\cdot)$ reaches its maximum; to aviod the trivial cases where it is always efficient not to trade, we let $x^*(B_m, S_1) = 1$. Also, we know that at profile (B_1, S_m), $x^*(\cdot)$ reaches its minimum; to aviod the trivial cases where it is always efficient to trade, we let $x^*(B_1, S_m) = 0$. Moreover, we set $x^*(B_1, S_1) = 1$ so that it is sometimes efficient to trade and sometimes not;[23] since $x^*(B_i, S_1)$ is nondecreasing in B_i, we obtain $x^*(B_i, S_1) = 1$ for all i. We obtain the following lemma:

Lemma 14 $x^*(B_i, S_{i+1}) = 0$ *and* $x^*(B_{i+1}, S_{i+1}) = 1$ *for all* $i = 1, \cdots, m-1$.

Proof: The proof is in the Appendix. ■

From Lemma 14, we obtain the following table:

$x^*(\cdot)$	S_m	S_{m-1}	S_{m-2}	$\cdots$	S_2	S_1
B_m	1					1
B_{m-1}	0	1				1
B_{m-2}		0	1			1
$\vdots$			$\ddots$	$\ddots$		$\vdots$
B_2				$\ddots$	1	1
B_1	0				0	1

Recall that for all i, $x^*(B_i, S_j)$ is nonincreasing in S_j. Then, since $x^*(B_i, S_{i+1}) = 0$, we obtain $x^*(B_i, S_j) = 0$ for all $j \geqslant i+1$; similarly, since $x^*(B_i, S_i) = 1$, we have $x^*(B_i, S_j) = 1$ for all $j \leqslant i$. We thus conclude that the efficient decision rule must be the one depicted in Table 2. This completes the proof of Step 3. ■ We complete the proof of the proposition by Steps 1,2, and 3. ■

3.3.2 A Necessary Condition for a Finite Type Space

For a general finite type space, we obtain the following necessary condition for the existence of mechanisms satisfying BIC, IIR, EFF, and BB:

[23] Otherwise, $x^*(B_1, S_1) = 0$ implies that $x^*(B_1, S_j) = 0$ for all j, because $x^*(B_1, S_j)$ is nonincreasing in S_j; then, NTNP property implies $t(B_1, S_j) = 0$ for all j. As a result, type B_1 is redundant as the payments associated with it are all zero.

Theorem 6 *There exists a mechanism satisfying BIC, IIR, EFF, and BB only if*

$$\sum_{i=1}^{m} g(B_i, S_1)v_b(B_1, S_1) + \sum_{i=2}^{m}\sum_{j=2}^{i} g(B_i, S_j)v_b(B_i, S_j)$$
$$\geq \sum_{j=1}^{m} g(B_m, S_j)v_s(B_m, S_m) + \sum_{j=1}^{m-1}\sum_{i=j}^{m-1} g(B_i, S_j)v_s(B_i, S_j). \quad (15)$$

Remark 1 *Notice that Theorem 6 is also true when there is no interdependence in valuation or no correlation in distribution. In particular, when there is no interdependence in valuation, each agent's valuation depends only on his own type and the valuation functions become $v_b(B_i)$ and $v_s(S_j)$; then, inequality (15) is rewritten as*

$$\sum_{i=1}^{m} g(B_i, S_1)v_b(B_1) + \sum_{i=2}^{m}\sum_{j=2}^{i} g(B_i, S_j)v_b(B_i)$$
$$\geq \sum_{j=1}^{m} g(B_m, S_j)v_s(S_m) + \sum_{j=1}^{m-1}\sum_{i=j}^{m-1} g(B_i, S_j)v_s(S_j).$$

On the other hand, when there is no correlation in distribution, $g(B_i, S_j) = g_b(B_i)g_s(S_j)$ for any (i, j); then, inequality (15) becomes

$$\sum_{i=1}^{m} g_b(B_i)g_s(S_1)v_b(B_1, S_1) + \sum_{i=2}^{m}\sum_{j=2}^{i} g_b(B_i)g_s(S_j)v_b(B_i, S_j)$$
$$\geq \sum_{j=1}^{m} g_b(B_m)g_s(S_j)v_s(B_m, S_m) + \sum_{j=1}^{m-1}\sum_{i=j}^{m-1} g_b(B_i)g_s(S_j)v_s(B_i, S_j).$$

Proof: In order to satisfy BB, we let $t_b(B_i, S_j) = -t_s(B_i, S_j) = t(B_i, S_j)$ for all $(B_i, S_j) \in B \times S$. According to the NTNP property, $x^*(B_i, S_j) = 0$ for any $i < j$ implies $t(B_i, S_j) = 0$ for any $i < j$. We write down the following IIR constraints:

$$IR_{B_1}: \quad g(B_1, S_1)\left(v_b(B_1, S_1) - t(B_1, S_1)\right) \geq 0;$$
$$IR_{S_m}: \quad g(B_m, S_m)t(B_m, S_m) \geq g(B_m, S_m)v_s(B_m, S_m).$$

To stop the buyer from deviating to B_1, we have the following BIC constriants: for any $i \neq 1$,

$$\sum_{j=1}^{i} g(B_i, S_j)\left(v_b(B_i, S_j) - t(B_i, S_j)\right) \geq g(B_i, S_1)\left(v_b(B_i, S_1) - t(B_1, S_1)\right)$$
$$\Rightarrow \sum_{j=2}^{i} g(B_i, S_j)\left(v_b(B_i, S_j) - t(B_i, S_j)\right) + g(B_i, S_1)\left(t(B_1, S_1) - t(B_i, S_1)\right) \geq 0$$

To stop the seller from deviating to S_m, we have the following BIC constraints: for any $j \neq m$,

$$\sum_{i=j}^{m} g(B_i,S_j)t(B_i,S_j) + \sum_{i=1}^{j-1} g(B_i,S_j)v_s(B_i,S_j) \geq g(B_m,S_j)t(B_m,S_m) + \sum_{i=1}^{m-1} g(B_i,S_j)v_s(B_i,S_j)$$

$$\Rightarrow \sum_{i=j}^{m-1} g(B_i,S_j)t(B_i,S_j) + g(B_m,S_j)\left(t(B_m,S_j) - t(B_m,S_m)\right) \geq \sum_{i=j}^{m-1} g(B_i,S_j)v_s(B_i,S_j).$$

In the following, we will show that after computation, inequality (15) is obtained as a necessary condition. First, we multiply IR_{B_1} by $\sum_{i=1}^{m} g(B_i,S_1)/g(B_1,S_1)$ and obtain

$$\sum_{i=1}^{m} g(B_i,S_1)\left(v_b(B_1,S_1) - t(B_1,S_1)\right) \geq 0. \tag{16}$$

Second, we multiply IR_{S_m} by $\sum_{j=1}^{m} g(B_m,S_j)/g(B_m,S_m)$ and obtain

$$\sum_{j=1}^{m} g(B_m,S_j)\left(t(B_m,S_m) - v_s(B_m,S_m)\right) \geq 0. \tag{17}$$

Third, adding up $IC_{B_i \to B_1}$ from $i = 2$ up to $i = m$, we have

$$\sum_{i=2}^{m}\sum_{j=2}^{i} g(B_i,S_j)v_b(B_i,S_j) - \sum_{i=2}^{m}\sum_{j=2}^{i} g(B_i,S_j)t(B_i,S_j) + \sum_{i=2}^{m} g(B_i,S_1)\left(t(B_1,S_1) - t(B_i,S_1)\right) \geq 0$$

$$\Rightarrow \sum_{i=2}^{m}\sum_{j=2}^{i} g(B_i,S_j)v_b(B_i,S_j) - \sum_{i=2}^{m}\sum_{j=1}^{i} g(B_i,S_j)t(B_i,S_j) + \sum_{i=2}^{m} g(B_i,S_1)t(B_1,S_1) \geq 0;$$

notice that the second term can be decomposed into the following:

$$-\sum_{i=2}^{m-1}\sum_{j=1}^{i} g(B_i,S_j)t(B_i,S_j) - \sum_{j=1}^{m-1} g(B_m,S_j)t(B_m,S_j) - g(B_m,S_m)t(B_m,S_m);$$

therefore, the inequality can be rewritten as follows:

$$\sum_{i=2}^{m}\sum_{j=2}^{i} g(B_i,S_j)v_b(B_i,S_j) - \sum_{i=2}^{m-1}\sum_{j=1}^{i} g(B_i,S_j)t(B_i,S_j) - \sum_{j=1}^{m-1} g(B_m,S_j)t(B_m,S_j)$$

$$-g(B_m,S_m)t(B_m,S_m) + \sum_{i=2}^{m} g(B_i,S_1)t(B_1,S_1) \geq 0 \tag{18}$$

Fourth, adding up $IC_{S_j \to S_m}$ from $j = 1$ up to $j = m-1$, we have

$$\sum_{j=1}^{m-1}\sum_{i=j}^{m-1} g(B_i,S_j)t(B_i,S_j) + \sum_{j=1}^{m-1} g(B_m,S_j)\left(t(B_m,S_j) - t(B_m,S_m)\right) \geq \sum_{j=1}^{m-1}\sum_{i=j}^{m-1} g(B_i,S_j)v_s(B_i,S_j)$$

$$\Rightarrow \sum_{j=1}^{m-1}\sum_{i=j}^{m} g(B_i,S_j)t(B_i,S_j) - \sum_{j=1}^{m-1} g(B_m,S_j)t(B_m,S_m) - \sum_{j=1}^{m-1}\sum_{i=j}^{m-1} g(B_i,S_j)v_s(B_i,S_j) \geq 0;$$

notice that the first term can be decomposed into the following:

$$\sum_{j=2}^{m-1}\sum_{i=j}^{m} g(B_i,S_j)t(B_i,S_j)+\sum_{i=2}^{m} g(B_i,S_1)t(B_i,S_1)+g(B_1,S_1)t(B_1,S_1);$$

therefore, the inequality can be rewritten as follows:

$$\begin{aligned}&\sum_{j=2}^{m-1}\sum_{i=j}^{m} g(B_i,S_j)t(B_i,S_j)+\sum_{i=2}^{m} g(B_i,S_1)t(B_i,S_1)+g(B_1,S_1)t(B_1,S_1)\\ &-\sum_{j=1}^{m-1} g(B_m,S_j)t(B_m,S_m)-\sum_{j=1}^{m-1}\sum_{i=j}^{m-1} g(B_i,S_j)v_s(B_i,S_j)\geq 0. \qquad (19)\end{aligned}$$

Finally, we obtain the following lemma:

Lemma 15 *Adding (16), (17), (18) with (19), we obtain the necessary condition, that is, inequality (15):*

$$\begin{aligned}&\sum_{i=1}^{m} g(B_i,S_1)v_b(B_1,S_1)+\sum_{i=2}^{m}\sum_{j=2}^{i} g(B_i,S_j)v_b(B_i,S_j)\\ \geq\ &\sum_{j=1}^{m} g(B_m,S_j)v_s(B_m,S_m)+\sum_{j=1}^{m-1}\sum_{i=j}^{m-1} g(B_i,S_j)v_s(B_i,S_j).\end{aligned}$$

Proof: The proof is in the Appendix. ■

This completes the proof. ■

3.4 The Case of Two Types

In this section, we confine our attention to a two-type model, i.e., each agent has only two types. In Subsection 3.4.1, we show that the identified necessary condition in the previous section turns out to be sufficient as well. Subsection 3.4.2 conducts the comparative statics for how the existence of desirable mechanisms rely on the interdependence of valuations and correlated signals in a two-type model with linear valuations.

3.4.1 Sufficiency for the Existence of Mechanisms Satisfying BIC, IIR, EFF, and BB

Proposition 6 *When each agent has only two types, i.e., $m = 2$. Then, the general necessary condition in Theorem 6 is also sufficient for the existence of mechanisms satisfying*

BIC, IIR, EFF, and BB.

Proof: If $m = 2$, the efficient decision rule in Table 2 is reduced to the following:

$x^*(\cdot)$	S_2	S_1
B_2	1	1
B_1	0	1

Next, the general necessary condition in Theorem 6 becomes

$$\sum_{i=1}^{2} g(B_i, S_1)v_b(B_1, S_1) + g(B_2, S_2)v_b(B_2, S_2) \geqslant \sum_{j=1}^{2} g(B_2, S_j)v_s(B_2, S_2) + g(B_1, S_1)v_s(B_1, S_1).$$

Dividing both sides of the above inequality by $g(B_2, S_1)$, we obtain

$$\left(\frac{g(B_1, S_1)}{g(B_2, S_1)} + 1\right) v_b(B_1, S_1) + \frac{g(B_2, S_2)}{g(B_2, S_1)} v_b(B_2, S_2) \geqslant \left(1 + \frac{g(B_2, S_2)}{g(B_2, S_1)}\right) v_s(B_2, S_2) + \frac{g(B_1, S_1)}{g(B_2, S_1)} v_s(B_1, S_1).$$

We set $\alpha_1 = g(B_1, S_1)/g(B_2, S_1)$ and $\alpha_2 = g(B_2, S_2)/g(B_2, S_1)$. Then, the general necessary condition can be further rewritten as

$$(1 + \alpha_1)\, v_b(B_1, S_1) + \alpha_2 v_b(B_2, S_2) \geqslant (1 + \alpha_2)\, v_s(B_2, S_2) + \alpha_1 v_s(B_1, S_1). \qquad (20)$$

Now we prove that (20) is also sufficient. Let $t(B_2, S_2) = v_s(B_2, S_2)$, $t(B_1, S_1) = v_b(B_1, S_1)$, and pick $t(B_2, S_1)$ such that

$$\begin{aligned} &\max\left\{v_b(B_1, S_1), v_s(B_2, S_2) - \alpha_{S_1}\left(v_b(B_1, S_1) - v_s(B_1, S_1)\right)\right\} \\ \leqslant &t(B_2, S_1) \leqslant \min\left\{v_s(B_2, S_2), v_b(B_1, S_1) + \alpha_{B_2}\left(v_b(B_2, S_2) - v_s(B_2, S_2)\right)\right\} \end{aligned}$$

For simplicity, let $A_1 \equiv v_s(B_2, S_2) - \alpha_{S_L}\left(v_b(B_1, S_1) - v_s(B_1, S_1)\right)$ and $A_2 \equiv v_b(B_1, S_1) + \alpha_{B_2}\left(v_b(B_2, S_2) - v_s(B_2, S_2)\right)$. First, we show that $t(B_2, S_1)$ is well-defined.

1. $v_b(B_1, S_1) - v_s(B_2, S_2) < 0$ because $v_b(B_1, S_1) \leqslant v_b(B_1, S_2) < v_s(B_1, S_2) \leqslant v_s(B_2, S_2)$.

2. $$\begin{aligned} v_b(B_1, S_1) - A_2 &= v_b(B_1, S_1) - v_b(B_1, S_1) - \alpha_{B_2}\left(v_b(B_2, S_2) - v_s(B_2, S_2)\right) \\ &= -\alpha_{B_2}\left(v_b(B_2, S_2) - v_s(B_2, S_2)\right) \\ &< 0 \end{aligned}$$

   because $v_b(B_2, S_2) - v_s(B_2, S_2) > 0$.

3. $A_1 - v_s(B_2, S_2) = v_s(B_2, S_2) - \alpha_{S_1}(v_b(B_1, S_1) - v_s(B_1, S_1)) - v_s(B_2, S_2)$

$$= -\alpha_{S_1}(v_b(B_1, S_1) - v_s(B_1, S_1)) < 0$$

because $v_b(B_1, S_1) - v_s(B_1, S_1) > 0$.

4. $A_1 - A_2 = v_s(B_2, S_2) - \alpha_{S_1}(v_b(B_1, S_1) - v_s(B_1, S_1))$

$$- v_b(B_1, S_1) - \alpha_{B_2}\alpha_{B_2}(v_b(B_2, S_2) - v_s(B_2, S_2))$$

$$= -(1 + \alpha_{S_1})v_b(B_1, S_1) - \alpha_{B_2}v_b(B_2, S_2) + \alpha_{S_1}v_s(B_1, S_1) + (1 + \alpha_{B_2})v_s(B_2, S_2)$$

$$\leqslant 0$$

because of (20).

Below we verify that the resulting mechanism (x^*, t) satisfies all the BIC and IIR constraints.

1. IR_{B_2} : LHS $= \alpha_{B_2}(v_b(B_2, S_2) - v_s(B_2, S_2)) + (v_b(B_2, S_1) - t(B_2, S_1))$

$$\geqslant \alpha_{B_2}(v_b(B_2, S_2) - v_s(B_2, S_2)) + v_b(B_2, S_1)$$

$$- v_b(B_1, S_1) - \alpha_{B_2}(v_b(B_2, S_2) - v_s(B_2, S_2)) \ (\because t(B_2, S_1) \leqslant A_2)$$

$$= v_b(B_2, S_1) - v_b(B_1, S_1)$$

$$> 0$$

because $v_b(B_i, S_j)$ is strictly increasing in B_i. Hence, IR_{B_2} is satisfied.

2. IR_{B_1}: LHS $= \alpha_{B_1}(v_b(B_1, S_1) - v_b(B_1, S_1)) = 0$. Hence, IR_{B_1} is satisfied.

3. IR_{S_2}: LHS $= \alpha_{S_2}(v_s(B_2, S_2) - v_s(B_2, S_2)) = 0$. Hence, IR_{S_2} is satisfied.

4. IR_{S_1} : LHS $= (t(B_2, S_1) - v_s(B_2, S_1)) + \alpha_{S_1}(v_b(B_1, S_1) - v_s(B_1, S_1))$

$$\geqslant v_s(B_2, S_2) - \alpha_{S_1}(v_b(B_1, S_1) - v_s(B_1, S_1)) - v_s(B_2, S_1)$$

$$+ \alpha_{S_1}(v_b(B_1, S_1) - v_s(B_1, S_1)) \ (\because t(B_2, S_1) \geqslant A_1)$$

$$= v_s(B_2, S_2) - v_s(B_2, S_1)$$

$$> 0$$

because $v_s(B_i, S_j)$ is strictly increasing in S_j. Hence, IR_{S_1} is satisfied.

5. $IC_{B_2 \to B_1}$: LHS $\geqslant v_b(B_2, S_1) - v_b(B_1, S_1)$ according to IR_{B_2};

RHS $= v_b(B_2, S_1) - v_b(B_1, S_1)$

Hence, $IC_{B_2 \to B_1}$ is satisfied.

6. $IC_{B_1 \to B_2}$: LHS $= 0$ according to IR_{B_1};

$$\begin{aligned} \text{RHS} &= v_b(B_1, S_2) - v_s(B_2, S_2) + \alpha_{B_1} \left(v_b(B_1, S_1) - t(B_2, S_1) \right) \\ &\leqslant v_b(B_1, S_2) - v_s(B_2, S_2) \; (\because t(B_2, S_1) \geqslant v_b(B_1, S_1)) \\ &< 0 \; (\because v_b(B_1, S_2) < v_s(B_1, S_2) \leqslant v_s(B_2, S_2)) \end{aligned}$$

Hence, $IC_{B_1 \to B_2}$ is satisfied.

7. $IC_{S_2 \to S_1}$: LHS $= 0$ according to IR_{S_2};

$$\begin{aligned} \text{RHS} &= \alpha_{S_2} \left(t(B_2, S_1) - v_s(B_2, S_2) \right) + v_b(B_1, S_1) - v_s(B_1, S_2) \\ &\leqslant v_b(B_1, S_1) - v_s(B_1, S_2) \; (\because t(B_2, S_1) \leqslant v_s(B_2, S_2)) \\ &< 0 \; (\because v_b(B_1, S_1) \leqslant v_b(B_1, S_2) < v_s(B_1, S_2)) \end{aligned}$$

Hence, $IC_{S_2 \to S_1}$ is satisfied.

8. $IC_{S_1 \to S_2}$: LHS $\geqslant v_s(B_2, S_1) - v_s(B_2, S_1)$ acoording to IR_{S_1};

$$\text{RHS} = v_s(B_2, S_2) - v_s(B_2, S_1)$$

Hence, $IC_{S_1 \to S_2}$ is satisfied.

This completes the proof. ■

Remark 2 *Consider a private-value setup where $v_b(B_i, S_j) = B_i$ and $v_s(B_i, S_j) = S_j$. Moreover, we assume that the agents' types are independently distributed. So, let g_b and g_s be the priors over $B = \{B_1, B_2\}$ and $S = \{S_1, S_2\}$, respectively, such that*

$$g_b(B_i) = \begin{cases} \delta & \text{if } B_i = B_1 \\ 1 - \delta & \text{if } B_i = B_2 \end{cases}$$

and

$$g_s(S_j) = \begin{cases} \varepsilon & \text{if } S_j = S_2 \\ 1 - \varepsilon & \text{if } S_j = S_1 \end{cases}$$

Then

$$\alpha_{S_1} = \frac{g(B_1, S_1)}{g(B_2, S_1)} = \frac{g_b(B_1)}{g_b(B_2)} = \frac{\delta}{1 - \delta}$$

and

$$\alpha_{B_2} = \frac{g(B_2, S_2)}{g(B_2, S_1)} = \frac{g_s(S_2)}{g_s(S_1)} = \frac{\varepsilon}{1 - \varepsilon}.$$

The necessary and sufficient condition (20) we identify becomes

$$\left(1+\frac{\delta}{1-\delta}\right)B_1+\frac{\varepsilon}{1-\varepsilon}B_2 \;\;\geqslant\;\; \frac{\delta}{1-\delta}S_1+\left(1+\frac{\varepsilon}{1-\varepsilon}\right)S_2.$$

This inequality is further summarized as

$$\varepsilon(1-\varepsilon)B_1+(1-\delta)B_2 \;\;\geqslant\;\; \delta(1-\varepsilon)S_1+(1-\delta)S_2,$$

which is the same as Matsuo (1989).

3.4.2 Comparative Statics under Linear Valuations

In this section, we restrict attention to a linear valuation setup and assume that the valuation functions are given as $v_b(B_i, S_j) = B_i + \gamma_b S_j$ and $v_s(B_i, S_j) = S_j + \gamma_s B_i$ where $\gamma_b \in [0,1]$ denotes the buyer's degree of interdependence of preferences and $\gamma_s \in [0,1]$ denotes the seller's counterpart. This specification suggests that the single crossing condition holds and each agent's valuation function is strictly increasing in his own type and nondecreasing in the other agent's type.

The agents' types are drawn from the following joint probability distribution $g(\cdot,\cdot)$:

$g(\cdot,\cdot)$	S_2	S_1
B_2	a	$1/2-a$
B_1	$1/2-a$	a

where $0 \leq a \leq 1/2$. Then, the marginal distribution $g(\cdot,\cdot)$ is given as follows:

$$g_b(B_2) = g_b(B_1) = g_s(S_2) = g_s(S_1) = 1/2,$$

and the following probability ratios are given as

$$\alpha_{S_1} = \alpha_{B_2} = \frac{g(B_1,S_1)}{g(B_2,S_1)} = \frac{a}{1/2-a}.$$

Notice that if $a = 1/4$, $g(B_i, S_j) = g_b(B_i)g_s(S_j)$ for each $(B_i, S_j) \in B \times S$, i.e., the agents' types are independently distributed. On the other hand, if $a = 1/2$, then $g(B_1, S_2) = g(B_2, S_1) = 0$ so that the types of both the buyer and seller are perfectly correlated, i.e., either they both have a high type or they both have a low type. The analysis of this section considers the following efficient decision rule, that is,

$x^*(\cdot)$	S_2	S_1
B_2	1	1
B_1	0	1

Lemma 16 *For each* $a \in [0, 1/2]$ *and* $\gamma_b \in [0, 1]$*, there exists* $\bar{\gamma}_s(a, \gamma_b) \in \mathbb{R}$ *such that if* $\bar{\gamma}_s(a, \gamma_b) \in [0, 1]$*, there exists a NTNP mechanism satisfying BIC, IIR, EFF, and BB in any model specified by* (a, γ_b, γ_s) *where* $\gamma_s \in [0, \bar{\gamma}_s]$*.*

Proof: Recall the necessary and sufficient condition in Proposition 6:

$$(1 + \alpha_{S_1})v_b(B_1, S_1) + \alpha_{B_2}v_b(B_2, S_2) \geq \alpha_{S_1}v_s(B_1, S_1) + (1 + \alpha_{B_2})v_s(B_2, S_2).$$

Plugging the probability ratios $\alpha_{S_1} = \alpha_{B_2} = a/(0.5 - a)$ into the above inequalities, we obtain

$$\left(1 + \frac{a}{0.5 - a}\right) v_b(B_1, S_1) + \frac{a}{0.5 - a} v_b(B_2, S_2) \geqslant \frac{a}{0.5 - a} v_s(B_1, S_1) + \left(1 + \frac{a}{0.5 - a}\right) v_s(B_2, S_2).$$

This can be summarized as follow:

$$0.5v_b(B_1, S_1) + av_b(B_2, S_2) \geqslant av_s(B_1, S_1) + 0.5v_s(B_2, S_2).$$

Plugging $v_b(B_i, S_j) = B_i + \gamma_b S_j$ and $v_s(B_i, S_j) = S_j + \gamma_s B_i$ into the above inequalities and thereafter summarizing the terms for γ_s, we obtain

$$\gamma_s \leq \frac{0.5(B_1 - S_2) + a(B_2 - S_1)}{aB_1 + 0.5B_2} + \frac{0.5S_1 + aS_2}{aB_1 + 0.5B_2}\gamma_b.$$

So, we set

$$\bar{\gamma}_s(a, \gamma_b) = \frac{0.5(B_1 - S_2) + a(B_2 - S_1)}{aB_1 + 0.5B_2} + \frac{0.5S_1 + aS_2}{aB_1 + 0.5B_2}\gamma_b.$$

This completes the proof. ■

Therefore, the necessary and sufficient condition in two-type space implies an upper bound on γ_s. The intuition is as follows: if γ_s is very large, the seller is asking such a high price that the buyer would rather quit the mechanism. In other words, the buyer's IIR constraint is violated.

We establish the following comparative statics results.

Proposition 7 *The following comparative statics results are obtained.*

1. *For any $a \in [0, 1/2]$, $\bar{\gamma}_s(a, \gamma_b)$ is increasing in γ_b.*

2. *Assume that $B_1S_2 - B_2S_1 > 0$. For any $\gamma_b \in [0, 1]$, $\bar{\gamma}_s(a, \gamma_b)$ is increasing in a.*

Proof: This proof follows straightforwardly from the following:

$$\begin{aligned}
\frac{\partial \bar{\gamma}_s(a, \gamma_b)}{\partial \gamma_b} &= \frac{0.5S_1 + aS_2}{aB_1 + 0.5B_2} > 0; \\
\frac{\partial \hat{\gamma}_2(\gamma_b, a)}{\partial a} &= \frac{0.5\left[(B_2 - B_1)(B_2 + B_1) + (B_1S_2 - B_2S_1) + 0.5\gamma_b(B_2S_2 - B_1S_1)\right]}{(aB_1 + 0.5B_2)^2} \geq 0.
\end{aligned}$$

■

The intuition is as follows:

1. when γ_b increases, the buyer has a higher valuation on average and is willing to pay a higher price; hence, the seller can ask for a higher price and γ_s increases.

2. If a increases, after the seller knows his own type, he also has some information about the buyer's type which he can use to extract full surplus. For example, if a increases, a high-type seller is more likely to meet the high-type buyer and thus he can ask for a price which is very close to buyer's valuation. Hence, the upper bound of γ_s increases.

3.5 Dropping IIR: A Linear Programming Approach

In this section, we drop IIR as part of the requirement for desirable mechanisms and characterize the set of valuation functions and correlated beliefs that guaranteeing mechanisms satisfying BIC, EFF, and BB as a system of inequalities. We solve this system of inequalities by using Farkas Lemma, which allows us to check whether such a system of inequalities admits a solution or not. Throughout this section, we restrict our attention to a model with two types for each agent.

First of all, in order to satisfy BB, we let $t_b(B_i, S_j) = -t_s(B_i, S_j) = t(B_i, S_j)$ for all $(B_i, S_j) \in B \times S$.

Suppose each agent have m types. Then we write down the BIC constraints for each agent. If the buyer's true type is B_i and he deviates to B_k, then BIC requires that the buyer obtain a higher expected utility under truth-telling, that is,

$$\begin{aligned}&\sum_{j=1}^{m} g_s(S_j|B_i)\left(x^*(B_i,S_j)v_b(B_i,S_j)-t(B_i,S_j)\right)\\ \geqslant\ &\sum_{j=1}^{m} g_s(S_j|B_i)\left(x^*(B_k,S_j)v_b(B_i,S_j)-t(B_k,S_j)\right);\end{aligned}$$

after rearrangment, we obtain

$$\begin{aligned}&\sum_{j=1}^{m} g_s(S_j|B_i)t(B_i,S_j)-\sum_{j=1}^{m} g_s(S_j|B_i)t(B_k,S_j)\\ \leqslant\ &\sum_{j=1}^{m} g_s(S_j|B_i)\left(x^*(B_i,S_j)-x^*(B_k,S_j)\right)v_b(B_i,S_j).\end{aligned}$$

Analogously, if the seller's true type is S_j and he deviates to S_l, then BIC requires that the seller obtain a higher expected utility under truth-telling, that is,

$$\begin{aligned}&\sum_{i=1}^{m} g_b(B_i|S_j)\left[\left(1-x^*(B_i,S_j)\right)v_s(B_i,S_j)-t_s(B_i,S_j)\right]\\ \geqslant\ &\sum_{i=1}^{m} g_b(B_i|S_j)\left[\left(1-x^*(B_i,S_l)\right)v_s(B_i,S_j)-t_s(B_i,S_l)\right];\end{aligned}$$

after rearrangement, we obtain

$$\begin{aligned}&\sum_{i=1}^{m} g_b(B_i|S_j)t_s(B_i,S_j)-\sum_{i=1}^{m} g_b(B_i|S_j)t_s(B_i,S_l)\\ \leqslant\ &\sum_{i=1}^{m} g_b(B_i|S_j)\left(x^*(B_i,S_l)-x^*(B_i,S_j)\right)v_s(B_i,S_j);\end{aligned}$$

since $t_s(B_i,S_j)=-t(B_i,S_j)$ for all $(B_i,S_j)\in B\times S$, the seller's BIC constraint can be further rewritten as follows:

$$\begin{aligned}&\sum_{i=1}^{m} g_b(B_i|S_j)t(B_i,S_l)-\sum_{i=1}^{m} g_b(B_i|S_j)t(B_i,S_j)\\ \leqslant\ &\sum_{i=1}^{m} g_b(B_i|S_j)\left(x^*(B_i,S_l)-x^*(B_i,S_j)\right)v_s(B_i,S_j).\end{aligned}$$

Putting all the BIC constraints together, we have

$$\sum_{j=1}^{m} g_s(S_j|B_i)t(B_i,S_j)-\sum_{j=1}^{m} g_s(S_j|B_i)t(B_k,S_j)\leqslant\sum_{j=1}^{m} g_s(S_j|B_i)\left(x^*(B_i,S_j)-x^*(B_k,S_j)\right)v_b(B_i,S_j)$$

for any $i \neq k$, and

$$\sum_{i=1}^{m} g_b(B_i|S_j)t(B_i,S_l) - \sum_{i=1}^{m} g_b(B_i|S_j)t(B_i,S_j) \leqslant \sum_{i=1}^{m} g_b(B_i|S_j)\left(x^*(B_i,S_l) - x^*(B_i,S_j)\right)v_s(B_i,S_j)$$

for any $j \neq l$.

In what follows, we will consider the two-type space and show that the system of BIC and BB constraints has a solution.

Proposition 8 *When each agent has only two types, there exists a mechanism (x^*, t) satisfying BIC, EFF, and BB, where x^* is the efficient decision rule.*

Remark 3 *The proposition remains correct even when the agents' types are independently distributed, because the proof also works under independent distribution.*

Proof: The proof consists of three steps. In Step 1, we write down all the constraints that the transfer rule t must satisfy in order to satisfy BIC and BB. In Step 2, we treat these constraints as a system of inequalities and write down its Farkas alternative. In Step 3, we show that the Farkas alternative has no solution. Finally, by Farkas Lemma, the original system of inequalities must have a solution.

Step 1: We write down all the constraints that the transfer rule t must satisfy in order to achieve BIC and BB.

Proof: Putting all the BIC constraints together, we have

$$\begin{aligned}
\sum_{j=1}^{2} g_s(S_j|B_1)t(B_1,S_j) - \sum_{j=1}^{2} g_s(S_j|B_1)t(B_2,S_j) &\leqslant \sum_{j=1}^{2} g_s(S_j|B_1)\left(x^*(B_1,S_j) - x^*(B_2,S_j)\right)v_b(B_1,S_j);\\
\sum_{j=1}^{2} g_s(S_j|B_2)t(B_2,S_j) - \sum_{j=1}^{2} g_s(S_j|B_2)t(B_1,S_j) &\leqslant \sum_{j=1}^{2} g_s(S_j|B_2)\left(x^*(B_2,S_j) - x^*(B_1,S_j)\right)v_b(B_2,S_j);\\
\sum_{i=1}^{2} g_b(B_i|S_1)t(B_i,S_2) - \sum_{i=1}^{2} g_b(B_i|S_1)t(B_i,S_1) &\leqslant \sum_{i=1}^{2} g_b(B_i|S_1)\left(x^*(B_i,S_2) - x^*(B_i,S_1)\right)v_s(B_i,S_1);\\
\sum_{i=1}^{2} g_b(B_i|S_2)t(B_i,S_1) - \sum_{i=1}^{2} g_b(B_i|S_2)t(B_i,S_2) &\leqslant \sum_{i=1}^{2} g_b(B_i|S_2)\left(x^*(B_i,S_1) - x^*(B_i,S_2)\right)v_s(B_i,S_2).
\end{aligned}$$

Recall that the efficient decision rule in Table 2 becomes the following when $m = 2$:

$x^*(\cdot)$	S_2	S_1
B_2	1	1
B_1	0	1

Then we can simplify the right-hand sides of the BIC constraints and obtain the following:

$$\begin{aligned}
\sum_{j=1}^{2} g_s(S_j|B_1)t(B_1,S_j) - \sum_{j=1}^{2} g_s(S_j|B_1)t(B_2,S_j) &\leqslant -g_s(S_2|B_1)v_b(B_1,S_2);\\
\sum_{j=1}^{2} g_s(S_j|B_2)t(B_2,S_j) - \sum_{j=1}^{2} g_s(S_j|B_2)t(B_1,S_j) &\leqslant g_s(S_2|B_2)v_b(B_2,S_2);\\
\sum_{i=1}^{2} g_b(B_i|S_1)t(B_i,S_2) - \sum_{i=1}^{2} g_b(B_i|S_1)t(B_i,S_1) &\leqslant -g_b(B_1|S_1)v_s(B_1,S_1);\\
\sum_{i=1}^{2} g_b(B_i|S_2)t(B_i,S_1) - \sum_{i=1}^{2} g_b(B_i|S_2)t(B_i,S_2) &\leqslant g_b(B_1|S_2)v_s(B_1,S_2). \qquad (21)
\end{aligned}$$

Therefore, finding a transfer rule t which, together with the efficient decision rule in Table 2, satisfies EFF, BIC, and BB, is equivalent to check whether (21) has a solution or not. ■

Step 2: We write down the Farkas alternative of (21).

Proof: First, we will convert (21) into the standard form $\{x : Ax = b, x \geqslant 0\}$. Observe that any inequality in (21) can be turned into an equation by the addition of a slack variable ε. That is, define four new variables $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4 \geqslant 0$ such that

$$\begin{aligned}
\sum_{j=1}^{2} g_s(S_j|B_1)t(B_1,S_j) - \sum_{j=1}^{2} g_s(S_j|B_1)t(B_2,S_j) + \varepsilon_1 &= -g_s(S_2|B_1)v_b(B_1,S_2);\\
\sum_{j=1}^{2} g_s(S_j|B_2)t(B_2,S_j) - \sum_{j=1}^{2} g_s(S_j|B_2)t(B_1,S_j) + \varepsilon_2 &= g_s(S_2|B_2)v_b(B_2,S_2);\\
\sum_{i=1}^{2} g_b(B_i|S_1)t(B_i,S_2) - \sum_{i=1}^{2} g_b(B_i|S_1)t(B_i,S_1) + \varepsilon_3 &= -g_b(B_1|S_1)v_s(B_1,S_1);\\
\sum_{i=1}^{2} g_b(B_i|S_2)t(B_i,S_1) - \sum_{i=1}^{2} g_b(B_i|S_2)t(B_i,S_2) + \varepsilon_4 &= g_b(B_1|S_2)v_s(B_1,S_2).
\end{aligned}$$

Beside, the monetary transfer $t(B_i, S_j)$ for any (i, j) that is unrestricted in sign can be replaced by two non-negative variables a_{ij} and b_{ij} by setting $t(B_i, S_j) = a_{ij} - b_{ij}$. Then,

(21) becomes

$$
\begin{aligned}
\sum_{j=1}^{2} g_s(S_j|B_1)\,(a_{1j}-b_{1j}) - \sum_{j=1}^{2} g_s(S_j|B_1)\,(a_{2j}-b_{2j}) + \varepsilon_1 &= -g_s(S_2|B_1)v_b(B_1,S_2);\\
\sum_{j=1}^{2} g_s(S_j|B_2)\,(a_{2j}-b_{2j}) - \sum_{j=1}^{2} g_s(S_j|B_2)\,(a_{1j}-b_{1j}) + \varepsilon_2 &= g_s(S_2|B_2)v_b(B_2,S_2);\\
\sum_{i=1}^{2} g_b(B_i|S_1)\,(a_{i2}-b_{i2}) - \sum_{i=1}^{2} g_b(B_i|S_1)\,(a_{i1}-b_{i1}) + \varepsilon_3 &= -g_b(B_1|S_1)v_s(B_1,S_1);\\
\sum_{i=1}^{2} g_b(B_i|S_2)\,(a_{i1}-b_{i1}) - \sum_{i=1}^{2} g_b(B_i|S_2)\,(a_{i2}-b_{i2}) + \varepsilon_4 &= g_b(B_1|S_2)v_s(B_1,S_2),
\end{aligned}
$$

where $a_{ij}, b_{ij} \geqslant 0$ for all (i,j) and $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4 \geqslant 0$. We can also rewrite it into matrix form: let

$$
\begin{aligned}
A &= \begin{bmatrix}
g_s(S_1|B_1) & -g_s(S_1|B_1) & g_s(S_2|B_1) & -g_s(S_2|B_1) & -g_s(S_1|B_1) & g_s(S_1|B_1) & -g_s(S_2|B_1) & g_s(S_2|B_1) & 1 & 0 & 0 & 0\\
-g_s(S_1|B_2) & g_s(S_1|B_2) & -g_s(S_2|B_2) & g_s(S_2|B_2) & g_s(S_1|B_2) & -g_s(S_1|B_2) & g_s(S_2|B_2) & -g_s(S_2|B_2) & 0 & 1 & 0 & 0\\
-g_b(B_1|S_1) & g_b(B_1|S_1) & g_b(B_1|S_1) & -g_b(B_1|S_1) & -g_b(B_2|S_1) & g_b(B_2|S_1) & g_b(B_2|S_1) & -g_b(B_2|S_1) & 0 & 0 & 1 & 0\\
g_b(B_1|S_2) & -g_b(B_1|S_2) & -g_b(B_1|S_2) & g_b(B_1|S_2) & g_b(B_2|S_2) & -g_b(B_2|S_2) & -g_b(B_2|S_2) & g_b(B_2|S_2) & 0 & 0 & 0 & 1
\end{bmatrix}\\
x &= \begin{bmatrix} a_{11} & b_{11} & a_{12} & b_{12} & a_{21} & b_{21} & a_{22} & b_{22} & \varepsilon_1 & \varepsilon_2 & \varepsilon_3 & \varepsilon_4 \end{bmatrix}^T\\
b &= \begin{bmatrix} -g_s(S_2|B_1)v_b(B_1,S_2)\\ g_s(S_2|B_2)v_b(B_2,S_2)\\ -g_b(B_1|S_1)v_s(B_1,S_1)\\ g_b(B_1|S_2)v_s(B_1,S_2) \end{bmatrix}
\end{aligned}
$$

Then, (21) is converted into the standard form $\{x \in \mathbb{R}^{12} : Ax = b, x \geqslant 0\}$.

According to Farkas Lemma, either (21) has a solution or there exists $y \in \mathbb{R}^4$ such that $yA \geqslant 0$ and $y \cdot b < 0$. Let

$$
y = \begin{bmatrix} y_1 & y_2 & y_3 & y_4 \end{bmatrix}
$$

Then, $yA \geqslant 0$ implies

$$
\begin{aligned}
y_1 g_s(S_1|B_1) - y_2 g_s(S_1|B_2) - y_3 g_b(B_1|S_1) + y_4 g_b(B_1|S_2) &\geqslant 0; && (22)\\
-y_1 g_s(S_1|B_1) + y_2 g_s(S_1|B_2) + y_3 g_b(B_1|S_1) - y_4 g_b(B_1|S_2) &\geqslant 0; && (23)\\
y_1 g_s(S_2|B_1) - y_2 g_s(S_2|B_2) + y_3 g_b(B_1|S_1) - y_4 g_b(B_1|S_2) &\geqslant 0; && (24)\\
-y_1 g_s(S_2|B_1) + y_2 g_s(S_2|B_2) - y_3 g_b(B_1|S_1) + y_4 g_b(B_1|S_2) &\geqslant 0; && (25)\\
-y_1 g_s(S_1|B_1) + y_2 g_s(S_1|B_2) - y_3 g_b(B_2|S_1) + y_4 g_b(B_2|S_2) &\geqslant 0; && (26)\\
y_1 g_s(S_1|B_1) - y_2 g_s(S_1|B_2) + y_3 g_b(B_2|S_1) - y_4 g_b(B_2|S_2) &\geqslant 0; && (27)\\
-y_1 g_s(S_2|B_1) + y_2 g_s(S_2|B_2) + y_3 g_b(B_2|S_1) - y_4 g_b(B_2|S_2) &\geqslant 0; && (28)\\
y_1 g_s(S_2|B_1) - y_2 g_s(S_2|B_2) - y_3 g_b(B_2|S_1) + y_4 g_b(B_2|S_2) &\geqslant 0; && (29)\\
y_1, y_2, y_3, y_4 &\geqslant 0. && (30)
\end{aligned}
$$

Notice that

1. Inequality (22) and (23) imply

$$y_1 g_s(S_1|B_1) - y_2 g_s(S_1|B_2) - y_3 g_b(B_1|S_1) + y_4 g_b(B_1|S_2) = 0;$$

2. Inequality (24) and (25) imply

$$y_1 g_s(S_2|B_1) - y_2 g_s(S_2|B_2) + y_3 g_b(B_1|S_1) - y_4 g_b(B_1|S_2) = 0;$$

3. Inequality (26) and (27) imply

$$y_1 g_s(S_1|B_1) - y_2 g_s(S_1|B_2) + y_3 g_b(B_2|S_1) - y_4 g_b(B_2|S_2) = 0;$$

4. Inequality (28) and (29) imply

$$y_1 g_s(S_2|B_1) - y_2 g_s(S_2|B_2) - y_3 g_b(B_2|S_1) + y_4 g_b(B_2|S_2) = 0.$$

On the other hand $y \cdot b < 0$ implies

$$-y_1 g_s(S_2|B_1) v_b(B_1, S_2) + y_2 g_s(S_2|B_2) v_b(B_2, S_2) - y_3 g_b(B_1|S_1) v_s(B_1, S_1) + y_4 g_b(B_1|S_2) v_s(B_1, S_2) < 0.$$

To summarize, the Farkas alternative of (21) is

$$\begin{aligned} y_1 g_s(S_1|B_1) - y_2 g_s(S_1|B_2) - y_3 g_b(B_1|S_1) + y_4 g_b(B_1|S_2) &= 0; \quad (31)\\ y_1 g_s(S_2|B_1) - y_2 g_s(S_2|B_2) + y_3 g_b(B_1|S_1) - y_4 g_b(B_1|S_2) &= 0; \quad (32)\\ y_1 g_s(S_1|B_1) - y_2 g_s(S_1|B_2) + y_3 g_b(B_2|S_1) - y_4 g_b(B_2|S_2) &= 0; \quad (33)\\ y_1 g_s(S_2|B_1) - y_2 g_s(S_2|B_2) - y_3 g_b(B_2|S_1) + y_4 g_b(B_2|S_2) &= 0; \quad (34)\\ -y_1 g_s(S_2|B_1) v_b(B_1,S_2) + y_2 g_s(S_2|B_2) v_b(B_2,S_2) &\\ -y_3 g_b(B_1|S_1) v_s(B_1,S_1) + y_4 g_b(B_1|S_2) v_s(B_1,S_2) &< 0, \quad (35) \end{aligned}$$

where $y_1, y_2, y_3, y_4 \geqslant 0$. ■

Step 3: We show that the Farkas alternative of (21) has no solutions.

Proof: Multiplying (32) by $-v_s(B_1, S_2)$, we obtain

$$v_s(B_1, S_2)\left[-y_1 g_s(S_2|B_1) + y_2 g_s(S_2|B_2) - y_3 g_b(B_1|S_1) + y_4 g_b(B_1|S_2)\right] = 0. \quad (36)$$

Define

$$\begin{aligned} A(y_1, y_2, y_3) &= -y_1 g_s(S_2|B_1)\left(v_b(B_1, S_2) - v_s(B_1, S_2)\right) + y_2 g_s(S_2|B_2)\left(v_b(B_2, S_2) - v_s(B_1, S_2)\right)\\ &\quad -y_3 g_b(B_1|S_1)\left(v_s(B_1, S_1) - v_s(B_1, S_2)\right). \end{aligned}$$

Notice that $A(y_1, y_2, y_3) > 0$ because

1. $v_b(B_1, S_2) - v_s(B_1, S_2) < 0$ because $x^*(B_1, S_2) = 0$;

2. $v_b(B_2, S_2) - v_s(B_1, S_2) > 0$ because $v_b(B_2, S_2) > v_s(B_2, S_2) \geqslant v_s(B_1, S_2)$;

3. $v_s(B_1, S_1) - v_s(B_1, S_2) < 0$ because $v_s(B_i, S_j)$ is strictly increasing in S_j.

Adding $A(y_1, y_2, y_3)$ to the left-hand side of (36), we obtain

$$\begin{aligned} -y_1 g_s(S_2|B_1) v_b(B_1, S_2) + y_2 g_s(S_2|B_2) v_b(B_2, S_2) &\\ -y_3 g_b(B_1|S_1) v_s(B_1, S_1) + y_4 g_b(B_1|S_2) v_s(B_1, S_2) &> 0, \end{aligned}$$

contradicting (35). Therefore, the Farkas alternative of (21) has no solutions. Then, by Farkas Lemma, (21) must have a solution. ■

Steps 1 through 3 show that there exists a transfer rule t which, together with the efficient decision rule in Table 2, satisfies BIC, EFF, and BB in a two-type model. This completes the proof of the proposition. ■

This proposition says that there generally exists a mechanism satisfying BIC, EFF, and BB. Therefore, IIR imposes a stringent restriction on the class of desirable mechanisms together with the other three properties.

3.6 Concluding Remarks

This paper identifies a necessary condition for the existence of mechanisms satisfying BIC, IIR, EFF, and BB in a bilateral trade model with a discrete type space. In the rest of the paper, however, we confine our attention to a two-type model to further obtain the implications of the paper's model. What clearly needs to be done is to extend the whole analysis of the paper to a general finite type space. Although this is a technically challenging question, we are currently working on this extension.

4 Reference

1. Börgers, T., *An Introduction to the Theory of Mechanism Design*, Oxford University Press, 2015.

2. Börgers, T., and P. Norman, "A Note on Budget Balance under Interim Participation Constraints: the Case of Independent Types," *Economic Theory*, (2009), 477-489.

3. Clarke, E., "Multipart Pricing of Public Goods," *Public Choice*, vol. 2, (1971), 19-33.

4. Cramton, R., R. Gibbons, and P. Klemperer, "Dissolving a Partnership Efficiently," *Econometrica*, 55, (1987), 615-632.

5. Cremer, J. and R.P. McLean, "Full extraction of the surplus in Bayesian and dominant strategy auctions," *Econometrica*, vol. 56, (1988), 1247-1257.

6. Cremer, J. and R.P. McLean, "Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist When Demands Are Interdependent," *Econometrica*, vol. 53, (1985), 345-361.

7. Dasgupta, P., and E. Maskin, "Efficient Auctions," *Quarterly Journal of Economics*, vol. 115, (2000), 341-388.

8. d'Aspremont, C. and L.A. Gerard-Varet, "Incentives and Incomplete Information," *Journal of Public Economics*, vol. 11, (1979), 25-45.

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

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

11. Green, J. and J.J. Laffont, "Characterization of Satisfactory Mechanisms for the Revelation of Preferences for Public Goods," *Econometrica*, vol. 45, (1977), 427-438.

12. Gresik, T.A., "Efficient bilateral trade with statistically dependent beliefs," *Journal of Economic Theory*, vol. 53, (1991a), 199-205.

13. Gresik, T.A., "Ex ante incentive efficient trading mechanisms without the private valuation restriction," *Journal of Economic Theory*, vol. 55, (1991b), 41-63.

14. Groves, T., "Incentives in Teams," *Econometrica*, vol. 41, (1973), 617-631.

15. Hellwig, M., "Public-Good Provision with Many Participants," *The Review of Economic Studies*, vol. 70, (2003), 589-614.

16. Kos, N. and M. Manea, "Efficient Trade Mechanism with Discrete Values," *working paper*, 2009.

17. Krishana, V., and M. Perry, "Efficient Mechanism Design," *Working Paper*, 2000.

18. Krishna, V., *Auction theory, the second edition*, Academic press, 2009.

19. Ledyard, J.O. and T.R. Palfrey, "Voting and Lottery Drafts as Efficient Public Goods Mechanisms," *The Review of Economic Studies*, vol. 61, (1994), 327-355.

20. Mailath, G.J. and A. Postlewaite, "Asymmetric Information Bargaining Problems with Many Agents," *The Review of Economic Studies*, vol. 57, (1990), 351-367.

21. Mas-Colell, A., M. Whinston, and J. Green, *Microeconomic Theory*, Oxford University Press, 1995.

22. Matsuo, T., "On incentive compatible, individually rational, and ex post efficient mechanisms for bilateral trading," *Journal of Economic theory*, vol. 49, (1989), 189-194.

23. Matsushima, H., and S. Noda, "Mechanism Design with Blockchain Enforcement," *Working Paper*, 2020.

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

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

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

27. Mirsky, L., *An introduction to linear algebra*, Courier Corporation, 2012.

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

29. Osborne, M.J. and A. Rubinstein, *A Course in Game Theory*, (1994), The MIT Press.

30. Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, McGraw Hill, 1991.

31. Saijo, T. and T. Yamato, "A voluntary participation game with a non-excludable public good," *Journal of Economic Theory*, vol. 84, (1999), 227-242.

32. Vickrey, W., "Counterspeculation, Auctions, and Competitive Sealed Tenders," *Journal of Finance*, vol. 16, (1961), 8-37.

33. Vohra, R.V., *Mechanism design: a linear programming approach*, Cambridge University Press, 2011.

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

35. Yenmez, M.B., "Incentive-compatible matching mechanisms: consistency with various stability notions," *American Economic Journal: Microeconomics*, vol. 5, (2013), 120-141.

A Appendix to Chapter 1

In the Appendix, we provide all the proofs which are omitted from the main body of the paper.

A.1 Proof of Lemma 1

Proof: Fix $i \in \mathcal{N}$ and $\theta^m, \theta^n \in \Theta$ arbitrarily. By definition, we obtain

$$\begin{aligned} \bar{x}_i^*(\theta^m) &= \sum_{\theta_{-i} \in \Theta^{N-1}} P^{N-1}(\theta_{-i}) x^*(\theta^m, \theta_{-i}) = \sum_{\theta_{-i}: \sum_{j \neq i} \theta_j \geqslant c(N) - \theta^m} P^{N-1}(\theta_{-i}); \\ \bar{x}_i^*(\theta^n) &= \sum_{\theta_{-i} \in \Theta^{N-1}} P^{N-1}(\theta_{-i}) x^*(\theta^n, \theta_{-i}) = \sum_{\theta_{-i}: \sum_{j \neq i} \theta_j \geqslant c(N) - \theta^n} P^{N-1}(\theta_{-i}). \end{aligned}$$

Observe that as $N \to \infty$,

$$\lim_{N \to \infty} (c(N)/N - \theta^m/N) = \lim_{N \to \infty} (c(N)/N - \theta^n/N) = \lim_{N \to \infty} c(N)/N.$$

because $\theta^1 < \lim_{N \to \infty} c(N)/N < \theta^M$. For each $\theta^m \in \Theta$ and $N \geq 2$, we have

$$\left\{ \theta_{-i} \in \Theta^{N-1} \middle| \sum_{j \neq i} \theta_j \geq c(N) - \theta^m \right\} = \left\{ \theta_{-i} \in \Theta^{N-1} \middle| \frac{1}{N} \sum_{j \neq i} \theta_j \geq \frac{c(N)}{N} - \frac{\theta^m}{N} \right\}$$

Therefore, for each $\theta^m \in \Theta$, as $N \to \infty$,

$$\sum_{\theta_{-i}: \sum_{j \neq i} \theta_j \geq c(N) - \theta^m} P^{N-1}(\theta_{-i}) \approx \sum_{\theta_{-i}: \frac{1}{N} \sum_{j \neq i} \theta_j \geq \frac{c(N)}{N}} P^{N-1}(\theta_{-i}),$$

which does not depend on θ^m. This implies $\lim_{N \to \infty} \bar{x}_i^*(\theta^m) = \lim_{N \to \infty} \bar{x}_i^*(\theta^n)$. ■

A.2 Proof of Lemma 2

Proof: For the ease of notation, we define

$$\Theta_*^N \equiv \left\{ \theta \in \Theta^N : \sum_{i \in \mathcal{N}} \theta_i \geq c(N) \right\},$$

and for each $\theta_i \in \Theta$,

$$\Theta_*^{N-1}(\theta_i) \equiv \left\{ \theta_{-i} \in \Theta^{N-1} : \sum_{j \neq i} \theta_j \geq c(N) - \theta_i \right\}.$$

Then we obtain the desired expression in Lemma 2 as follows:

$$
\begin{aligned}
\text{LHS} &= \lim_{N\to\infty} \sum_{\theta\in\Theta^N} P^N(\theta) x^*[N](\theta) \\
&= \lim_{N\to\infty} \sum_{\theta\in\Theta^N_*} P^N(\theta) \text{ (recall the definition of } \bar{x}^*(\theta)) \\
&= \lim_{N\to\infty} \sum_{\theta_i\in\Theta} \sum_{\theta_{-i}\in\Theta^{N-1}_*(\theta_i)} P(\theta_i)P^{N-1}(\theta_{-i}) \text{ } (\because \text{ types are independently distributed)} \\
&= \lim_{N\to\infty} \sum_{\theta_i\in\Theta} P(\theta_i)\bar{x}^*_i[N](\theta_i) \left(\because \bar{x}_i{}^*[N](\theta_i) = \sum_{\theta_{-i}\in\Theta^{N-1}_*} P^{N-1}(\theta_{-i})\right) \\
&= \lim_{N\to\infty} \bar{x}^*_i[N](\theta^M) \sum_{\theta_i\in\Theta} P(\theta_i) \left(\because \lim_{N\to\infty} \bar{x}^*_i[N](\theta^m) = \lim_{N\to\infty} \bar{x}^*_i[N](\theta^n) \text{ for any } m\neq n \text{ by Lemma 1}\right) \\
&= \lim_{N\to\infty} \bar{x}^*_i[N](\theta^M) = \text{RHS} \left(\because \sum_{\theta_i\in\Theta} P(\theta_i) = 1\right).
\end{aligned}
$$

■

A.3 Proof of Lemma 3

Proof: We take the expression of $\Pi'_{ea}(x^*)$ from Theorem 1:

$$
\begin{aligned}
\Pi'_{ea}(x^*) = \sum_{m=1}^{M-1} \bar{x}^*_i(\theta^m) &\left(N\theta^m \sum_{l=m}^{M} P(\theta^l) - N\theta^{m+1} \sum_{l=m+1}^{M} P(\theta^l) - c(N)P(\theta^m)\right) \\
&+ \bar{x}^*_i(\theta^M)\left(N\theta^M - c(N)\right) P(\theta^M).
\end{aligned}
$$

In particular, when $M = 2$,

$$\begin{aligned}
\Pi'_{ea}(x^*) &= \bar{x}_i^*(\theta^1)\left(N\theta^1\sum_{l=1}^{2}P(\theta^l) - N\theta^2P(\theta^2) - c(N)P(\theta^1)\right) + \bar{x}_i^*(\theta^2)\big(N\theta^2 - c(N)\big)P(\theta^2) \\
&= \bar{x}_i^*(\theta^1)\big(N\theta^1 - (1-\varepsilon)N\theta^2 - \varepsilon c(N)\big) + \bar{x}_i^*(\theta^2)(1-\varepsilon)\big(N\theta^2 - c(N)\big) \\
&= \sum_{k=k^*(N)}^{N-1} p^{N-1}(k)\big(N\theta^1 - (1-\varepsilon)N\theta^2 - \varepsilon c(N)\big) + \sum_{k=k^*(N)-1}^{N-1} p^{N-1}(k)(1-\varepsilon)\big(N\theta^2 - c(N)\big) \\
&\quad \text{(recall the formulas of } \bar{x}_i^*(\theta^1) \text{ and } \bar{x}_i^*(\theta^2)) \\
&= \sum_{k=k^*(N)}^{N-1} p^{N-1}(k)\Big(N\theta^1 - (1-\varepsilon)N\theta^2 - \varepsilon c(N) + (1-\varepsilon)N\theta^2 - (1-\varepsilon)c(N)\Big) \\
&\quad + p^{N-1}(k^*(N)-1)(1-\varepsilon)\big(N\theta^2 - c(N)\big) \\
&= \sum_{k=k^*(N)}^{N-1} p^{N-1}(k)\big(N\theta^1 - c(N)\big) + p^{N-1}(k^*(N)-1)(1-\varepsilon)\big(N\theta^2 - c(N)\big).
\end{aligned}$$

We know that $(k^*(N)-1)\theta^2 + (N-k^*(N)+1)\theta^1 < c(N)$. Plugging this inequality in $\Pi'_{ea}(x^*)$, we obtain

$$\begin{aligned}
\Pi'_{ea}(x^*) &< \big(N\theta^1 - (k^*(N)-1)\theta^2 - (N-k^*(N)+1)\theta^1\big)\sum_{k=k^*(N)}^{N-1} p^{N-1}(k) \\
&\quad + (1-\varepsilon)\big(N\theta^2 - (k^*(N)-1)\theta^2 - (N-k^*(N)+1)\theta^1\big)\cdot p^{N-1}(k^*(N)-1) \\
&= -(k^*(N)-1)(\theta^2-\theta^1)\sum_{k=k^*(N)}^{N-1} p^{N-1}(k) + (1-\varepsilon)(N-k^*(N)+1)(\theta^2-\theta^1)p^{N-1}(k^*(N)-1) \\
&= (\theta^2-\theta^1)\left(-(k^*(N)-1)\sum_{k=k^*(N)}^{N-1} p^{N-1}(k) + (1-\varepsilon)(N-k^*(N)+1)p^{N-1}(k^*(N)-1)\right).
\end{aligned}$$

Defining

$$\bar{\Pi}_{ea}(x^*) \equiv -(k^*(N)-1)\sum_{k=k^*(N)}^{N-1} p^{N-1}(k) + (1-\varepsilon)(N-k^*(N)+1)p^{N-1}(k^*(N)-1),$$

we rewrite the expression for the upper bound of $\Pi'_{ea}(x^*)$ as follows:

$$\Pi'_{ea}(x^*) < (\theta^2-\theta^1)\bar{\Pi}_{ea}(x^*).$$

■

A.4 Proof of the Assertion in Footnote 13

Proof: Recall that $x^*[N]$ is the efficient decision rule in an economy with N agents. In this example,

$$\lim_{N\to\infty}\sum_{\theta\in\Theta^N} P^N(\theta)x^*[N](\theta) = \lim_{N\to\infty}\sum_{k=k^*(N)}^{N} p^N(k)$$

where $p^N(k)$ is the probability that k out of N agents are of type θ^2.

Moreover, as $N\to\infty$, $p^N(k)$ can be approximated by $f^N(k)$, which is the probability density function of the normal distribution with mean $N(1-\varepsilon)$ and variance $N(1-\varepsilon)\varepsilon$. This approximation is formally called De Moivre-Laplace theorem. See, for example, Papoulis (1991) for the details. Therefore,

$$\lim_{N\to\infty}\sum_{\theta\in\Theta^N} P^N(\theta)x^*[N](\theta) = \lim_{N\to\infty}\sum_{k=k^*(N)}^{N} p^N(k) = \lim_{N\to\infty}\int_{k^*(N)}^{N} f^N(k)dk > 0$$

because $k^*(N) = \lceil(1-\varepsilon)N\rceil$, which is the mean of the approximated normal distribution. Hence, Theorem 2 applies so that we obtain $\lim_{N\to\infty}\Pi'_{ea}(x^*[N])/N < 0$. ■

A.5 Proof of Theorem 3

Proof: From IR_{θ^1}, we know that for each agent $i\in\mathcal{N}$,

$$\bar{t}_i(\theta^1) \leqslant \theta^1\bar{x}_i(\theta^1) = \bar{t}_i^T(\theta^1).$$

If $\theta^m > \theta^1$, or equivalently, $m > 1$, then adding IR_{θ^1} and $IC_{\theta^l\to\theta^{l-1}}$ for every $2\leqslant l\leqslant m$, we obtain that for each agent $i\in\mathcal{N}$,

$$\bar{t}_i(\theta^m) \leqslant \sum_{l=1}^{m}\theta^l\left(\bar{x}_i(\theta^l)-\bar{x}_i(\theta^{l-1})\right) = \bar{t}_i^T(\theta^m).$$

Therefore, each agent has the largest interim expected payment in the tight mechanism, and thus the tight mechanism maximizes the ex ante expected budget surplus. Finally, we obtain the expression for the ex ante budget surplus of the tight mechanism (x, t^T):

$$
\begin{aligned}
\Pi^T_{ea}(x) &= \sum_{i\in\mathcal{N}}\sum_{\theta_i\in\Theta} P(\theta_i)\vec{t}^T_i(\theta_i) - c(N)\sum_{\theta_i\in\Theta} P(\theta_i)\bar{x}_i(\theta_i) \\
&= \sum_{i\in\mathcal{N}}\sum_{m=1}^{M} P(\theta^m)\vec{t}^T_i(\theta^m) - c(N)\sum_{m=1}^{M} P(\theta^m)\bar{x}_i(\theta^m) \\
&= N\sum_{m=1}^{M} P(\theta^m)\sum_{l=1}^{m}\theta^l\left(\bar{x}_i(\theta^l)-\bar{x}_i(\theta^{l-1})\right) - c(N)\sum_{m=1}^{M} P(\theta^m)\bar{x}_i(\theta^m) \\
&\quad \left(\text{recall the definition of } \vec{t}^T_i(\theta^m)\right) \\
&= N\sum_{m=1}^{M}\theta^m\left(\bar{x}_i(\theta^m)-\bar{x}_i(\theta^{m-1})\right)\sum_{l=m}^{M} P(\theta^l) - c(N)\sum_{m=1}^{M} P(\theta^m)\bar{x}_i(\theta^m) \\
&= \sum_{m=1}^{M-1}\bar{x}_i(\theta^m)\left(N\theta^m\sum_{l=m}^{M}P(\theta^l) - N\theta^{m+1}\sum_{l=m+1}^{M}P(\theta^l) - c(N)P(\theta^m)\right) \\
&\quad +\bar{x}_i(\theta^M)\left(N\theta^M - c(N)\right)P(\theta^M).
\end{aligned}
$$

■

A.6 Proof of Theorem 4

Proof: We take the expression for $\Pi^T_{ea}(x[N])$ from Theorem 3:

$$
\Pi^T_{ea}(x[N]) = \sum_{m=1}^{M-1}\bar{x}_i[N](\theta^m)\left(N\theta^m\sum_{l=m}^{M}P(\theta^l) - N\theta^{m+1}\sum_{l=m+1}^{M}P(\theta^l) - c(N)P(\theta^m)\right),
$$
$$
+\,\bar{x}_i[N](\theta^M)\left(N\theta^M - c(N)\right)P(\theta^M),
$$

where $\bar{x}_i[N](\theta^m) = \sum_{\theta_{-i}\in\Theta^{N-1}} P^{N-1}(\theta_{-i})x[N](\theta^m,\theta_{-i})$ for each $\theta^m \in \Theta$. In particular, when $M = 2$,

$$
\Pi^T_{ea}(x[N]) = \bar{x}_i[N](\theta^1)\left(N\theta^1 - N\theta^2(1-\varepsilon) - \varepsilon c(N)\right) + (1-\varepsilon)\bar{x}_i[N](\theta^2)\left(N\theta^2 - c(N)\right).
$$

Hence,

$$\begin{aligned}
\frac{\Pi^T_{ea}(x[N])}{N} &= \bar{x}_i[N](\theta^1)\left(\theta^1-\theta^2(1-\varepsilon)-\frac{\varepsilon c(N)}{N}\right)+(1-\varepsilon)\bar{x}_i[N](\theta^2)\left(\theta^2-\frac{c(N)}{N}\right)\\
&= \sum_{\theta_{-i}\in\Theta^{N-1}} P^{N-1}(\theta_{-i})x[N](\theta^1,\theta_{-i})\left(\theta^1-\theta^2(1-\varepsilon)-\frac{\varepsilon c(N)}{N}\right)\\
&\quad +(1-\varepsilon)\sum_{\theta_{-i}\in\Theta^{N-1}} P^{N-1}(\theta_{-i})x[N](\theta^2,\theta_{-i})\left(\theta^2-\frac{c(N)}{N}\right)\\
&\quad (\text{ recall the definition of } \bar{x}_i[N](\theta^1) \text{ and } \bar{x}_i[N](\theta^2))\\
&= \sum_{k=0}^{N-1} p^{N-1}(k)x[N](k)\left(\theta^1-\theta^2(1-\varepsilon)-\frac{\varepsilon c(N)}{N}\right)\\
&\quad +(1-\varepsilon)\sum_{k=0}^{N-1} p^{N-1}(k)x[N](k+1)\left(\theta^2-\frac{c(N)}{N}\right)\\
&\quad (x \text{ is anonymous by Assumption 2 })
\end{aligned}$$

where $p^{N-1}(k)$ is the probability that k out of $(N-1)$ agents are of type θ^2 and $x[N](k)$ is the probability of public good provision when there are k agents who are of type θ^2.

The expression can be further rewritten as follows:

$$\begin{aligned}
&\frac{\Pi^T_{ea}(x[N])}{N}\\
=& \sum_{k=1}^{N-1} x[N](k)\left[p^{N-1}(k)\left(\theta^1-\theta^2(1-\varepsilon)-\frac{\varepsilon c(N)}{N}\right)+p^{N-1}(k-1)(1-\varepsilon)\left(\theta^2-\frac{c(N)}{N}\right)\right]\\
&+(1-\varepsilon)p^{N-1}(N-1)x[N](N)\left(\theta^2-\frac{c(N)}{N}\right)+p^{N-1}(0)x[N](0)\left(\theta^1-\theta^2(1-\varepsilon)-\frac{\varepsilon c(N)}{N}\right)\\
=& \sum_{k=k^*(N)}^{N-1} x[N](k)\left[p^{N-1}(k)\left(\theta^1-\theta^2(1-\varepsilon)-\frac{\varepsilon c(N)}{N}\right)+p^{N-1}(k-1)(1-\varepsilon)\left(\theta^2-\frac{c(N)}{N}\right)\right]\\
&+(1-\varepsilon)^N x[N](N)\left(\theta^2-\frac{c(N)}{N}\right),
\end{aligned}$$

because $x[N](k)=0$ whenever $k<k^*(N)$ by Assumption 1.

Hence,

$$\begin{aligned}
&\lim_{N\to\infty}\frac{\Pi^T_{ea}(x[N])}{N}\\
=\ &\lim_{N\to\infty}\sum_{k=k^*(N)}^{N-1} x[N](k)\left[p^{N-1}(k)\left(\theta^1-\theta^2(1-\varepsilon)-\frac{\varepsilon c(N)}{N}\right)+p^{N-1}(k-1)(1-\varepsilon)\left(\theta^2-\frac{c(N)}{N}\right)\right]\\
&+\lim_{N\to\infty}(1-\varepsilon)^N x[N](N)\left(\theta^2-\frac{c(N)}{N}\right)\\
=\ &\lim_{N\to\infty}\sum_{k=k^*(N)}^{N-1} x[N](k)\left[p^{N-1}(k)\left(\theta^1-\theta^2(1-\varepsilon)-\frac{\varepsilon c(N)}{N}\right)+p^{N-1}(k-1)(1-\varepsilon)\left(\theta^2-\frac{c(N)}{N}\right)\right],
\end{aligned}$$

because $(1-\varepsilon)^N\to 0$ as $N\to\infty$. For simplicity, let

$$\begin{aligned}
A &\equiv \theta^1-\theta^2(1-\varepsilon)-\varepsilon\frac{c(N)}{N}=\varepsilon\left(\theta^1-\frac{c(N)}{N}\right)+(1-\varepsilon)(\theta^1-\theta^2)<0;\\
B &\equiv (1-\varepsilon)\left(\theta^2-\frac{c(N)}{N}\right)\geq 0.
\end{aligned}$$

Observe that

$$\begin{aligned}
|A|-|B| &= -A-B\\
&= -\varepsilon\left(\theta^1-\frac{c(N)}{N}\right)-(1-\varepsilon)(\theta^1-\theta^2)-(1-\varepsilon)\left(\theta^2-\frac{c(N)}{N}\right)\\
&= -\varepsilon\left(\theta^1-\frac{c(N)}{N}\right)-(1-\varepsilon)\left(\theta^1-\frac{c(N)}{N}\right)\\
&= -\left(\theta^1-\frac{c(N)}{N}\right)>0.
\end{aligned}$$

The expression of $\lim_{N\to\infty}\Pi^T_{ea}(x[N])/N$ is rewritten as follows:

$$\lim_{N\to\infty}\frac{\Pi^T_{ea}(x[N])}{N}=\lim_{N\to\infty}\sum_{k=k^*(N)}^{N-1} x[N](k)\left[p^{N-1}(k)A+p^{N-1}(k-1)B\right].$$

Since $\lim_{N\to\infty}k^*(N)/N=\alpha$ for some $\alpha\in(0,1]$, $k^*(N)$ can be approximated by αN for N large enough. So,

$$\lim_{N\to\infty}\frac{\Pi^T_{ea}(x[N])}{N}=\lim_{N\to\infty}\sum_{k=\lceil\alpha N\rceil}^{N-1} x[N](k)\left[p^{N-1}(k)A+p^{N-1}(k-1)B\right],$$

where $\lceil\alpha N\rceil$ denotes the least integer greater than or equal to αN for every N. Moreover, for any N large enough and any integer k, if $\lceil\alpha N\rceil\leq k\leq N-1$, there exists $\beta_{k,N}\in(0,1)$ such that k is approximated by $\beta_{k,N}N$. We thus obtain that for N large enough,

$$\frac{\beta_{k,N}N}{\beta_{k,N}N-1}\approx 1.$$

We also know that as $N \to \infty$, $p^{N-1}(k)$ can be approximated by $f^{N-1}(k)$, which is the probability density function of the normal distribution with mean $(N-1)(1-\varepsilon)$ and variance $(N-1)(1-\varepsilon)\varepsilon$. This approximation is formally called De Moivre-Laplace theorem. See, for example, Papoulis (1991) for the details. Since $f^{N-1}(\cdot)$ is continuous, for any N large enough and any integer k, if $\lceil \alpha N \rceil \leq k \leq N-1$,

$$f^{N-1}(\beta_{k,N}N) \approx f^{N-1}(\beta_{k,N}N - 1).$$

This implies that for any integer k and N, if $k^*(N) \leq k \leq N-1$,

$$\lim_{N\to\infty} p^{N-1}(k) = \lim_{N\to\infty} p^{N-1}(k-1).$$

In other words, the difference in the binomial probabilities of k and $(k-1)$ is negligible in large economies. We thus replace $p^{N-1}(k-1)$ by $p^{N-1}(k)$ in the expression of $\lim_{N\to\infty} \Pi^T_{ea}(x[N])/N$ and rewrite it as follows:

$$\begin{aligned}
\lim_{N\to\infty} \frac{\Pi^T_{ea}(x[N])}{N} &= \lim_{N\to\infty} \sum_{k=k^*(N)}^{N-1} x[N](k) \left[p^{N-1}(k)A + p^{N-1}(k)B\right] \\
&= \lim_{N\to\infty} \sum_{k=k^*(N)}^{N-1} p^{N-1}(k)x[N](k)(A+B) \\
&= (A+B) \lim_{N\to\infty} \sum_{k=k^*(N)}^{N-1} p^{N-1}(k)x[N](k).
\end{aligned}$$

Moreover, for N large enough, we have $N/(N-1) \approx 1$ so that we can replace $(N-1)$ by N and rewrite the expression of $\Pi^T_{ea}(x[N])/N$ as follows:

$$\lim_{N\to\infty} \frac{\Pi^T_{ea}(x[N])}{N} = (A+B) \lim_{N\to\infty} \sum_{k=k^*(N)}^{N} p^{N}(k)x[N](k),$$

where $p^N(k)$ is the probability that k out of N agents are of type θ^2. Recall that $A+B = -(-A-B) < 0$. Since $\{x[N]\}_N$ is a sequence of mechanisms satisfying BIC, IIR, and BB, by Corollary 3, we must have $\Pi^T_{ea}(x[N]) \geq 0$ for each N. This together with Assumption 1 implies that we obtain the following equivalence:

$$\lim_{N\to\infty} \Pi^T_{ea}(x[N])/N \geq 0$$

if and only if

$$\lim_{N\to\infty}\sum_{k=1}^{N} p^N(k)x[N](k) = 0.$$

This concludes that the ex ante probability that the public good is provided converges to zero as the population size of the economy gets large. ■

B Appendix to Chapter 2

B.1 Proof of Claim 4

Proof: We first identify the worst-off type for each agent by checking the following cases.

Case 1: $0 \leq \theta_1 \leq (1-\gamma_1)/(1-\gamma_2)$

We compute the following.

$$\begin{aligned}
& \mathbb{E}_{\theta_2}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right) \\
= & \int_0^{\frac{1-\gamma_2}{1-\gamma_1}\theta_1} \tilde{u}_1(\theta_1,\theta_2)d\theta_2 + \int_{\frac{1-\gamma_2}{1-\gamma_1}\theta_1}^{1} \tilde{u}_2(\theta_1,\theta_2)d\theta_2 \\
= & \int_0^{\frac{1-\gamma_2}{1-\gamma_1}\theta_1} (\theta_1+\gamma_1\theta_2)d\theta_2 + \int_{\frac{1-\gamma_2}{1-\gamma_1}\theta_1}^{1} (\gamma_2\theta_1+\theta_2)d\theta_2 \\
= & \frac{1-\gamma_2}{1-\gamma_1}(\theta_1)^2 + \frac{1}{2}\frac{\gamma_1(1-\gamma_2)^2}{(1-\gamma_1)^2}(\theta_1)^2 + \gamma_2\theta_1\left(1-\frac{1-\gamma_2}{1-\gamma_1}\theta_1\right) + \frac{1}{2}\left[1-\left(\frac{1-\gamma_2}{1-\gamma_1}\right)^2(\theta_1)^2\right] \\
= & \frac{1}{2} + \gamma_2\theta_1 + \frac{1}{2}\frac{(1-\gamma_2)^2}{1-\gamma_1}(\theta_1)^2. \qquad (37)
\end{aligned}$$

Hence, the objective function becomes

$$\begin{aligned}
\mathbb{E}_{\theta_2}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right) - U_1^O(\theta_1) &= \frac{1}{2} + \gamma_2\theta_1 + \frac{1}{2}\frac{(1-\gamma_2)^2}{1-\gamma_1}(\theta_1)^2 - \int_0^1 \tilde{u}_1(\theta_1,\theta_2)d\theta_2 \\
&= \frac{1}{2} + \gamma_2\theta_1 + \frac{1}{2}\frac{(1-\gamma_2)^2}{1-\gamma_1}(\theta_1)^2 - \theta_1 - \frac{1}{2}\gamma_1 \\
&= \frac{1}{2}\frac{(1-\gamma_2)^2}{1-\gamma_1}\left(\theta_1 - \frac{1-\gamma_1}{1-\gamma_2}\right)^2.
\end{aligned}$$

So, when $\theta_1 = (1-\gamma_1)/(1-\gamma_2)$, the objective function attains its minimum, which is zero. So, in this case, the seller's worst-off type is $\theta_1^w = (1-\gamma_1)/(1-\gamma_2)$.

Case 2: $(1-\gamma_1)/(1-\gamma_2) < \theta_1 \leq 1$

We compute the following:

$$\mathbb{E}_{\theta_2}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right) - U_1^O(\theta_1) = \int_0^1 \tilde{u}_1(\theta_1,\theta_2)d\theta_2 - \int_0^1 \tilde{u}_1(\theta_1,\theta_2)d\theta_2 = 0.$$

Therefore, in this case, the seller's worst-off type is $\theta_1^w = 1$.

We compute the expected loss for his worst-off type θ_1^w below:

$$\begin{aligned}
L_1 &\equiv U_1^O(\theta_1^w) - U_1^G(\theta_1^w) \\
&= -\left[\mathbb{E}_{\theta_2}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1^w,\theta_2),\theta_1^w,\theta_2)\right) - U_1^O(\theta_1^w)\right] + \frac{1}{2}\mathbb{E}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right) \\
&= 0 + \frac{1}{2}\mathbb{E}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right) \\
&= \frac{1}{2}\mathbb{E}_{\theta_2}\left[\mathbb{E}_{\theta_1}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right)\right],
\end{aligned}$$

where $\mathbb{E}_{\theta_1}$ denotes the expectation operator over Θ_1. Note that for each $\theta_2 \in \Theta_2$,

$$\begin{aligned}
&\mathbb{E}_{\theta_1}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right) \\
= &\int_0^{\frac{1-\gamma_1}{1-\gamma_2}\theta_2} \tilde{u}_2(\theta_1,\theta_2)d\theta_1 + \int_{\frac{1-\gamma_1}{1-\gamma_2}\theta_2}^1 \tilde{u}_1(\theta_1,\theta_2)d\theta_1 \\
= &\int_0^{\frac{1-\gamma_1}{1-\gamma_2}\theta_2} (\gamma_2\theta_1+\theta_2)d\theta_1 + \int_{\frac{1-\gamma_1}{1-\gamma_2}\theta_2}^1 (\theta_1+\gamma_1\theta_2)d\theta_1 \\
= &\frac{1}{2}\gamma_2\left(\frac{1-\gamma_1}{1-\gamma_2}\theta_2\right)^2 + \frac{1-\gamma_1}{1-\gamma_2}(\theta_2)^2 + \frac{1}{2}\left(1-\left(\frac{1-\gamma_1}{1-\gamma_2}\theta_2\right)^2\right) + \gamma_1\theta_2\left(1-\frac{1-\gamma_1}{1-\gamma_2}\theta_2\right) \\
= &\frac{1}{2} + \gamma_1\theta_2 + \frac{1}{2}\frac{(1-\gamma_1)^2}{1-\gamma_2}(\theta_2)^2. \qquad (38)
\end{aligned}$$

Therefore, we compute the expected loss for the seller's worst-off type θ_1^w:

$$\begin{aligned}
L_1 &= \frac{1}{2}\mathbb{E}_{\theta_2}\left[\mathbb{E}_{\theta_1}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right)\right] \\
&= \frac{1}{2}\int_0^1 \left[\frac{1}{2} + \gamma_1\theta_2 + \frac{1}{2}\frac{(1-\gamma_1)^2}{1-\gamma_2}(\theta_2)^2\right] d\theta_2 \\
&= \frac{1}{4} + \frac{1}{4}\gamma_1 + \frac{1}{12}\frac{(1-\gamma_1)^2}{1-\gamma_2}. \qquad (39)
\end{aligned}$$

Since $\gamma_1 > 0$ and $\gamma_2 < 1$, we obtain $L_1 > 0$, which implies that the seller is worse off after participating in the mechanism. On the other hand, we obtain the buyer's worst-off type θ_2^w from participating in the generalized two-stage Groves mechanism:

$$\begin{aligned}\theta_2^w &\in \arg\min_{\theta_2 \in \Theta_2} \left[\mathbb{E}_{\theta_1}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right) - U_2^O(\theta_2)\right] \\ &= \arg\min_{\theta_2 \in \Theta_2} \left[\frac{1}{2} + \gamma_1\theta_2 + \frac{1}{2}\frac{(1-\gamma_1)^2}{1-\gamma_2}(\theta_2)^2 - 0\right],\end{aligned}$$

where the equality follows from (38). It is easy to see that the buyer's worst-off type is $\theta_2^w = 0$. So, we compute the expected loss for his worst-off type $\theta_2^w = 0$ as follows:

$$\begin{aligned}L_2 &\equiv U_2^O(\theta_2^w) - U_2^G(\theta_2^w) \\ &= -\left[\mathbb{E}_{\theta_1}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2^w),\theta_1,\theta_2^w)\right) - U_2^O(\theta_2^w)\right] + \frac{1}{2}\mathbb{E}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right) \\ &= -\frac{1}{2} + \frac{1}{2}\mathbb{E}\left(\sum_{j=1}^{2} u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right) \\ &= -\frac{1}{4} + \frac{1}{4}\gamma_1 + \frac{1}{12}\frac{(1-\gamma_1)^2}{1-\gamma_2} \text{ (recall (39))}\end{aligned}$$

So, the total expected loss is

$$\begin{aligned}L_1 + L_2 &= \frac{1}{4} + \frac{1}{4}\gamma_1 + \frac{1}{12}\frac{(1-\gamma_1)^2}{1-\gamma_2} - \frac{1}{4} + \frac{1}{4}\gamma_1 + \frac{1}{12}\frac{(1-\gamma_1)^2}{1-\gamma_2} \\ &= \frac{1}{2}\gamma_1 + \frac{1}{6}\frac{(1-\gamma_1)^2}{1-\gamma_2} \\ &> 0,\end{aligned}$$

where the last inequality follows because $\gamma_1 > 0$ and $0 < \gamma_2 < 1$. This completes the proof of Claim 4. ■

B.2 Proof of Claim 5

Proof: We compute the seller's worst-off type from participating in the generalized two-stage Groves mechanism.

$$\begin{aligned}
\theta_1^w &\in \arg\min_{\theta_1\in\Theta_1}\left[\mathbb{E}_{\theta_2}\left(\sum_{j=1}^{2}u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right)-U_1^O(\theta_1)\right]\\
&= \arg\min_{\theta_1\in\Theta_1}\left[\int_0^{\frac{1-\gamma_2}{1-\gamma_1}\theta_1}\tilde{u}_1(\theta_1,\theta_2)d\theta_2+\int_{\frac{1-\gamma_2}{1-\gamma_1}\theta_1}^{1}\tilde{u}_2(\theta_1,\theta_2)d\theta_2-\int_0^1\tilde{u}_1(\theta_1,\theta_2)d\theta_2\right]\\
&= \arg\min_{\theta_1\in\Theta_1}\left[\frac{1}{2}+\gamma_2\theta_1+\frac{1}{2}\frac{(1-\gamma_2)^2}{1-\gamma_1}(\theta_1)^2-\theta_1-\frac{1}{2}\gamma_1\right]\ \text{(recall (37))}\\
&= \arg\min_{\theta_1\in\Theta_1}\left[\frac{1}{2}\frac{(1-\gamma_2)^2}{1-\gamma_1}\left(\theta_1-\frac{1-\gamma_1}{1-\gamma_2}\right)^2\right].
\end{aligned}$$

Note that $0<\gamma_1<\gamma_2<1$ implies $(1-\gamma_1)/(1-\gamma_2)>1$. Hence, the seller's worst-off type is $\theta_1^w=1$. We compute the expected loss for his worst-off type as follows:

$$\begin{aligned}
L_1 &\equiv U_1^O(\theta_1^w)-U_1^G(\theta_1^w)\\
&= -\left[\mathbb{E}_{\theta_2}\left(\sum_{j=1}^{2}u_j(x^*(\theta_1^w,\theta_2),\theta_1^w,\theta_2)\right)-U_1^O(\theta_1^w)\right]+\frac{1}{2}\mathbb{E}\left(\sum_{j=1}^{2}u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right)\\
&= -\frac{1}{2}\frac{(1-\gamma_2)^2}{1-\gamma_1}\left(1-\frac{1-\gamma_1}{1-\gamma_2}\right)^2+\frac{1}{2}\mathbb{E}\left(\sum_{j=1}^{2}u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right).
\end{aligned}$$

We further compute the following:

$$\begin{aligned}
\frac{1}{2}\mathbb{E}\left(\sum_{j=1}^{2}u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right) &= \frac{1}{2}\mathbb{E}_{\theta_1}\left[\mathbb{E}_{\theta_2}\left(\sum_{j=1}^{2}u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right)\right]\\
&= \frac{1}{2}\int_0^1\left[\int_0^{\frac{1-\gamma_2}{1-\gamma_1}\theta_1}\tilde{u}_1(\theta_1,\theta_2)d\theta_2+\int_{\frac{1-\gamma_2}{1-\gamma_1}\theta_1}^{1}\tilde{u}_2(\theta_1,\theta_2)d\theta_2\right]d\theta_1\\
&= \frac{1}{2}\int_0^1\left[\frac{1}{2}+\gamma_2\theta_1+\frac{1}{2}\frac{(1-\gamma_2)^2}{1-\gamma_1}(\theta_1)^2\right]d\theta_1\ \text{(recall (37))}\\
&= \frac{1}{4}+\frac{1}{4}\gamma_2+\frac{1}{12}\frac{(1-\gamma_2)^2}{1-\gamma_1}. \qquad (40)
\end{aligned}$$

Therefore, the expected loss for the seller's worst-off type is

$$\begin{aligned}
L_1 &= -\frac{1}{2}\frac{(1-\gamma_2)^2}{1-\gamma_1}\left(1-\frac{1-\gamma_1}{1-\gamma_2}\right)^2+\frac{1}{2}\mathbb{E}\left(\sum_{j=1}^{2}u_j(x^*(\theta_1,\theta_2),\theta_1,\theta_2)\right)\\
&= -\frac{1}{2}\frac{(\gamma_2-\gamma_1)^2}{1-\gamma_1}+\frac{1}{4}+\frac{1}{4}\gamma_2+\frac{1}{12}\frac{(1-\gamma_2)^2}{1-\gamma_1}.
\end{aligned}$$

On the other hand, the buyer's worst-off type from participating in the generalized two-stage Groves mechanism is given as follows:

$$\theta_2^w \in \arg\min_{\theta_2 \in \Theta_2} \left[\mathbb{E}_{\theta_1} \left(\sum_{j=1}^{2} u_j(x^*(\theta_1, \theta_2), \theta_1, \theta_2) \right) - U_2^O(\theta_2) \right] = \arg\min_{\theta_2 \in \Theta_2} \mathbb{E}_{\theta_1} \left(\sum_{j=1}^{2} u_j(x^*(\theta_1, \theta_2), \theta_1, \theta_2) \right).$$

We identify the worst-off type for each agent by the following cases.

Case 1: $0 < \theta_2 < (1-\gamma_2)/(1-\gamma_1)$

then

$$\begin{aligned} \mathbb{E}_{\theta_1} \left[\sum_{j=1}^{2} u_j(x^*(\theta_2, \theta_1), \theta_2, \theta_1) \right] &= \int_0^{\frac{1-\gamma_1}{1-\gamma_2}\theta_2} \tilde{u}_2(\theta_1, \theta_2) d\theta_1 + \int_{\frac{1-\gamma_1}{1-\gamma_2}\theta_2}^{1} \tilde{u}_1(\theta_1, \theta_2) d\theta_1 \\ &= \frac{1}{2} + \gamma_1 \theta_2 + \frac{1}{2} \frac{(1-\gamma_1)^2}{1-\gamma_2} (\theta_2)^2 \text{ (recall (38)).} \end{aligned}$$

It is easy to see that $\theta_2 = 0$ achieves its minimum, which is 1/2.

Case 2: $(1-\gamma_2)/(1-\gamma_1) \leq \theta_2 \leq 1$

then

$$\mathbb{E}_{\theta_1} \left[\sum_{j=1}^{2} u_j(x^*(\theta_2, \theta_1), \theta_2, \theta_1) \right] = \int_0^1 \tilde{u}_2(\theta_1, \theta_2) d\theta_1 = \int_0^1 (\gamma_2 \theta_1 + \theta_2) d\theta_1 = \frac{1}{2}\gamma_2 + \theta_2.$$

Clearly, $\theta_2 = (1-\gamma_2)/(1-\gamma_1)$ achieves its minimum, which is $\gamma_2/2 + (1-\gamma_2)/(1-\gamma_1)$.

Since

$$\frac{1}{2} - \left[\frac{1}{2}\gamma_2 + \frac{1-\gamma_2}{1-\gamma_1} \right] = -\frac{(1-\gamma_2)(1+\gamma_1)}{2(1-\gamma_1)} < 0,$$

the buyer's worst-off type is $\theta_2^w = 0$. We compute the expected loss for his worst-off type.

$$\begin{aligned} L_2 &\equiv U_2^O(\theta_2^w) - U_2^G(\theta_2^w) \\ &= -\left[\mathbb{E}_{\theta_1} \left(\sum_{j=1}^{2} u_j(x^*(\theta_1, \theta_2^w), \theta_1, \theta_2^w) \right) - U_2^O(\theta_2^w) \right] + \frac{1}{2}\mathbb{E} \left(\sum_{j=1}^{2} u_j(x^*(\theta_1, \theta_2), \theta_1, \theta_2) \right) \\ &= -\frac{1}{2} + \frac{1}{2}\mathbb{E} \left(\sum_{j=1}^{2} u_j(x^*(\theta_1, \theta_2), \theta_1, \theta_2) \right) \\ &= -\frac{1}{2} + \frac{1}{4} + \frac{1}{4}\gamma_2 + \frac{1}{12} \frac{(1-\gamma_2)^2}{1-\gamma_1} \text{ (recall (40)).} \end{aligned}$$

Therefore, the total expected loss is

$$\begin{aligned} L_1 + L_2 &= -\frac{1}{2} \frac{(\gamma_2-\gamma_1)^2}{1-\gamma_1} + \frac{1}{4} + \frac{1}{4}\gamma_2 + \frac{1}{12} \frac{(1-\gamma_2)^2}{1-\gamma_1} - \frac{1}{4} + \frac{1}{4}\gamma_2 + \frac{1}{12} \frac{(1-\gamma_2)^2}{1-\gamma_1} \\ &= \frac{1}{2(1-\gamma_1)} \left[\gamma_2(1-\gamma_1) - (\gamma_2-\gamma_1)^2 \right] + \frac{1}{6} \frac{(1-\gamma_2)^2}{1-\gamma_1}. \end{aligned}$$

Since $\gamma_2 > \gamma_2 - \gamma_1$ and $1 - \gamma_1 > \gamma_2 - \gamma_1$, we obtain $L_1 + L_2 > 0$. This completes the proof of Claim 5. ■

B.3 Proof of Claim 6

Proof: We divide our argument into the following two cases.

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

From Figure 1, we know that if $\theta_1 = 1$ and $\theta_2 = 0$, it is efficient not to trade and the buyer's transfer in the generalized two-stage Groves mechanism is given as follows:

$$\begin{aligned} & t_2^G(1,0;u_1,u_2) \\ = \; & u_1 - \frac{1}{2} h_2(1,0) \\ = \; & u_1 - \frac{1}{2}\left[\sum_{j=1}^{2} u_j\left(x^*(1,0),1,0\right) - \mathbb{E}_{\theta_1}\left(\sum_{j=1}^{2} u_j\left(x^*(\theta_1,0),\theta_1,0\right)\right) + \mathbb{E}_{\theta_2}\left(\sum_{j=1}^{2} u_j\left(x^*(1,\theta_2),1,\theta_2\right)\right)\right] \\ = \; & \tilde{u}_1(1,0) - \frac{1}{2}\left(\tilde{u}_1(1,0) - \int_0^1 \tilde{u}_1(\theta_1,0)d\theta_1 + \int_0^1 \tilde{u}_1(1,\theta_2)d\theta_2\right), \end{aligned}$$

where the third equality follows because $u_1 = u_1(x^*(1,0),1,0) = \tilde{u}_1(1,0)$, $u_2 = u_2(x^*(1,0),1,0) = 0$ and $x^*(\theta_1,0) = x^*(1,\theta_2) = 0$ for any $\theta_1, \theta_2 \in [0,1]$. Plugging the linear valuations in $t_2^G(1,0;u_1,u_2)$ above, we obtain

$$t_2^G(1,0;u_1,u_2) = 1 - \frac{1}{2}\left(1 - \int_0^1 \theta_1 d\theta_1 + \int_0^1 (1+\gamma_1\theta_2)d\theta_2\right) = \frac{1}{4}(1-\gamma_1) > 0,$$

where the last strict inequality follows because $1 > \gamma_1$ in Case (i). Hence, in the type profile $(\theta_1, \theta_2) = (1,0)$, the buyer receives positive subsidy under no trade, contradicting NTNP.

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

From Figure 2, we know that if $\theta_1 = 1$ and $\theta_2 = 0$, it is efficient not to trade. We then compute the buyer's transfer in the generalized two-stage Groves mechanism:

$$\begin{aligned}
& t_2^G(1,0;u_1,u_2) \\
= \;& u_1 - \frac{1}{2}\left[\sum_{j=1}^{2} u_j\left(x^*(1,0),1,0\right) - \mathbb{E}_{\theta_1}\left(\sum_{j=1}^{2} u_j\left(x^*(\theta_1,0),\theta_1,0\right)\right) + \mathbb{E}_{\theta_2}\left(\sum_{j=1}^{2} u_j\left(x^*(1,\theta_2),1,\theta_2\right)\right)\right] \\
= \;& \tilde{u}_1(1,0) - \frac{1}{2}\left(\tilde{u}_1(1,0) - \int_0^1 \tilde{u}_1(\theta_1,0)d\theta_1 + \int_0^{\frac{1-\gamma_2}{1-\gamma_1}} \tilde{u}_1(1,\theta_2)d\theta_2 + \int_{\frac{1-\gamma_2}{1-\gamma_1}}^1 \tilde{u}_2(1,\theta_2)d\theta_2\right),
\end{aligned}$$

where the last equality follows because $u_1 = u_1(x^*(1,0),1,0) = \tilde{u}_1(1,0)$, $u_2 = u_2(x^*(1,0),1,0) = 0$, $x^*(\theta_1,0) = 0$ for any $\theta_1 \in [0,1]$, $x^*(1,\theta_2) = 0$ if $\theta_2 < (1-\gamma_2)/(1-\gamma_1)$ and $x^*(1,\theta_2) = 1$ otherwise. Plugging the linear valuations in $t_2^G(1,0;u_1,u_2)$, we obtain

$$\begin{aligned}
t_2^G(1,0;u_1,u_2) &= 1 - \frac{1}{2}\left(1 - \int_0^1 \theta_1 d\theta_1 + \int_0^{\frac{1-\gamma_2}{1-\gamma_1}} (1+\gamma_1\theta_2)d\theta_2 + \int_{\frac{1-\gamma_2}{1-\gamma_1}}^1 (\theta_2+\gamma_2)d\theta_2\right) \\
&= 1 - \frac{1}{2}\left[1 - \frac{1}{2} + \frac{1-\gamma_2}{1-\gamma_1} + \frac{\gamma_1}{2}\left(\frac{1-\gamma_2}{1-\gamma_1}\right)^2 + \frac{1}{2} - \frac{1}{2}\left(\frac{1-\gamma_2}{1-\gamma_1}\right)^2 + \gamma_2 - \gamma_2\frac{1-\gamma_2}{1-\gamma_1}\right].
\end{aligned}$$

After rearranging the terms above, we simplify its expression:

$$t_2^G(1,0;u_1,u_2) = 1 - \frac{1}{2}\left(1 + \gamma_2 + \frac{1}{2}\frac{(1-\gamma_2)^2}{1-\gamma_1}\right),$$

which is strictly decreasing in γ_1. Since $\gamma_1 < \gamma_2$, then for any $\gamma_2 \in (0,1)$, $t_2^G(1,0;u_1,u_2)$ reaches its greatest lower bound when $\gamma_1 = \gamma_2$, i.e.,

$$t_2^G(1,0;u_1,u_2) > 1 - \frac{1}{2}\left(1 + \gamma_2 + \frac{1}{2}\frac{(1-\gamma_2)^2}{1-\gamma_2}\right) = \frac{1}{4}(1-\gamma_2) > 0,$$

where the last strict inequality holds because $1 > \gamma_2$ in Case (ii). Therefore, we conclude $t_2^G(1,0;u_1,u_2) > 0$, implying that, in the type profile $(\theta_1,\theta_2) = (1,0)$, the buyer receives positive subsidies under no trade. Hence, NTNP is violated. This completes the proof. ■

B.4 Proof of Claim 7

Proof: We divide our argument into two cases.

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

Fix $\hat{\theta}_1 \in [0,1]$ and let $\alpha, \beta \in [0,1]$ be two distinct types of the buyer such that $\alpha > \beta$ and $x^*(\hat{\theta}_1,\alpha) = x^*(\hat{\theta}_1,\beta) = 1$. Then, the difference between the buyer's transfer under

$(\hat{\theta}_1, \alpha)$ and that under $(\hat{\theta}_1, \beta)$ is computed below:

$$
\begin{aligned}
& t_2^G(\hat{\theta}_1, \alpha; u_1^\alpha, u_2^\alpha) - t_2^G(\hat{\theta}_1, \beta; u_1^\beta, u_2^\beta) \\
= \;& u_1^\alpha - \frac{1}{2}\left[\tilde{u}_2(\hat{\theta}_1, \alpha) - \mathbb{E}_{\theta_1}\left(\sum_{i=1}^{2} u_j(x^*(\theta_1, \alpha), \theta_1, \alpha)\right) + \mathbb{E}_{\theta_2}\left(\sum_{j=1}^{2} u_j(x^*(\hat{\theta}_1, \theta_2), \hat{\theta}_1, \theta_2)\right)\right] \\
& -u_1^\beta + \frac{1}{2}\left[\tilde{u}_2(\hat{\theta}_1, \beta) - \mathbb{E}_{\theta_1}\left(\sum_{i=1}^{2} u_j(x^*(\theta_1, \beta), \theta_1, \beta)\right) + \mathbb{E}_{\theta_2}\left(\sum_{j=1}^{2} u_j(x^*(\hat{\theta}_1, \theta_2), \hat{\theta}_1, \theta_2)\right)\right] \\
= \;& -\frac{1}{2}\left[\alpha - \beta - \mathbb{E}_{\theta_1}\left(\sum_{i=1}^{2} u_j(x^*(\theta_1, \alpha), \theta_1, \alpha)\right) + \mathbb{E}_{\theta_1}\left(\sum_{i=1}^{2} u_j(x^*(\theta_1, \beta), \theta_1, \beta)\right)\right], \qquad (41)
\end{aligned}
$$

where the second equality follows because $x^*(\hat{\theta}_1, \alpha) = x^*(\hat{\theta}_1, \beta) = 1$ implies $u_1^\alpha = u_1^\beta = 0$ and $\tilde{u}_2(\hat{\theta}_1, \alpha) - \tilde{u}_2(\hat{\theta}_1, \beta) = \alpha + \gamma_2\hat{\theta}_1 - \beta - \gamma_2\hat{\theta}_1 = \alpha - \beta$.

Moreover, we compute the following term in the above expression:

$$
\begin{aligned}
& -\mathbb{E}_{\theta_1}\left(\sum_{i=1}^{2} u_j(x^*(\theta_1, \alpha), \theta_1, \alpha)\right) + \mathbb{E}_{\theta_1}\left(\sum_{i=1}^{2} u_j(x^*(\theta_1, \beta), \theta_1, \beta)\right) \\
= \;& -\left(\int_0^{\frac{1-\gamma_1}{1-\gamma_2}\alpha} \tilde{u}_2(\theta_1, \alpha) d\theta_1 + \int_{\frac{1-\gamma_1}{1-\gamma_2}\alpha}^{1} \tilde{u}_1(\theta_1, \alpha) d\theta_1\right) + \left(\int_0^{\frac{1-\gamma_1}{1-\gamma_2}\beta} \tilde{u}_2(\theta_1, \beta) d\theta_1 + \int_{\frac{1-\gamma_1}{1-\gamma_2}\beta}^{1} \tilde{u}_1(\theta_1, \beta) d\theta_1\right) \\
= \;& -\int_0^{\frac{1-\gamma_1}{1-\gamma_2}\alpha} (\alpha + \gamma_2\theta_1) d\theta_1 - \int_{\frac{1-\gamma_1}{1-\gamma_2}\alpha}^{1} (\theta_1 + \gamma_1\alpha) d\theta_1 + \int_0^{\frac{1-\gamma_1}{1-\gamma_2}\beta} (\beta + \gamma_2\theta_1) d\theta_1 + \int_{\frac{1-\gamma_1}{1-\gamma_2}\beta}^{1} (\theta_1 + \gamma_1\beta) d\theta_1 \\
= \;& -\frac{1-\gamma_1}{1-\gamma_2}\alpha^2 - \frac{1}{2}\gamma_2\left(\frac{1-\gamma_1}{1-\gamma_2}\alpha\right)^2 - \frac{1}{2}\left(1 - \left(\frac{1-\gamma_1}{1-\gamma_2}\alpha\right)^2\right) - \gamma_1\alpha\left(1 - \frac{1-\gamma_1}{1-\gamma_2}\alpha\right) \\
& + \frac{1-\gamma_1}{1-\gamma_2}\beta^2 + \frac{1}{2}\gamma_2\left(\frac{1-\gamma_1}{1-\gamma_2}\beta\right)^2 + \frac{1}{2}\left(1 - \left(\frac{1-\gamma_1}{1-\gamma_2}\beta\right)^2\right) + \gamma_1\beta\left(1 - \frac{1-\gamma_1}{1-\gamma_2}\beta\right).
\end{aligned}
$$

After making a further rearrangement of the above expression, we obtain

$$
\begin{aligned}
& -\mathbb{E}_{\theta_1}\left(\sum_{i=1}^{2} u_j(x^*(\theta_1, \alpha), \theta_1, \alpha)\right) + \mathbb{E}_{\theta_1}\left(\sum_{i=1}^{2} u_j(x^*(\theta_1, \beta), \theta_1, \beta)\right) \\
= \;& -\gamma_1(\alpha - \beta) - \frac{1}{2}\frac{(1-\gamma_1)^2}{1-\gamma_2}\left(\alpha^2 - \beta^2\right).
\end{aligned}
$$

Plugging this back into (41), we obtain

$$
\begin{aligned}
t_2^G(\hat{\theta}_1, \alpha; u_1^\alpha, u_2^\alpha) - t_2^G(\hat{\theta}_1, \beta; u_1^\beta, u_2^\beta) &= -\frac{1}{2}\left[\alpha - \beta - \gamma_1(\alpha - \beta) - \frac{1}{2}\frac{(1-\gamma_1)^2}{1-\gamma_2}\left(\alpha^2 - \beta^2\right)\right] \\
&= -\frac{1}{4}\frac{1-\gamma_1}{1-\gamma_2}(\alpha - \beta)\left[2(1-\gamma_2) - (1-\gamma_1)(\alpha + \beta)\right] \\
&< 0,
\end{aligned}
$$

where the last strict inequality above follows because $\alpha > \beta$ and $2(1-\gamma_2)-(1-\gamma_1)(\alpha+\beta) > 0$, which is followed by the assumption that $2 > \alpha+\beta$ and $1-\gamma_2 \geq 1-\gamma_1 > 0$. Therefore, we show that the generalized two-stage Groves mechanism is monotone in this case.

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

Fix $\hat{\theta}_1 \in [0,1]$ and let $\alpha, \beta \in [0,1]$ be two distinct types of the buyer such that $\alpha > \beta$ and $x^*(\hat{\theta}_1, \alpha) = x^*(\hat{\theta}_1, \beta) = 1$. Let $\theta_2^* \in (\underline{\theta}_1, \bar{\theta}_2]$ be the unique cutoff point identified in Lemma 8. There are two subcases:

Case 1: $\alpha \leqslant \theta_2^*$

In this subcase, we can apply here the same argument in Case (i) and the buyer's payment is strictly increasing in his type report.

Case 2: $\beta \geqslant \theta_2^*$

The difference between the buyer's transfer under $(\hat{\theta}_1, \alpha)$ and $(\hat{\theta}_1, \beta)$ is computed below:

$$\begin{aligned}
& t_2^G(\hat{\theta}_1, \alpha; u_1^\alpha, u_2^\alpha) - t_2^G(\hat{\theta}_1, \beta; u_1^\beta, u_2^\beta) \\
= \; & u_1^\alpha - \frac{1}{2}\left[\tilde{u}_2(\hat{\theta}_1, \alpha) - \mathbb{E}_{\theta_1}\left(\sum_{i=1}^{2} u_j(x^*(\theta_1, \alpha), \theta_1, \alpha)\right) + \mathbb{E}_{\theta_2}\left(\sum_{j=1}^{2} u_j(x^*(\hat{\theta}_1, \theta_2), \hat{\theta}_1, \theta_2)\right)\right] \\
& -u_1^\beta + \frac{1}{2}\left[\tilde{u}_2(\hat{\theta}_1, \beta) - \mathbb{E}_{\theta_1}\left(\sum_{i=1}^{2} u_j(x^*(\theta_1, \beta), \theta_1, \beta)\right) + \mathbb{E}_{\theta_2}\left(\sum_{j=1}^{2} u_j(x^*(\hat{\theta}_1, \theta_2), \hat{\theta}_1, \theta_2)\right)\right] \\
= \; & -\frac{1}{2}\left[\alpha - \beta - \mathbb{E}_{\theta_1}\left(\sum_{i=1}^{2} u_j(x^*(\theta_1, \alpha), \theta_1, \alpha)\right) + \mathbb{E}_{\theta_1}\left(\sum_{i=1}^{2} u_j(x^*(\theta_1, \beta), \theta_1, \beta)\right)\right], \qquad (42)
\end{aligned}$$

where the second equality follows because $x^*(\hat{\theta}_1, \alpha) = x^*(\hat{\theta}_1, \beta) = 1$ implies $u_1^\alpha = u_1^\beta = 0$ and $\tilde{u}_2(\hat{\theta}_1, \alpha) - \tilde{u}_2(\hat{\theta}_1, \beta) = \alpha + \gamma_2\hat{\theta}_1 - \beta - \gamma_2\hat{\theta}_1 = \alpha - \beta$.

Moreover, we compute the following term in the above expression:

$$\begin{aligned}
& -\mathbb{E}_{\theta_1}\left(\sum_{i=1}^{2} u_j(x^*(\theta_1,\alpha),\theta_1,\alpha)\right) + \mathbb{E}_{\theta_1}\left(\sum_{i=1}^{2} u_j(x^*(\theta_1,\beta),\theta_1,\beta)\right) \\
= & -\int_0^1 \tilde{u}_2(\theta_1,\alpha)d\theta_1 + \int_0^1 \tilde{u}_2(\theta_1,\beta)d\theta_1 \\
= & -\int_0^1 (\alpha+\gamma_2\theta_1)d\theta_1 + \int_0^1 (\beta+\gamma_2\theta_1)d\theta_1 \\
= & -\alpha - \frac{1}{2}\gamma_2 + \beta + \frac{1}{2}\gamma_2 = -\alpha+\beta.
\end{aligned}$$

Plugging this back into (42), we obtain

$$t_2^G(\hat{\theta}_1,\alpha;u_1^\alpha,u_2^\alpha) - t_2^G(\hat{\theta}_1,\beta;u_1^\beta,u_2^\beta) = -\frac{1}{2}\left[\alpha-\beta-\alpha+\beta\right] = 0.$$

Therefore, the generalized two-stage Groves mechanism is monotone in this subcase. This completes the proof of the claim. ■

B.5 Proof of Lemma 9

Proof: Recall our Condition α says that

$$\begin{aligned}
& \int_{\Theta_1}\int_{\Theta_2^*(\theta_1)\backslash\Theta_2^{**}} \left(\tilde{u}_2(\theta_1,\theta_2) - \tilde{u}_1(\theta_1,\theta_2)\right) dF_2(\theta_2)dF_1(\theta_1) \\
& + \int_{\Theta_1}\int_{\Theta_2^{**}} \left(\tilde{u}_2(\theta_1,\theta_2^*) - \tilde{u}_1(\theta_1,\theta_2)\right) dF_2(\theta_2)dF_1(\theta_1) \\
\geqslant & \;\; 0,
\end{aligned}$$

where $\theta_2^* \in (\underline{\theta}_2,\bar{\theta}_2]$ is the cutoff point identified in Lemma 8, $\Theta_2^{**} = [\theta_2^*,\bar{\theta}_2]$, and for each $\theta_1 \in \Theta_1$,

$$\Theta_2^*(\theta_1) = \begin{cases} \{\bar{\theta}_2\} & \text{if } \{\theta_2 \in \Theta_2 : x^*(\theta_1,\theta_2)=1\} = \emptyset \\ \{\theta_2 \in \Theta_2 : x^*(\theta_1,\theta_2)=1\} & \text{otherwise,} \end{cases}$$

Here, the cutoff point θ_2^* is equal to $\min\{(1-\gamma_2)/(1-\gamma_1),1\}$ above which it is always efficient to trade and below which it is efficient not to trade for some $\theta_1 \in \Theta_1$. Moreover, we have $\Theta_2^*(\theta_1) = [\min\{(1-\gamma_2)\theta_1/(1-\gamma_1),1\},1]$ and $\Theta_2^{**} = [\min\{(1-\gamma_2)/(1-\gamma_1),1\},1]$. So, $\Theta_2^*(\theta_1)\backslash\Theta_2^{**} = [\min\{(1-\gamma_2)\theta_1/(1-\gamma_1),1\},\min\{(1-\gamma_2)/(1-\gamma_1),1\}]$.

We divide our argument into the following two cases:

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

From Figure 1, we know that if $\theta_1 > (1-\gamma_1)/(1-\gamma_2)$, then $(1-\gamma_2)\theta_1/(1-\gamma_1) > 1$; hence,

$$\Theta_2^*(\theta_1) = \left[\min\left\{\frac{1-\gamma_2}{1-\gamma_1}\theta_1, 1\right\}, 1\right] = \begin{cases} \{1\} & \text{if } \theta_1 > (1-\gamma_1)/(1-\gamma_2) \\ \left[\frac{1-\gamma_2}{1-\gamma_1}\theta_1, 1\right] & \text{otherwise.} \end{cases}$$

Moreover, in Case (i), we know $(1-\gamma_2)/(1-\gamma_1) > 1$; hence,

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

As a result,

$$\Theta_2^*(\theta_1)\backslash\Theta_2^{**} = \left[\min\left\{\frac{1-\gamma_2}{1-\gamma_1}\theta_1, 1\right\}, \min\left\{\frac{1-\gamma_2}{1-\gamma_1}, 1\right\}\right] = \begin{cases} \emptyset & \text{if } \theta_1 > (1-\gamma_1)/(1-\gamma_2) \\ \left[\frac{1-\gamma_2}{1-\gamma_1}\theta_1, 1\right) & \text{otherwise.} \end{cases}$$

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

$$\int_0^{\frac{1-\gamma_1}{1-\gamma_2}} \int_{\frac{1-\gamma_2}{1-\gamma_1}\theta_1}^1 ((1-\gamma_1)\theta_2 - (1-\gamma_2)\theta_1)\, d\theta_2 d\theta_1 \geqslant 0.$$

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

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

We continue our computation below:

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

which is strictly positive. Therefore, Condition α is satisfied in Case (i).

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

From Figure 2, we know that $(1-\gamma_2)\theta_1/(1-\gamma_1) < 1$ for all $\theta_1 \in [0,1]$; hence,

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

for all $\theta_1 \in [0,1]$. Moreover, in Case (ii), we know $(1-\gamma_2)/(1-\gamma_1) < 1$; hence,

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

As a result, we have that for all $\theta_1 \in \Theta_1$,

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

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

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

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

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

We continue our computation below:

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

Rearranging the terms above, we obtain

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

Therefore, our Condition α is reduced to

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

■

B.6 Proof of Step 1 in the Proof of Theorem 5

Step 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.

Proof:

Consider the seller of type θ_1. Then, the expected utility of the seller of type θ_1 under truth-telling is

$$\int_{\Theta_2\backslash\Theta_2^*(\theta_1)}(\tilde{u}_1(\theta_1,\theta_2)+0)\,dF_2(\theta_2)+\int_{\Theta_2^*(\theta_1)\backslash\Theta_2^{**}}(0+\tilde{u}_2(\theta_1,\theta_2))\,dF_2(\theta_2)+\int_{\Theta_2^{**}}(0-g(\theta_1))\,dF_2(\theta_2).$$

On the other hand, if the seller deviates to $\theta_1^r\neq\theta_1$ and trade occurs, the second-stage report by the buyer of type θ_2 becomes $u_2^r=u_2^r(x^*(\theta_1^r,\theta_2),\theta_1,\theta_2)=\tilde{u}_2(\theta_1,\theta_2)$. Since $\tilde{u}_2(\cdot)$ is strictly increasing in θ_1, then $u_2^r=\tilde{u}_2(\theta_1,\theta_2)\neq\tilde{u}_2(\theta_1^r,\theta_2)$ and the seller must pay a penalty ψ according to the transfer rule t_1^M. Therefore, the expected utility of the seller of type θ_1 becomes

$$\int_{\Theta_2\backslash\Theta_2^*(\theta_1^r)}(\tilde{u}_1(\theta_1,\theta_2)+0)\,dF_2(\theta_2)+\int_{\Theta_2^*(\theta_1^r)}(0-\psi)dF_2(\theta_2).$$

By Lemma 8, we divide our argument into the following two cases:

Case 1: $\theta_2^* = \bar{\theta}_2$, i.e., $\int_{\Theta_1} x^*(\theta_1, \theta_2) dF_1(\theta_1) < 1$ for all $\theta_2 < \bar{\theta}_2$.

Then, the expected utility of the seller of type θ_1 becomes

$$\int_{\Theta_2 \backslash \Theta_2^*(\theta_1)} (\tilde{u}_1(\theta_1, \theta_2) + 0) \, dF_2(\theta_2) + \int_{\Theta_2^*(\theta_1)} (0 + \tilde{u}_2(\theta_1, \theta_2)) \, dF_2(\theta_2),$$

where

$$\Theta_2^*(\theta_1) = \begin{cases} \{\bar{\theta}_2\} & \text{if } \{\theta_2 \in \Theta_2 : x^*(\theta_1, \theta_2) = 1\} = \emptyset \\ \{\theta_2 \in \Theta_2 : x^*(\theta_1, \theta_2) = 1\} & \text{otherwise.} \end{cases}$$

Since $\psi > 0$, the best possible deviation the seller of type θ_1 can achieve is to announce θ_1^r such that $\Theta_2^*(\theta_1^r) = \emptyset$. This implies that the seller keeps the good so that the seller's expected payoff becomes $\int_{\Theta_2} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2)$. However, we claim that this expected utility is at most the same as that under truth-telling. To see this, we compute the difference between the seller's expected utility under truth-telling and that under the best deviation:

$$\begin{aligned}
& \int_{\Theta_2 \backslash \Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) + \int_{\Theta_2^*(\theta_1)} \tilde{u}_2(\theta_1, \theta_2) dF_2(\theta_2) - \int_{\Theta_2} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) \\
= & \int_{\Theta_2 \backslash \Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) + \int_{\Theta_2^*(\theta_1)} \tilde{u}_2(\theta_1, \theta_2) dF_2(\theta_2) \\
& - \left[\int_{\Theta_2 \backslash \Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) + \int_{\Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) \right] \\
= & \int_{\Theta_2^*(\theta_1)} (\tilde{u}_2(\theta_1, \theta_2) - \tilde{u}_1(\theta_1, \theta_2)) \, dF_2(\theta_2) \\
\geq & \; 0,
\end{aligned}$$

where the weak inequality follows because whenever $\theta_2 \in \Theta_2^*(\theta_1)$, it is efficient to trade, implying that $\tilde{u}_2(\theta_1, \theta_2) - \tilde{u}_1(\theta_1, \theta_2) > 0$. So, the seller will never be better off after such a deviation so that he has no incentive to deviate from truth-telling.

Case 2: $\theta_2^* \in (\underline{\theta}_2, \bar{\theta}_2)$ such that for any $\theta_2 \in \Theta_2$,

$$\int_{\Theta_1} x^*(\theta_1, \theta_2) dF_1(\theta_1) \begin{cases} < 1 & \text{if } \theta_2 < \theta_2^* \\ = 1 & \text{if } \theta_2 \geq \theta_2^* \end{cases}$$

To stop the seller of type θ_1 from deviating to θ_1^r, the penalty ψ must be large enough so that the seller always receives at most the same expected utility as that under truth-telling whenever he deviates. That is, what we want is that for any $\theta_1, \theta_1^r \in \Theta_1$,

$$\begin{aligned}
&\int_{\Theta_2 \setminus \Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) + \int_{\Theta_2^*(\theta_1) \setminus \Theta_2^{**}} \tilde{u}_2(\theta_1, \theta_2) dF_2(\theta_2) - \int_{\Theta_2^{**}} g(\theta_1) dF_2(\theta_2) \\
\geq & \int_{\Theta_2 \setminus \Theta_2^*(\theta_1^r)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) - \psi \int_{\Theta_2^*(\theta_1^r)} dF_2(\theta_2),
\end{aligned}$$

where

$$\Theta_2^*(\theta_1) = \begin{cases} \{\bar{\theta}_2\} & \text{if } \{\theta_2 \in \Theta_2 : x^*(\theta_1, \theta_2) = 1\} = \emptyset \\ \{\theta_2 \in \Theta_2 : x^*(\theta_1, \theta_2) = 1\} & \text{otherwise,} \end{cases}$$

and $\Theta_2^{**} = [\theta_2^*, \bar{\theta}_2]$. After rearranging the terms for ψ, we obtain

$$\begin{aligned}
\psi \geq & \frac{1}{\int_{\Theta_2^*(\theta_1^r)} dF_2(\theta_2)} \left(\int_{\Theta_2 \setminus \Theta_2^*(\theta_1^r)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) - \int_{\Theta_2 \setminus \Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) \right) \\
& + \frac{1}{\int_{\Theta_2^*(\theta_1^r)} dF_2(\theta_2)} \left(- \int_{\Theta_2^*(\theta_1) \setminus \Theta_2^{**}} \tilde{u}_2(\theta_1, \theta_2) dF_2(\theta_2) + \int_{\Theta_2^{**}} g(\theta_1) dF_2(\theta_2) \right).
\end{aligned}$$

Then, it remains to find an upper bound of the right-hand side of the above inequality. We obtain this upper bound as follows:

$$\begin{aligned}
\psi \geq & \sup_{\substack{\theta_1 \in \Theta_1 \\ \theta_1^r \in \Theta_1}} \left[\frac{1}{\int_{\Theta_2^*(\theta_1^r)} dF_2(\theta_2)} \left(\int_{\Theta_2 \setminus \Theta_2^*(\theta_1^r)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) - \int_{\Theta_2 \setminus \Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) \right) \right. \\
& \left. + \frac{1}{\int_{\Theta_2^*(\theta_1^r)} dF_2(\theta_2)} \left(- \int_{\Theta_2^*(\theta_1) \setminus \Theta_2^{**}} \tilde{u}_2(\theta_1, \theta_2) dF_2(\theta_2) + \int_{\Theta_2^{**}} g(\theta_1) dF_2(\theta_2) \right) \right].
\end{aligned}$$

In this case, we know $[\theta_2^*, \bar{\theta}_2] \subseteq \Theta_2^*(\theta_1^r)$ for all $\theta_1^r \in \Theta_1$.[24] Therefore, $\Theta_2^*(\theta_1^r)$ is nonempty and carries positive measure under $F_2(\cdot)$ so that the denominator $\int_{\Theta_2^*(\theta_1^r)} dF_2(\theta_2)$ is strictly positive. Moreover, since the type space Θ is bounded and each valuation function $\tilde{u}_i(\cdot, \theta)$ is bounded, the right-hand side of the above inequality is bounded and we denote it by A_1. So, if

$$\psi \geq A_1,$$

the seller will never be better off after such a deviation so that he has no incentive to deviate from truth-telling. This completes the proof of Step 1.

■

[24] In the first general case, the denominator may be zero because $\Theta_2^*(\theta_1)$ may be a singleton for some $\theta_1 \in \Theta_1$.

B.7 Proof of Step 2 in the Proof of Theorem 5

Step 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.

Proof:

Consider the buyer of type $\theta_2 < \theta_2^*$. Then, the buyer's expected utility under truth-telling, denoted by $U_2(\theta_2)$, is

$$U_2(\theta_2) = \int_{\Theta_1^*(\theta_2)} (\tilde{u}_2(\theta_1, \theta_2) - \tilde{u}_2(\theta_1, \theta_2))\, dF_1(\theta_1) = 0,$$

where $\Theta_1^*(\theta_2) = \{\theta_1 \in \Theta_1 : x^*(\theta_1, \theta_2) = 1\}$. On the other hand, if the buyer of type θ_2 deviates to $\theta_2^r \neq \theta_2$ such that $\theta_2^r < \theta_2^*$ and no trade occurs, the second-stage report of the seller of type θ_1 becomes $u_1^r = u_1^r(x^*(\theta_1, \theta_2^r), \theta_1, \theta_2) = \tilde{u}_1(\theta_1, \theta_2)$. Since the seller's utility function $\tilde{u}_1(\cdot)$ is strictly increasing in θ_2, then $u_1^r = \tilde{u}_1(\theta_1, \theta_2) \neq \tilde{u}_1(\theta_1, \theta_2^r)$ and the buyer must pay a penalty ψ according to the transfer rule t_2^M. Therefore, the expected utility of the buyer of type θ_2 when announcing θ_2^r becomes

$$\int_{\Theta_1^*(\theta_2^r)} (\tilde{u}_2(\theta_1, \theta_2) - \tilde{u}_2(\theta_1, \theta_2^r))\, dF_1(\theta_1) + \int_{\Theta_1 \backslash \Theta_1^*(\theta_2^r)} (0 - \psi) dF_1(\theta_1).$$

To stop the buyer from deviating, the penalty ψ must be large enough so that the buyer always receives at most the same expected utility as that under truth-telling whenever he deviates. That is, for any $\theta_2 < \theta_2^*$ and $\theta_2^r < \theta_2^*$,

$$0 \geq \int_{\Theta_1^*(\theta_2^r)} (\tilde{u}_2(\theta_1, \theta_2) - \tilde{u}_2(\theta_1, \theta_2^r))\, dF_1(\theta_1) - \psi \int_{\Theta_1 \backslash \Theta_1^*(\theta_2^r)} dF_1(\theta_1).$$

After rearranging the terms for ψ in the above inequality, we obtain

$$\psi \geq \frac{\int_{\Theta_1^*(\theta_2^r)} (\tilde{u}_2(\theta_1, \theta_2) - \tilde{u}_2(\theta_1, \theta_2^r))\, dF_1(\theta_1)}{\int_{\Theta_1 \backslash \Theta_1^*(\theta_2^r)} dF_1(\theta_1)},$$

where $\Theta_1 \backslash \Theta_1^*(\theta_2^r) = \{\theta_1 \in \Theta_1 : x^*(\theta_1, \theta_2^r) = 0\}$. We know that for any $\theta_2^r < \theta_2^*$, there must exist some $\theta_1 \in \Theta_1$ such that $x^*(\theta_1, \theta_2^r) = 0$. Therefore, $\Theta_1 \backslash \Theta_1^*(\theta_2^r)$ is nonempty and carries positive measure under $F_1(\cdot)$ so that the denominator is strictly positive. Then, it remains to find an upper bound of the right-hand side of the above inequality. Such an upper bound can be found as follows:

$$\psi \geq \sup_{\substack{\theta_2 \in [\underline{\theta}_2, \theta_2^*) \\ \theta_2^r \in [\underline{\theta}_2, \theta_2^*)}} \frac{\int_{\Theta_1^*(\theta_2^r)} (\tilde{u}_2(\theta_1, \theta_2) - \tilde{u}_2(\theta_1, \theta_2^r))\, dF_1(\theta_1)}{\int_{\Theta_1 \backslash \Theta_1^*(\theta_2^r)} dF_1(\theta_1)}.$$

since $[\underline{\theta}_2, \theta_2^*)$ is bounded and each $\tilde{u}_2(\cdot, \theta)$ is bounded, the numerator is also bounded. Therefore, the right-hand side of the above inequality is bounded and we denote its upper bound by A_2. So, if

$$\psi \geq A_2,$$

the buyer of type $\theta_2 < \theta_2^*$ will never be better off after such a deviation so that he has no incentive to deviate to $\theta_2^r < \theta_2^*$ from truth-telling.

Moreover, if the buyer deviates to $\theta_2^r \geq \theta_2^*$, it is always efficient to trade and the expected utility of the buyer of type θ_2 when announcing θ_2^r, denoted by $U_2(\theta_2, \theta_2^r)$, becomes

$$U_2(\theta_2, \theta_2^r) = \int_{\Theta_1} (\tilde{u}_2(\theta_1, \theta_2) + g(\theta_1))\, dF_1(\theta_1).$$

Then, the difference between the expected utility of the buyer of type θ_2 under truth-telling and that under deviation to θ_2^r is

$$\begin{aligned} U_2(\theta_2) - U_2(\theta_2, \theta_2^r) &= 0 - \int_{\Theta_1} (\tilde{u}_2(\theta_1, \theta_2) + g(\theta_1))\, dF_1(\theta_1) \\ &= -\int_{\Theta_1} \tilde{u}_2(\theta_1, \theta_2) dF_1(\theta_1) - \int_{\Theta_1} g(\theta_1) dF_1(\theta_1). \end{aligned}$$

By Lemma 8, we divide our argument into the following two cases:

Case 1: $\theta_2^* = \bar{\theta}_2$, i.e., $\int_{\Theta_1} x^*(\theta_1, \theta_2) dF_1(\theta_1) < 1$ for all $\theta_2 < \bar{\theta}_2$.

Recall that in this case, $g(\theta_1) = -\tilde{u}_2(\theta_1, \bar{\theta}_2)$. Then, we evaluate the utility difference.

$$\begin{aligned} U_2(\theta_2) - U_2(\theta_2, \theta_2^r) &= -\int_{\Theta_1} \tilde{u}_2(\theta_1, \theta_2) dF_1(\theta_1 + \int_{\Theta_1} \tilde{u}_2(\theta_1, \bar{\theta}_2) dF_1(\theta_1) \\ &= \int_{\Theta_1} \big(\tilde{u}_2(\theta_1, \bar{\theta}_2) - \tilde{u}_2(\theta_1, \theta_2)\big)\, dF_1(\theta_1) \\ &\geqslant 0, \end{aligned}$$

where the last inequality follows because $\theta_2 \leqslant \bar{\theta}_2$ and $\tilde{u}_2(\cdot)$ is strictly increasing in θ_2. 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 $\theta_2^r \geq \theta_2^*$ in this case.

Case 2: $\theta_2^* \in (\underline{\theta}_2, \bar{\theta}_2)$ such that for any $\theta_2 \in \Theta_2$,

$$\int_{\Theta_1} x^*(\theta_1, \theta_2) dF_1(\theta_1) \begin{cases} < 1 & \text{if } \theta_2 < \theta_2^* \\ = 1 & \text{if } \theta_2 \geq \theta_2^* \end{cases}$$

Recalling the definition of $g(\theta_1)$, we obtain

$$
\begin{aligned}
& (1-F_2(\theta_2^*))\int_{\Theta_1} g(\theta_1)dF_1(\theta_1) \\
= & \int_{\Theta_1}\int_{\Theta_2^*(\theta_1)\backslash\Theta_2^{**}} \tilde{u}_2(\theta_1,\theta_2)dF_2(\theta_2)dF_1(\theta_1) - \int_{\Theta_1}\int_{\Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1,\theta_2)dF_2(\theta_2)dF_1(\theta_1) \\
& -\int_{\Theta_1}\int_{\Theta_2^*(\theta_1)\backslash\Theta_2^{**}} \left(\tilde{u}_2(\theta_1,\theta_2)-\tilde{u}_1(\theta_1,\theta_2)\right)dF_2(\theta_2)dF_1(\theta_1) \\
& -\int_{\Theta_1}\int_{\Theta_2^{**}} \left(\tilde{u}_2(\theta_1,\theta_2^*)-\tilde{u}_1(\theta_1,\theta_2)\right)dF_2(\theta_2)dF_1(\theta_1) \\
= & \int_{\Theta_1}\int_{\Theta_2^*(\theta_1)\backslash\Theta_2^{**}} \tilde{u}_2(\theta_1,\theta_2)dF_2(\theta_2)dF_1(\theta_1) - \int_{\Theta_1}\int_{\Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1,\theta_2)dF_2(\theta_2)dF_1(\theta_1) \\
& -\int_{\Theta_1}\int_{\Theta_2^*(\theta_1)\backslash\Theta_2^{**}} \tilde{u}_2(\theta_1,\theta_2)dF_2(\theta_2)dF_1(\theta_1) + \int_{\Theta_1}\int_{\Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1,\theta_2)dF_2(\theta_2)dF_1(\theta_1) \\
& -\int_{\Theta_1}\int_{\Theta_2^{**}} \tilde{u}_2(\theta_1,\theta_2^*)dF_2(\theta_2)dF_1(\theta_1).
\end{aligned}
$$

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

$$
\begin{aligned}
(1-F_2(\theta_2^*))\int_{\Theta_1} g(\theta_1)dF_1(\theta_1) &= -\int_{\Theta_1}\int_{\Theta_2^{**}} \tilde{u}_2(\theta_1,\theta_2^*)dF_2(\theta_2)dF_1(\theta_1) \\
&= -(1-F_2(\theta_2^*))\int_{\Theta_1} \tilde{u}_2(\theta_1,\theta_2^*)dF_1(\theta_1)
\end{aligned}
$$

Therefore, we obtain

$$
\int_{\Theta_1} g(\theta_1)dF_1(\theta_1) = -\int_{\Theta_1} \tilde{u}_2(\theta_1,\theta_2^*)dF_1(\theta_1).
$$

Plugging this back into the utility difference, we obtain

$$
\begin{aligned}
U_2(\theta_2)-U_2(\theta_2,\theta_2^r) &= -\int_{\Theta_1} \tilde{u}_2(\theta_1,\theta_2)dF_1(\theta_1) + \int_{\Theta_1} \tilde{u}_2(\theta_1,\theta_2^*)dF_1(\theta_1) \\
&= \int_{\Theta_1} \left(\tilde{u}_2(\theta_1,\theta_2^*)-\tilde{u}_2(\theta_1,\theta_2)\right)dF_1(\theta_1) \\
&> 0,
\end{aligned}
$$

where the last strict inequality follows because $\theta_2 < \theta_2^*$ and $\tilde{u}_2(\cdot)$ is strictly increasing in θ_2. 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 $\theta_2^r \geq \theta_2^*$.

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

$$\int_{\Theta_1} (\tilde{u}_2(\theta_1, \theta_2) + g(\theta_1))\, dF_1(\theta_1).$$

On the other hand, if the buyer deviates to $\theta_2^r \neq \theta_2$ such that $\theta_2^r \geq \theta_2^*$, then it is still always efficient to trade regardless of the seller's type. Thus, the expected utility of the buyer of type θ_2 under the deviation to θ_2^r is

$$\int_{\Theta_1} (\tilde{u}_2(\theta_1, \theta_2) + g(\theta_1))\, dF_1(\theta_1),$$

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 the buyer of type θ_2 deviates to $\theta_2^r < \theta_2^*$ and trade does not occur, the second-stage report of the seller of type θ_1 becomes $u_1^r = u_1^r(x^*(\theta_1, \theta_2^r), \theta_1, \theta_2) = \tilde{u}_1(\theta_1, \theta_2)$. Since $\tilde{u}_1(\cdot)$ is strictly increasing in θ_2, we have that $u_1^r = \tilde{u}_1(\theta_1, \theta_2) \neq \tilde{u}_1(\theta_1, \theta_2^r)$ so that the buyer must pay a penalty ψ according to the transfer rule t_2^M. Therefore, the expected utility of the buyer of type θ_2 when announcing θ_2^r becomes

$$\int_{\Theta_1^*(\theta_2^r)} (\tilde{u}_2(\theta_1, \theta_2) - \tilde{u}_2(\theta_1, \theta_2^r))\, dF_1(\theta_1) + \int_{\Theta_1 \setminus \Theta_1^*(\theta_2^r)} (0 - \psi) dF_1(\theta_1).$$

To stop the buyer from deviating, the penality ψ must be large enough so that the buyer always receives at most the same expected utility as that under truth-telling whenever he deviates. That is, what we want to have is that for any $\theta_2 \geq \theta_2^*$ and $\theta_2^r < \theta_2^*$,

$$\int_{\Theta_1} (\tilde{u}_2(\theta_1, \theta_2) + g(\theta_1))\, dF_1(\theta_1) \quad \geq \quad \int_{\Theta_1^*(\theta_2^r)} (\tilde{u}_2(\theta_1, \theta_2) - \tilde{u}_2(\theta_1, \theta_2^r))\, dF_1(\theta_1) - \psi \int_{\Theta_1 \setminus \Theta_1^*(\theta_2^r)} dF_1(\theta_1).$$

As we argued previously, we know that $\int_{\Theta_1} g(\theta_1) dF_1(\theta_1) = -\int_{\Theta_1} \tilde{u}_2(\theta_1, \theta_2^*) dF_1(\theta_1)$. Hence, the above inequality can be rewritten as

$$\int_{\Theta_1} (\tilde{u}_2(\theta_1, \theta_2) - \tilde{u}_2(\theta_1, \theta_2^*))\, dF_1(\theta_1) \quad \geq \quad \int_{\Theta_1^*(\theta_2^r)} (\tilde{u}_2(\theta_1, \theta_2) - \tilde{u}_2(\theta_1, \theta_2^r))\, dF_1(\theta_1) - \psi \int_{\Theta_1 \setminus \Theta_1^*(\theta_2^r)} dF_1(\theta_1).$$

After rearranging the terms above for ψ in the above inequality, we obtain

$$\psi \geq \frac{1}{\int_{\Theta_1 \setminus \Theta_1^*(\theta_2^r)} dF_1(\theta_1)} \left[- \int_{\Theta_1 \setminus \Theta_1^*(\theta_2^r)} \tilde{u}_2(\theta_1, \theta_2) dF_1(\theta_1) - \int_{\Theta_1^*(\theta_2^r)} \tilde{u}_2(\theta_1, \theta_2^r) dF_1(\theta_1) + \int_{\Theta_1} \tilde{u}_2(\theta_1, \theta_2^*) dF_1(\theta_1) \right].$$

Then, it remains to find an upper bound of the right-hand side of the above inequality. Therefore, we want to satisfy the following inequality:

$$\psi \geq \sup_{\substack{\theta_2\in[\theta_2^*,\bar{\theta}_2]\\ \theta_2^r\in[\underline{\theta}_2,\theta_2^*)}} \left[\frac{1}{\int_{\Theta_1\backslash\Theta_1^*(\theta_2^r)} dF_1(\theta_1)}\left(-\int_{\Theta_1\backslash\Theta_1^*(\theta_2^r)} \tilde{u}_2(\theta_1,\theta_2)dF_1(\theta_1) - \int_{\Theta_1^*(\theta_2^r)} \tilde{u}_2(\theta_1,\theta_2^r)dF_1(\theta_1)\right)\right.$$
$$\left. + \frac{1}{\int_{\Theta_1\backslash\Theta_1^*(\theta_2^r)} dF_1(\theta_1)}\int_{\Theta_1} \tilde{u}_2(\theta_1,\theta_2^*)dF_1(\theta_1)\right].$$

Previously, we have argued that if $\theta_2^r < \theta_2^*$, then $\Theta_1\backslash\Theta_1^*(\theta_2^r)$ carries positive measure under $F_1(\cdot)$, that is, $\int_{\Theta_1\backslash\Theta_1^*(\theta_2^r)} dF_1(\theta_1) > 0$. Moreover, since $[\theta_2^*, \bar{\theta}_2]$ and $[\underline{\theta}_2, \theta_2^*)$ are bounded and $\tilde{u}_2(\cdot)$ is bounded, the right-hand side of the above inequality is bounded. We denote this upper bound by A_3. So, if

$$\psi \geq A_3,$$

the buyer will never be better off after such a deviation so that he has no incentive to deviate to $\theta_2^r < \theta_2^*$. This completes the proof of Step 2. ■

B.8 Proof of Step 3 in the Proof of Theorem 5

Step 3: The two-stage mechanism (Θ, Π, x^*, t) also satisfies IIR.

Proof: By Steps 1 and 2, we set $\psi = \max\{A_1, A_2, A_3\}$. We first show that IIR is satisfied for the seller. 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, denoted by $U_1(\theta_1)$, is

$$U_1(\theta_1) = \int_{\Theta_2\backslash\Theta_2^*(\theta_1)} (\tilde{u}_1(\theta_1,\theta_2)+0)\, dF_2(\theta_2) + \int_{\Theta_2^*(\theta_1)\backslash\Theta_2^{**}} (0+\tilde{u}_2(\theta_1,\theta_2))\, dF_2(\theta_2)$$
$$+ \int_{\Theta_2^{**}} (0 - g(\theta_1))\, dF_2(\theta_2).$$

By Lemma 8, we continue our discussion by considering the following two cases:

Case 1: $\theta_2^* = \bar{\theta}_2$. That is, $\int_{\Theta_1} x^*(\theta_1,\theta_2)dF_1(\theta_1) < 1$ for all $\theta_2 < \bar{\theta}_2$.

Then, the expected utility of the seller of type θ_1 after participating in the mechanism becomes

$$U_1(\theta_1) = \int_{\Theta_2\backslash\Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1,\theta_2)dF_2(\theta_2) + \int_{\Theta_2^*(\theta_1)} \tilde{u}_2(\theta_1,\theta_2)dF_2(\theta_2).$$

Then, we compute the difference between the expected utility of the seller of type θ_1 after participating in the mechanism and θ_1's outside option utility:

$$\begin{aligned}
& U_1(\theta_1) - U_1^O(\theta_1) \\
= & \int_{\Theta_2 \setminus \Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) + \int_{\Theta_2^*(\theta_1)} \tilde{u}_2(\theta_1, \theta_2) dF_2(\theta_2) - \int_{\Theta_2} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) \\
= & \int_{\Theta_2 \setminus \Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) + \int_{\Theta_2^*(\theta_1)} \tilde{u}_2(\theta_1, \theta_2) dF_2(\theta_2) \\
& - \left[\int_{\Theta_2 \setminus \Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) + \int_{\Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) \right] \\
= & \int_{\Theta_2^*(\theta_1)} \left(\tilde{u}_2(\theta_1, \theta_2) - \tilde{u}_1(\theta_1, \theta_2) \right) dF_2(\theta_2) \\
\geq & \ 0,
\end{aligned}$$

where the weak inequality follows because whenever $\theta_2 \in \Theta_2^*(\theta_1)$, it is efficient to trade, implying $\tilde{u}_2(\theta_1, \theta_2) - \tilde{u}_1(\theta_1, \theta_2) > 0$.

Case 2: $\theta_2^* \in (\underline{\theta}_2, \bar{\theta}_2)$ such that

$$\int_{\Theta_1} x^*(\theta_1, \theta_2) dF_1(\theta_1) \begin{cases} < 1 & \text{if } \theta_2 < \theta_2^* \\ = 1 & \text{if } \theta_2 \geq \theta_2^*. \end{cases}$$

We compute the difference between the expected utility of the seller of type θ_1 after participating in the mechanism and θ_1's outside option utility:

$$\begin{aligned}
& U_1(\theta_1) - U_1^O(\theta_1) \\
= & \int_{\Theta_2 \setminus \Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) + \int_{\Theta_2^*(\theta_1) \setminus \Theta_2^{**}} \tilde{u}_2(\theta_1, \theta_2) dF_2(\theta_2) - \int_{\Theta_2^{**}} g(\theta_1) dF_2(\theta_2) \\
& - \int_{\Theta_2} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) \\
= & - \int_{\Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) + \int_{\Theta_2^*(\theta_1) \setminus \Theta_2^{**}} \tilde{u}_2(\theta_1, \theta_2) dF_2(\theta_2) - \int_{\Theta_2^{**}} g(\theta_1) dF_2(\theta_2) \\
= & - \int_{\Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) + \int_{\Theta_2^*(\theta_1) \setminus \Theta_2^{**}} \tilde{u}_2(\theta_1, \theta_2) dF_2(\theta_2) - g(\theta_1) \left(1 - F_2(\theta_2^*)\right).
\end{aligned}$$

Plugging the formula of $g(\theta_1)\left(1 - F_2(\theta_2^*)\right)$ in the above expression, we obtain

$$
\begin{aligned}
U_1(\theta_1) - U_1^O(\theta_1) &= -\int_{\Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) + \int_{\Theta_2^*(\theta_1)\setminus\Theta_2^{**}} \tilde{u}_2(\theta_1, \theta_2) dF_2(\theta_2) \\
&\quad - \int_{\Theta_2^*(\theta_1)\setminus\Theta_2^{**}} \tilde{u}_2(\theta_1, \theta_2) dF_2(\theta_2) + \int_{\Theta_2^*(\theta_1)} \tilde{u}_1(\theta_1, \theta_2) dF_2(\theta_2) \\
&\quad + \int_{\Theta_1} \int_{\Theta_2^*(\theta_1)\setminus\Theta_2^{**}} \left(\tilde{u}_2(\theta_1, \theta_2) - \tilde{u}_1(\theta_1, \theta_2)\right) dF_2(\theta_2) dF_1(\theta_1) \\
&\quad + \int_{\Theta_1} \int_{\Theta_2^{**}} \left(\tilde{u}_2(\theta_1, \theta_2^*) - \tilde{u}_1(\theta_1, \theta_2)\right) dF_2(\theta_2) dF_1(\theta_1).
\end{aligned}
$$

Further rearranging the terms above, we obtain

$$
\begin{aligned}
U_1(\theta_1) - U_1^O(\theta_1) &= \int_{\Theta_1} \int_{\Theta_2^*(\theta_1)\setminus\Theta_2^{**}} \left(\tilde{u}_2(\theta_1, \theta_2) - \tilde{u}_1(\theta_1, \theta_2)\right) dF_2(\theta_2) dF_1(\theta_1) \\
&\quad + \int_{\Theta_1} \int_{\Theta_2^{**}} \left(\tilde{u}_2(\theta_1, \theta_2^*) - \tilde{u}_1(\theta_1, \theta_2)\right) dF_2(\theta_2) dF_1(\theta_1).
\end{aligned}
$$

Then, by our Condition α, we conclude

$$U_1(\theta_1) - U_1^O(\theta_1) \geq 0.$$

Therefore, in both cases, 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. If $\theta_2 < \theta_2^*$ and both agents report truthfully in both stages, the expected utility of the buyer of type θ_2 after participating in the mechanism is

$$\int_{\Theta_1^*(\theta_2)} \left(\tilde{u}_2(\theta_1, \theta_2) - \tilde{u}_2(\theta_1, \theta_2)\right) dF_1(\theta_1) = 0 = U_2^O(\theta_2).$$

Hence, if $\theta_2 < \theta_2^*$, by participating in the mechanism, the buyer receives exactly the same expected utility as his outside option utility.

If $\theta_2 \geq \theta_2^*$, the expected utility of the buyer of type θ_2 after participating in the mechanism is

$$\int_{\Theta_1} \left(\tilde{u}_2(\theta_1, \theta_2) + g(\theta_1)\right) dF_1(\theta_1) = \int_{\Theta_1} \tilde{u}_2(\theta_1, \theta_2) dF_1(\theta_1) + \int_{\Theta_1} g(\theta_1) dF_1(\theta_1).$$

By Lemma 8, we divide our argument into the following two cases:

Case 1: $\theta_2^* = \bar{\theta}_2$, i.e., $\int_{\Theta_1} x^*(\theta_1, \theta_2) dF_1(\theta_1) < 1$ for all $\theta_2 < \bar{\theta}_2$.

Recall that in this case, $g(\theta_1) = -\tilde{u}_2(\theta_1, \bar{\theta}_2)$. Then, the expected utility of the buyer of type $\theta_2 = \bar{\theta}_2$ after participation is

$$\int_{\Theta_1} \tilde{u}_2(\theta_1, \bar{\theta}_2) dF_1(\theta_1) + \int_{\Theta_1} g(\theta_1) dF_1(\theta_1) = \int_{\Theta_1} \tilde{u}_2(\theta_1, \bar{\theta}_2) dF_1(\theta_1) - \int_{\Theta_1} \tilde{u}_2(\theta_1, \bar{\theta}_2) dF_1(\theta_1) = 0.$$

Therefore, if $\theta_2 \geq \theta_2^*$, by participating in the mechanism, the buyer of type θ_2 receives exactly the same expected utility as his outside option utility.

Case 2: $\theta_2^* \in (\underline{\theta}_2, \bar{\theta}_2)$ such that for any $\theta_2 \in \Theta_2$,

$$\int_{\Theta_1} x^*(\theta_1, \theta_2) dF_1(\theta_1) \begin{cases} < 1 & \text{if } \theta_2 < \theta_2^* \\ = 1 & \text{if } \theta_2 \geq \theta_2^*. \end{cases}$$

As we argued previously, we know that

$$\int_{\Theta_1} g(\theta_1) dF_1(\theta_1) = -\int_{\Theta_1} \tilde{u}_2(\theta_1, \theta_2^*) dF_1(\theta_1).$$

Hence, if $\theta_2 \geq \theta_2^*$, the expected utility of the buyer of type θ_2 after participating in the mechanism is

$$\begin{aligned} \int_{\Theta_1} \tilde{u}_2(\theta_1, \theta_2) dF_1(\theta_1) + \int_{\Theta_1} g(\theta_1) dF_1(\theta_1) &= \int_{\Theta_1} \tilde{u}_2(\theta_1, \theta_2) dF_1(\theta_1) - \int_{\Theta_1} \tilde{u}_2(\theta_1, \theta_2^*) dF_1(\theta_1) \\ &= \int_{\Theta_1} \left(\tilde{u}_2(\theta_1, \theta_2) - \tilde{u}_2(\theta_1, \theta_2^*)\right) dF_1(\theta_1) \\ &\geq 0 = U_2^O(\theta_2), \end{aligned}$$

where the weak inequality follows because $\theta_2 \geq \theta_2^*$ and $\tilde{u}_2(\cdot)$ is strictly increasing in θ_2. Therefore, if $\theta_2 \geq \theta_2^*$, by participating in the mechanism, the buyer of type θ_2 receives at least the same expected utility as his outside option utility. We thus conclude that IIR is satisfied for the buyer. This completes the proof. ■

B.9 Proof of Claim 8

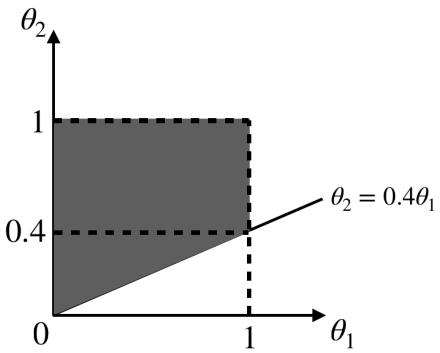
Proof: We show that the two-stage mechanism we propose in Subsection 2.4.2 violates the seller's IIR constraint.

In this case, $\tilde{u}_2(\theta_1, \theta_2) > \tilde{u}_1(\theta_1, \theta_2)$ if and only if $\theta_2 > (1-\gamma_2)\theta_1/(1-\gamma_1) = 0.4\theta_1$. Hence, the efficient decision rule dictates that, for each $(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2$,

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

Figure 8 below illustrates the decision at different type profiles in this case. In particular, the shaded region represents $\Theta^* = \{(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2 : x^*(\theta_1, \theta_2) = 1\}$, which describes the set of possible type profiles for which it is efficient to trade. Note that if $\theta_2 \geq 0.4$, it is always efficient to trade regardless of the seller's type.

Figure 8



Observe that in this case, $\theta_2^* = 0.4$ and $\Theta_2^{**} = [0.4, 1]$. Moreover, the sum of the last two terms in expression (10) in the definition of $G(\theta_1^r)$ (See Subsection 2.4.2) is exactly the negative of the left-hand side of inequality (11) in our Condition α, which is equal to 1/60. Then, expression (10) can be rewritten as the following: for each $\theta_1^r \in [0, 1]$,

$$\begin{aligned} G(\theta_1^r) &= \int_{0.4\theta_1^r}^{0.4} (\theta_2 + 0.8\theta_1^r) d\theta_2 - \int_{0.4\theta_1^r}^{1} (\theta_1^r + 0.5\theta_2) d\theta_2 + \frac{1}{60} \\ &= 0.08\left(1 - (\theta_1^r)^2\right) + 0.32\theta_1^r(1-\theta_1^r) - \theta_1^r(1-0.4\theta_1^r) - 0.25\left(1 - (0.4\theta_1^r)^2\right) + \frac{1}{60}. \end{aligned}$$

Rearranging the terms, we obtain: for each $\theta_1^r \in [0, 1]$,

$$G(\theta_1^r) = -0.68\theta_1^r - 0.17 + \frac{1}{60} + 0.04(\theta_1^r)^2.$$

Then, we have

$$\begin{aligned} g(\theta_1^r) &= \left(-0.68\theta_1^r - 0.17 + \frac{1}{60} + 0.04(\theta_1^r)^2\right)/(1-0.4) \\ &= -\frac{17}{15}\theta_1^r - \frac{23}{90} + \frac{1}{15}(\theta_1^r)^2. \end{aligned}$$

Consider the IIR constraint for the seller of type θ_1. If both agents report truthfully in both stages, the seller's expected utility after participation in the mechanism, denoted by $U_1(\theta_1)$, is

$$U_1(\theta_1) = \int_0^{0.4\theta_1} \tilde{u}_1(\theta_1, \theta_2) d\theta_2 + \int_{0.4\theta_1}^{0.4} \tilde{u}_2(\theta_1, \theta_2) d\theta_2 - \int_{0.4}^{1} g(\theta_1) d\theta_2.$$

Then, the difference between the seller's expected utility after participation in the mechanism and his outside option utility is computed as follows: for any $\theta_1 \in [0, 1]$,

$$\begin{aligned} U_1(\theta_1) - U_1^O(\theta_1) &= \int_0^{0.4\theta_1} \tilde{u}_1(\theta_1, \theta_2) d\theta_2 + \int_{0.4\theta_1}^{0.4} \tilde{u}_2(\theta_1, \theta_2) d\theta_2 - \int_{0.4}^{1} g(\theta_1) d\theta_2 - \int_0^1 \tilde{u}_1(\theta_1, \theta_2) d\theta_2 \\ &= \int_{0.4\theta_1}^{0.4} \tilde{u}_2(\theta_1, \theta_2) d\theta_2 - \int_{0.4}^{1} g(\theta_1) d\theta_2 - \int_{0.4\theta_1}^{1} \tilde{u}_1(\theta_1, \theta_2) d\theta_2. \end{aligned}$$

Plugging the specific valuation functions and $g(\cdot)$ function into the above equation, we obtain

$$\begin{aligned} U_1(\theta_1) - U_1^O(\theta_1) &= \int_{0.4\theta_1}^{0.4} (\theta_2 + 0.8\theta_1) d\theta_2 - \int_{0.4}^{1} \left(-\frac{17}{15}\theta_1 - \frac{23}{90} + \frac{1}{15}(\theta_1)^2 \right) d\theta_2 - \int_{0.4\theta_1}^{1} (\theta_1 + 0.5\theta_2) d\theta_2 \\ &= 0.08\left(1 - (\theta_1)^2\right) + 0.32\theta_1(1 - \theta_1) - 0.6\left(-\frac{17}{15}\theta_1 - \frac{23}{90} + \frac{1}{15}(\theta_1)^2 \right) \\ &\quad -\theta_1(1 - 0.4\theta_1) - 0.25\left(1 - (0.4\theta_1)^2\right). \end{aligned}$$

Rearranging the terms above further, we obtain

$$U_1(\theta_1) - U_1^O(\theta_1) = -\frac{1}{60} < 0,$$

implying that the seller's IIR constraint is violated. ■

B.10 Proof of Lemma 12

Proof: We will first show that if inequality (13) is satisfied, our Condition α is satisfied. In our bilateral trade model, inequality (13) becomes the following condition: for all $\hat{\theta}_1 \neq \theta_1$, there exists $M_1 > 0$ such that

$$\mathbb{E}_{\theta_2}\left[\mathbb{1}_{\{\theta_2 | x^*(\hat{\theta}_1, \theta_2) = 0\}}(\theta_2)\left(\tilde{u}_1(\bar{\theta}_1, \theta_2) - \tilde{u}_1(\hat{\theta}_1, \theta_2)\right)\right] \leqslant M_1 \mathbb{E}_{\theta_2}\left[\mathbb{1}_{\{\theta_2 | x^*(\hat{\theta}_1, \theta_2) = 1, \tilde{u}_2(\theta_1, \theta_2) \neq \tilde{u}_2(\hat{\theta}_1, \theta_2)\}}(\theta_2)\right], \tag{43}$$

and for all $\hat{\theta}_2 \neq \theta_2$, there exists $\tilde{M}_1 > 0$ such that

$$\mathbb{E}_{\theta_1}\left[\mathbb{1}_{\{\theta_1|x^*(\theta_1,\hat{\theta}_2)=1\}}(\theta_1)\left(\tilde{u}_2(\theta_1,\bar{\theta}_2)-\tilde{u}_2(\theta_1,\hat{\theta}_2)\right)\right] \leqslant \tilde{M}_1\mathbb{E}_{\theta_1}\left[\mathbb{1}_{\{\theta_1|x^*(\theta_1,\hat{\theta}_2)=0,\tilde{u}_1(\theta_1,\theta_2)\neq\tilde{u}_1(\theta_1,\hat{\theta}_2)\}}(\theta_1)\right]. \tag{44}$$

Since $\tilde{u}_1(\cdot)$ is assumed to be strictly increasing in θ_2 in our paper, we have that $\tilde{u}_1(\theta_1,\theta_2) \neq \tilde{u}_1(\theta_1,\hat{\theta}_2)$ for all $\theta_1 \in \Theta_1$ and all $\hat{\theta}_2 \neq \theta_2$. Then, inequality (44) can slightly be simplified as follows: for all $\hat{\theta}_2 \neq \theta_2$, there exists $\tilde{M}_1 > 0$ such that

$$\mathbb{E}_{\theta_1}\left[\mathbb{1}_{\{\theta_1|x^*(\theta_1,\hat{\theta}_2)=1\}}(\theta_1)\left(\tilde{u}_2(\theta_1,\bar{\theta}_2)-\tilde{u}_2(\theta_1,\hat{\theta}_2)\right)\right] \leqslant \tilde{M}_1\mathbb{E}_{\theta_1}\left[\mathbb{1}_{\{\theta_1|x^*(\theta_1,\hat{\theta}_2)=0\}}(\theta_1)\right].$$

Suppose on the contrary that our Condition α is violated. Then, there exists $\hat{\theta}_2 < \bar{\theta}_2$ such that $\int_{\Theta_1} x^*(\theta_1,\hat{\theta}_2)dF_1(\theta_1) = 1$, or equivalently, $x^*(\theta_1,\hat{\theta}_2) = 1$ for all $\theta_1 \in \Theta_1$. As a result, the above inequality becomes

$$\mathbb{E}_{\theta_1}\left[\tilde{u}_2(\theta_1,\bar{\theta}_2)-\tilde{u}_2(\theta_1,\hat{\theta}_2)\right] \leqslant \tilde{M}_1\mathbb{E}_{\theta_1}[0] = 0.$$

Since $\bar{\theta}_2 > \hat{\theta}_2$, by the strict increasingness of $\tilde{u}_2(\cdot)$ in θ_2, we have $\tilde{u}_2(\theta_1,\bar{\theta}_2)-\tilde{u}_2(\theta_1,\hat{\theta}_2) > 0$ for all $\theta_1 \in \Theta_1$. Thus, the left-hand side of the above inequality is strictly positive, leading to a contradiction. So, if inequality (13) is satisfied, then our Condition α is also satisfied.

Second, we will show that inequality (14) is automatically satisfied in our bilateral trade model. First we reproduce inequality (14): there exists $M_2 \geq 0$ such that for all $i \in \{1,2\}$, all $\theta_i, \hat{\theta}_i \in \Theta_i$ with $\hat{\theta}_i \neq \theta_i$,

$$\sum_{j\neq i}\mathbb{E}_{\theta_{-i}}\left[\mathbb{1}_{\{\theta_{-i}|j=m(\hat{\theta}_i,\theta_{-i}),\tilde{u}_j(\theta_i,\theta_{-i})=\tilde{u}_j(\hat{\theta}_i,\theta_{-i})\}}(\theta_{-i})\right] \leqslant M_2\sum_{j\neq i}\mathbb{E}_{\theta_{-i}}\left[\mathbb{1}_{\{\theta_{-i}|j=m(\hat{\theta}_i,\theta_{-i}),\tilde{u}_j(\theta_i,\theta_{-i})\neq\tilde{u}_j(\hat{\theta}_i,\theta_{-i})\}}(\theta_{-i})\right],$$

where $m(\theta) = \max\{\arg\max_j \tilde{u}_j(\theta)\}$. We assume throughout that each agent's valuation is strictly increasing in the other agent's type. Thus, for all $j \neq i$ and all $\hat{\theta}_i \neq \theta_i$, it is impossible to have $\tilde{u}_j(\theta_i,\theta_{-i}) = \tilde{u}_j(\hat{\theta}_i,\theta_{-i})$ so that the left-hand side of the above inequality is zero. On the other hand, the right-hand side of the above inequality is always nonnegative. Therefore, the above inequality is automatically satisfied in our bilateral trade model. This completes the proof. ■

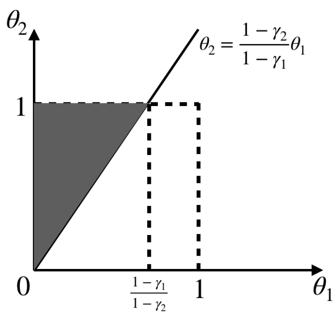
B.11 Proof of Lemma 13

Proof: We revisit the example in Section 2.3. We divide our argument into the following three cases.

Case 1: $0 < \gamma_2 < \gamma_1 < 1$

Recall that Figure 9 below illustrates the decision at different type profiles when $\gamma_2 < \gamma_1$. In particular, the shaded region in the figure represents $\Theta^* = \{(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2 : x^*(\theta_1, \theta_2) = 1\}$, which describes the set of possible type profiles under which it is efficient to trade. We will show that inequality (13) in GMO's Assumption 5.1 is violated in this case.

Figure 9: $0 < \gamma_2 < \gamma_1 < 1$



If the seller's true type is $\theta_1 = 1$ and he deviates to report $\hat{\theta}_1 = (1 - \gamma_1)/(1 - \gamma_2)$, then it is always efficient not to trade under $\hat{\theta}_1$, i.e., $x^*(\hat{\theta}_1, \theta_2) = 0$ for any $\theta_2 \in \Theta_2$. As a result, inequality (13) becomes

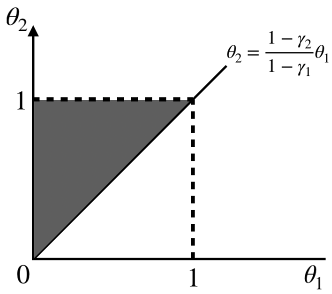
$$\mathbb{E}_{\theta_2} \left[\tilde{u}_1(\bar{\theta}_1, \theta_2) - \tilde{u}_1(\hat{\theta}_1, \theta_2) \right] \leqslant 0.$$

However, the left-hand side of the above inequality is strictly positive because $\bar{\theta}_1 > \hat{\theta}_1$ implies $\tilde{u}_1(\bar{\theta}_1, \theta_2) - \tilde{u}_1(\hat{\theta}_1, \theta_2) > 0$ by strict increasingness of $\tilde{u}_1(\cdot)$ in θ_1. This is a contradiction. Therefore, inequality (13) in GMO's Assumption 5.1 is violated in this case.

Case 2: $0 < \gamma_2 = \gamma_1 < 1$

Figure 10 illustrates the decision at different type profiles when $\gamma_1 = \gamma_2$. In particular, the shaded region represents $\Theta^* = \{(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2 : x^*(\theta_1, \theta_2) = 1\}$, which describes the set of possible type profiles under which it is efficient to trade. We will show that inequality (13) in GMO's Assumption 5.1 is satisfied in this case.

Figure 10: $0 < \gamma_1 = \gamma_2 < 1$



We first consider the seller. In this case, we know that for any $\hat{\theta}_1 < \bar{\theta}_1$, there exists $\theta_2 \in \Theta_2$ such that $x^*(\hat{\theta}_1, \theta_2) = 1$. Thus, inequality (13) can be rewritten as

$$M_1 \geqslant \frac{\mathbb{E}_{\theta_2}\left[\mathbb{1}_{\{\theta_2 | x^*(\hat{\theta}_1,\theta_2)=0\}}(\theta_2)\left(\tilde{u}_1(\bar{\theta}_1,\theta_2) - \tilde{u}_1(\hat{\theta}_1,\theta_2)\right)\right]}{\mathbb{E}_{\theta_2}\left[\mathbb{1}_{\{\theta_2 | x^*(\hat{\theta}_1,\theta_2)=1\}}(\theta_2)\right]}.$$

Since its denominator is positive and its numerator is bounded, the right-hand side of the above inequality is well defined so that we can choose M_1 appropriately. Moreover, if $\hat{\theta}_1 = \bar{\theta}_1$, then $\tilde{u}_1(\bar{\theta}_1, \theta_2) - \tilde{u}_1(\hat{\theta}_1, \theta_2) = 0$ so that the left-hand side of inequality (13) is zero. Since the right-hand side of inequality (13) is always nonnegative, there exists $M_1 > 0$ such that inequality (13) is satisfied.

Next consider the buyer. In this case, we know that for any $\hat{\theta}_2 < \bar{\theta}_2$, there exists some $\theta_1 \in \Theta_1$ such that $x^*(\hat{\theta}_1, \theta_2) = 0$. Thus, inequality (13) can be rewritten as

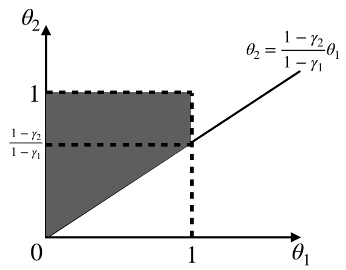
$$\tilde{M}_1 \geqslant \frac{\mathbb{E}_{\theta_1}\left[\mathbb{1}_{\{\theta_1 | x^*(\theta_1,\hat{\theta}_2)=1\}}(\theta_1)\left(\tilde{u}_2(\theta_1,\bar{\theta}_2) - \tilde{u}_2(\theta_1,\hat{\theta}_2)\right)\right]}{\mathbb{E}_{\theta_1}\left[\mathbb{1}_{\{\theta_1 | x^*(\theta_1,\hat{\theta}_2)=0\}}(\theta_1)\right]}.$$

Since its denominator is positive and its numerator is bounded, the right-hand side is well defined so that we can choose $\tilde{M}_1$ appropriately. Moreover, if $\hat{\theta}_2 = \bar{\theta}_2$, then $\tilde{u}_2(\theta_1, \bar{\theta}_2) - \tilde{u}_2(\theta_1, \hat{\theta}_2) = 0$ so that the left-hand side of inequality (13) is zero. Since the right-hand side of the above inequality is always nonnegative, there always exists $\tilde{M}_1 > 0$ such that inequality (13) is satisfied in this case.

Case 3: $0 < \gamma_1 < \gamma_2 < 1$

Figure 11 illustrates the decision at different type profiles when $\gamma_1 < \gamma_2$. In particular, the shaded region represents $\Theta^* = \{(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2 : x^*(\theta_1, \theta_2) = 1\}$, which describes the set of possible type profiles for which it is efficient to trade. We will show that inequality (13) is violated in this case.

Figure 11: $0 < \gamma_1 < \gamma_2 < 1$



If the buyer's true type is $\theta_2 = 1$ and he deviates to report $\hat{\theta}_2 = (1-\gamma_2)/(1-\gamma_1)$, then it is always efficient to trade under $\hat{\theta}_2$, i.e., $x^*(\theta_1, \hat{\theta}_2) = 1$ for any $\theta_1 \in \Theta_1$. As a result, inequality (13) becomes

$$\mathbb{E}_{\theta_1}\left[\mathbb{1}_{\{x^*(\theta_1,\hat{\theta}_2)=1\}}\left(\tilde{u}_2(\theta_1,\bar{\theta}_2) - \tilde{u}_2(\theta_1,\hat{\theta}_2)\right)\right] \leqslant 0.$$

However, the left-hand side of the above inequality is strictly positive because $\bar{\theta}_2 > \hat{\theta}_2$ implies $\tilde{u}_2(\theta_1, \bar{\theta}_2) - \tilde{u}_2(\theta_1, \hat{\theta}_2) > 0$ by the strict increasingness of $\tilde{u}_2(\cdot)$ in θ_2. This is a contradiction. Therefore, inequality (13) is violated in this case.

In the example in Section 2.3, we conclude that inequality (13) is satisfied if and only if $\gamma_1 = \gamma_2$. This completes the proof. ■

C Appendix to Chapter 3

C.1 Proof of Lemma 14

Proof: We will prove by induction that $x^*(B_i, S_{i+1}) = 0$ and $x^*(B_{i+1}, S_{i+1}) = 1$ for all $i = 1, \cdots, m-1$. First, we consider the base case where $i = 1$.

1. We show $x^*(B_1, S_2) = 0$.

   Suppose on the contrary that $x^*(B_1, S_2) = 1$. Since $x^*(B_i, S_2)$ is nondecreasing in B_i, we have $x^*(B_i, S_2) = 1$ for all i. Notice that in this case,

   $$x^*(B_i, S_1) = x^*(B_i, S_2) = 1$$

   for any i, violating the condition that the seller has no redundant type. Therefore, $x^*(B_1, S_2) = 0$ must be satisfied.

2. We show $x^*(B_2, S_2) = 1$.

   Suppose on the contrary that $x^*(B_2, S_2) = 0$. Since $x^*(B_2, S_j)$ is nonincreasing in S_j, we have $x^*(B_2, S_j) = 0$ for all $j \geqslant 3$; similarly, since $x^*(B_1, S_2) = 0$, we also have $x^*(B_1, S_j) = 0$ for all $j \geqslant 3$. Notice that in this case,

   $$x^*(B_1, S_j) = x^*(B_2, S_j)$$

   for any j, violating the condition that the buyer has no redundant type. Therefore, $x^*(B_2, S_2) = 1$ must be satisfied.

Second, we consider the inductive step and show that if $x^*(B_i, S_{i+1}) = 0$ and $x^*(B_{i+1}, S_{i+1}) = 1$, then $x^*(B_{i+1}, S_{i+2}) = 0$ and $x^*(B_{i+2}, S_{i+2}) = 1$.

1. Suppose on the contrary that $x^*(B_{i+1}, S_{i+2}) = 1$. We will show that the allocation outcomes under S_{i+1} and S_{i+2} are exactly the same and thus one of them is redundant.

   Since $x^*(B_i, S_j)$ is nonincreasing in S_j and $x^*(B_i, S_{i+1}) = 0$, we have $x^*(B_i, S_{i+2}) = 0$. Furthermore, since $x^*(B_k, S_{i+2})$ is nondecreasing in B_k and $x^*(B_i, S_{i+2}) = 0$, we have $x^*(B_k, S_{i+2}) = 0$ for any $k \leqslant i$; similarly, since $x^*(B_{i+1}, S_{i+2}) = 1$, we obtain $x^*(B_k, S_{i+2}) = 1$ for any $k \geqslant i+1$. Lastly, since $x^*(B_k, S_{i+1})$ is nondecreasing in B_k and $x^*(B_i, S_{i+1}) = 0$, we have $x^*(B_k, S_{i+1})$ for any $k \leqslant i$; similarly, since $x^*(B_{i+1}, S_{i+1}) = 1$, we have $x^*(B_k, S_{i+1}) = 1$ for any $k \geqslant i+1$. The above argument is also illustrated in the following table.

$x^*(\cdot)$	S_m	$\cdots$	S_{i+2}	S_{i+1}	$\cdots$	S_1
B_m			1	1		1
$\vdots$			$\vdots$	$\vdots$		$\vdots$
B_{i+1}			1	1		1
B_i			0	0		1
$\vdots$			$\vdots$	$\vdots$		$\vdots$
B_1	0		0	0		1

To summarize,

$$x^*(B_k, S_{i+1}) = x^*(B_k, S_i)$$

for any k, violating the condition that the seller has no redundant type. Therefore, $x^*(B_{i+1}, S_{i+2}) = 0$ must be satisfied.

2. Suppose on the contrary that $x^*(B_{i+2}, S_{i+2}) = 0$. We will show that the allocation outcomes under B_{i+2} and B_{i+1} are exactly the same and thus one of them is redundant.

   Since $x^*(B_k, S_i)$ is nondecreasing in B_k and $x^*(B_{i+1}, S_{i+1}) = 1$, we have $x^*(B_{i+2}, S_{i+1}) = 1$. Moreover, since $x^*(B_{i+2}, S_j)$ is nonincreasing in S_j and $x^*(B_{i+2}, S_{i+2}) = 0$, we have $x^*(B_{i+2}, S_j) = 0$ for all $j \geqslant i+2$; similarly, since $x^*(B_{i+2}, S_{i+1}) = 1$, we have $x^*(B_{i+2}, S_{i+1}) = 1$ for all $j \neq i+1$. Lastly, since $x^*(B_{i+1}, S_j)$ is nonincreasing in S_j and $x^*(B_{i+1}, S_{i+2}) = 0$, we have $x^*(B_{i+1}, S_j) = 0$ for all $j \geqslant i+2$; similarly, since $x^*(B_{i+1}, S_{i+1}) = 1$, we have $x^*(B_{i+1}, S_j) = 1$ for all $j \leqslant i+1$. The above argument is also illustrated in the following table.

   To summarize,

$$x^*(B_{i+2}, S_j) = x^*(B_{i+1}, S_j)$$

   for any j, violating the condition that the buyer has no redundant type. Therefore, $x^*(B_{i+2}, S_{i+2}) = 1$ must be satisfied.

$x^*(\cdot)$	S_m	$\cdots$	S_{i+2}	S_{i+1}	$\cdots$	S_1
B_m						1
$\vdots$						$\vdots$
B_{i+2}	0	$\cdots$	0	1	$\cdots$	1
B_{i+1}	0	$\cdots$	0	1	$\cdots$	1
$\vdots$						$\vdots$
B_1	0					1

Finally, since both the base case and the inductive steo have been performed, by mathematical induction, $x^*(B_i, S_{i+1}) = 0$ and $x^*(B_{i+1}, S_{i+1}) = 1$ hold for all $i = 1, \cdots, m-1$. ■

C.2 Proof of Lemma 15

Proof: Notice that inequality (15) is equal to the summation of the constant terms in inequalities (16), (17), (18) and (19). Below we will show that by adding up inequalities (16) to (19), all the terms regarding the payment $t(B_i, S_j)$ where $i \geq j$ are cancelled out. To do so, we can divide the the payment into three groups: $t(B_1, S_1)$, $t(B_m, S_m)$, and the payment at all the other type profiles.

1. For the payment $t(B_1, S_1)$:

   Notice that the payment$t(B_1, S_1)$ appears in inequality (16) with coefficient $-\sum_{i=1}^m g(B_i, S_1)$; it also appears in inequality (18) with coefficient $\sum_{i=2}^m g(B_i, S_1)$; finally, it appears in inequality (19) with coefficient $g(B_1, S_1)$. Therefore, the summation of the coefficients is

   $$-\sum_{i=1}^m g(B_i, S_1) + \sum_{i=2}^m g(B_i, S_1) + g(B_1, S_1) = 0,$$

   and hence the terms regarding $t(B_1, S_1)$ are cancelled out.

2. For the payment $t(B_m, S_m)$:

   Notice that the payment$t(B_m, S_m)$ appears in inequality (17) with coefficient $\sum_{j=1}^m g(B_m, S_j)$; it also appears in inequality (18) with coefficient $-g(B_m, S_m)$; finally, it appears in inequality (19) with coefficient $-\sum_{j=1}^{m-1} g(B_m, S_j)$. Therefore, the summation of the coefficients is

$$\sum_{j=1}^{m} g(B_m, S_j) - g(B_m, S_m) - \sum_{j=1}^{m-1} g(B_m, S_j) = 0,$$

and hence the terms regarding $t(B_m, S_m)$ are cancelled out.

3. For all the other payment $t(B_i, S_j)$:

   Notice that any other payment $t(B_i, S_j)$ appears in inequality (18) with coefficient $-g(B_i, S_j)$; it also appears in inequality (19) with coefficient $g(B_i, S_j)$. Therefore, the summation of the coefficients is

$$-g(B_i, S_j) + g(B_i, S_j) = 0,$$

   and hence the terms regarding all the other payments $t(B_i, S_j)$ are cancelled out.

We conclude that the summation of inequalities (16) to (19) is equal to the summation of the constant terms in these inequalities and thus we obtain inequality (15). ■