

# Mechanism Design with Ambiguous Transfers

Huiyi Guo\*

Department of Economics, University of Iowa

huiyi-guo@uiowa.edu

First draft: November 4, 2016

Current draft: September 15, 2017

## Abstract

This paper introduces ambiguous transfers to study the problem of full surplus extraction and (partial) implementation of efficient allocations. We show that full surplus extraction can be guaranteed via a mechanism with ambiguous transfers if and only if the Beliefs Determine Preferences (BDP) property is satisfied by all agents. We also show that any efficient allocation rule is implementable via an interim individually rational and ex-post budget-balanced mechanism with ambiguous transfers if and only if the BDP property holds for all agents. This property is generic in a type space with at least two agents. It is weaker than the necessary and sufficient conditions for full surplus extraction and implementation via Bayesian mechanisms. Therefore, ambiguous transfers may provide a solution for situations where Bayesian mechanism design is impossible. In particular, with ambiguous transfers, efficient allocations become implementable generically in two-agent problems, which contrasts the impossibility results in the literature.

**Keywords:** Full surplus extraction; Bayesian (partial) implementation; Ambiguous transfers; Correlated beliefs; Individual rationality; Budget balance

---

\*I am indebted to Nicholas Yannelis for his guidance and encouragement. I also thank Rabah Amir, Chris Chambers, Yi-Chun Chen, Yifan Dai, Luciano De Castro, Wei He, Toomas Hinnoaar, Jiangtao Li, Heng Liu, Mallesh Pai, Max Stinchcombe, and conference participants at 2017 Asian Meeting of the Econometric Society and 2017 China Meeting of the Econometric Society for helpful comments.

# 1 Introduction

One could perceive ambiguity in many forms of transactions. For instance, Priceline Express Deals offer travelers a known price for a hotel stay, but the exact name of the hotel remains unknown until the completion of payment. Another example is the scratch-and-save promotion run by some stores: a consumer receives a scratch card at the time of check out, which reveals a discount, and thus the cost of her purchase and even its distribution are mysterious when she decides to buy. As a third example, eBay allows sellers of auction-style listings to set hidden reserve prices. In all the above mechanisms, the ambiguity about the allocation or (and) transfer rule are generated by the mechanism designers. It is reasonable to imply that the mechanism designers engineer ambiguity to benefit themselves. To know how far such a practice can go, we would like to ask the specific question that if introducing ambiguity can help the mechanism designers achieve the first-best outcomes.

This paper introduces ambiguous transfers to study two problems: full surplus extraction and implementation of an efficient allocation rule via an interim individually rational and ex-post budget-balanced mechanism. In both problems, the mechanism designer informs agents of the exact allocation rule, but the communication is ambiguous so that agents only know a set of potential transfer rules. Agents are assumed to be ambiguity-averse and thus make decisions based on the worst-case transfer rule.

In this paper, the Beliefs Determine Preferences (BDP) property is the key condition for first-best mechanism design with ambiguous transfers. The property, introduced by Neeman (2004), requires that an agent with different types should have distinct beliefs. We show that full surplus extraction can be guaranteed via a mechanism with ambiguous transfers if and only if the BDP property is satisfied by all agents. In addition, any efficient allocation rule is implementable via an interim individually rational and ex-post budget-balanced mechanism with ambiguous transfers if and only if the BDP property holds for all agents. These two are the primary results of this paper. As an extension, we show any efficient allocation rule under a private value environment is implementable if and only if at most one agent does not satisfy the BDP property. Then, we investigate an environment without a common prior and establish sufficient conditions for individually rational and budget-balanced implementation of any efficient allocation rule under private value environments.

Our key condition, the BDP property, is weaker than Crémer and McLean (1988)'s Convex Independence condition, which is necessary and sufficient for full surplus extraction via a Bayesian mechanism. Convex Independence, together with the Identifiability condition established by Kosenok and Severinov (2008), is necessary and sufficient for implementing any efficient allocation rule via an individually rational and budget-balanced Bayesian im-

plementation. In a type space with fixed finite dimension and more than one agent, the BDP property holds generically. Without restricting the dimension, models satisfying the BDP property are topologically generic, as is shown by Chen and Xiong (2011).

We summarize several advantages of the BDP property below. Firstly, compared to Convex Independence, the BDP property imposes weaker restrictions on the cardinality of the type space. For example, in a two agent problem, where one agent has two types and the other has three, Convex Independence fails for one agent for sure, but the BDP property holds for both generically. Secondly, the Identifiability condition is relaxed, and hence so is its associated restriction on cardinality of the type space.<sup>1</sup> For example, in a three-agent problem where each agent has two types, the Identifiability property fails with positive probability, but the BDP property holds generically. Thirdly, the Bayesian mechanism design literature documents several negative results on individually rational and budget-balanced implementation with two agents, but the BDP property and ambiguous transfers provide a generic solution to such problems, which are fundamental and important in view of the many bilateral trades and bargains occurring every day.<sup>2</sup> Fourthly, the BDP property is very easy to check. To verify this property for an agent, we only need to make sure that she never has identical beliefs under different types.

In this paper, we let the mechanism designer announce a fixed efficient allocation rule and introduce ambiguity in transfer rules only. To see why we impose this restriction, notice that the allocation rule in an implementation problem is exogenous, and thus, it is natural for the mechanism designer to commit to that particular allocation rule. When the mechanism designer aims to extract full surplus instead, she endogenously chooses an ex-post efficient allocation rule, which is often unique in a finite-type framework. Hence, we do not give the mechanism designer the freedom to use ambiguous allocation rules in full surplus extraction either. In a related paper, Di Tillio et al. (2017) study how second-best outcomes under independent beliefs could be improved if the mechanism designer introduces ambiguity in both allocation and transfer rules. We discuss more on the relationship with that paper in Section 1.1. As a by-product, the restriction on no ambiguity in allocation rules also helps us to clarify the scope and limitation of ambiguous transfers.

The paper proceeds as follows. We review the literature in Section 1.1 and introduce

---

<sup>1</sup>For the first two points, see Section 5 for more details.

<sup>2</sup>For example, Myerson and Satterthwaite (1983) demonstrate the impossibility of efficient bilateral trading with independent information. Matsushima (2007) provides a sufficient condition under which individually rational and budget-balanced implementation with two agents cannot be achieved. Kosenok and Severinov (2008)'s necessary and sufficient conditions never hold simultaneously in two-agent environments, which could also be interpreted as an impossibility result even if correlated information is allowed.

the environment in Section 2. After providing an example on how ambiguous transfers work for full surplus extraction and implementation in Section 3, we formalize the mechanism with ambiguous transfers in Section 4. The BDP property is introduced and shown to be necessary and sufficient for full surplus extraction and implementation in Section 5. Section 6 extends our primary results along several directions. The Appendix collects all proofs and some examples.

## 1.1 Literature review

### 1.1.1 Efficient mechanisms with independent information

How to implement efficient allocations is a classical topic in mechanism design theory that has been widely studied in situations such as public good provision and bilateral trading. Individual rationality is a natural requirement as agents can opt out of the mechanism. As a resource constraint, budget balance requires that agents should finance within themselves for the efficient outcome rather than rely on an outside budget-breaker. When either individual rationality or budget balance is required, the literature provides positive results for efficient mechanism design in private value environments. For instance, the VCG mechanism (Vickrey (1961), Clarke (1971), and Groves (1973)) is ex-post individually rational. The AGV mechanism (d’Aspremont and Gérard-Varet (1979)) is ex-post budget-balanced.

However, the literature documents a tension between efficiency, individual rationality, and budget balance, when agents have independent information. For example, in a private value bilateral trading framework, Myerson and Satterthwaite (1983) prove that it is impossible to achieve efficiency with an individually rational and budget-balanced mechanism in general. With multi-dimensional and interdependent values, Dasgupta and Maskin (2000) and Jehiel and Moldovanu (2001) prove that efficient allocations are generically non-implementable.

One goal of the current paper is to design an efficient, individually rational, and budget-balanced mechanism. But instead of assuming independent information, we show that correlation is necessary and sufficient to achieve the goal.

### 1.1.2 Mechanism design with correlated information

With correlated information, first-best mechanism design becomes possible. Crémer and McLean (1985, 1988) establish two conditions to fully extract agents’ surplus in private value auctions, among which the Convex Independence condition is necessary and sufficient for Bayesian mechanism design. In a fixed finite dimension type space, if there are at least two agents, and if no agent has more types than all others’ type profiles, the condition is

satisfied generically. Without restricting the dimension, different notions of genericity are adopted in the literature and various conclusions on genericity of Convex Independence are made (e.g., Neeman (2004), Heifetz and Neeman (2006), Barelli (2009), Chen and Xiong (2011, 2013)). With continuous types, McAfee and Reny (1992) show that approximate full surplus extraction can be achieved. In addition, the recent papers of Liu (2014) and Noda (2015) prove an intertemporal variant of Convex Independence is sufficient for first-best mechanism design in dynamic environments. In Sections 5.1, by introducing ambiguous transfers, the current paper shows that the weaker condition, the BDP property, becomes necessary and sufficient for full surplus extraction.

Unlike full surplus extraction, in a problem of (partial) implementation, the allocation rule is exogenously given, and the mechanism designer constructs incentive compatible transfers to achieve the desired outcome. Under the context of exchange economies, McLean and Postlewaite (2002, 2003a,b) propose the notion of informational size and prove the existence of incentive compatible and approximately efficient outcomes when agents have small informational size.<sup>3</sup> Under a mechanism design framework, McLean and Postlewaite (2004, 2015) implement efficient allocation rules via individually rational mechanisms under the BDP property. In their mechanisms, small outside money is needed even when agents are informationally small. We do not address the issue of informational size, but our mechanism for implementation in Section 5 is exactly efficient, individually rational, and budget-balanced.

A few papers study budget-balanced mechanisms with or without independent information, including Matsushima (1991), Aoyagi (1998), Chung (1999), d’Aspremont et al. (2004), etc.<sup>4</sup> Among these works, d’Aspremont et al. (2004) propose necessary and sufficient conditions for budget-balanced mechanisms. All these papers do not require individual rationality. Also, they assume that there are at least three agents. Actually, d’Aspremont et al. (2004) indicate an impossibility result in implementing efficient allocations via budget-balanced mechanisms with two agents under correlated information. However, we do require individual rationality, and our mechanism with ambiguous transfers works for environments with at least two agents.

Matsushima (2007) and Kosenok and Severinov (2008) design individually rational and budget-balanced mechanisms. The latter work proposes the Identifiability condition, which along with the Convex Independence condition, is necessary and sufficient for implementing any ex-ante socially rational allocation rule via an individually rational and budget-balanced

---

<sup>3</sup>For related results, see also Sun and Yannelis (2007, 2008).

<sup>4</sup> Matsushima (1991), Chung (1999), and d’Aspremont et al. (2004) only consider private value utility functions. In this case, incentive compatibility can be achieved via a VCG mechanism, rather than via information correlation. Thus, they allow for independent information.

Bayesian mechanism. The Identifiability condition is generic with at least three agents and under some restrictions on the dimension of agents' types, but Convex Independence and Identifiability never hold simultaneously in a two-agent setting. Thus Kosenok and Severinov (2008) imply an impossibility result in efficient, individually rational, and budget-balanced two-agent mechanism design. In our paper, the BDP property is weaker than Convex Independence, and we do not need Identifiability. Moreover, the BDP property holds generically with at least two agents, and thus we make the impossible possible for two-agent implementation problems.

### 1.1.3 Mechanism design under ambiguity

In the growing literature on mechanism design with ambiguity-averse agents, most of the works assume exogenously that agents hold ambiguous beliefs of others' types. For example, Bose et al. (2006) prove that when agents are more ambiguity-averse than the auctioneer, a full insurance transfer rule is optimal in a private value auction. Bose and Daripa (2009) achieve almost full surplus extraction in a dynamic auction by exploiting the dynamic inconsistency of prior-by-prior updating. Bodoh-Creed (2012) characterizes the revenue-maximizing mechanism with a payoff equivalence theorem. De Castro and Yannelis (2009) prove that all Pareto efficient allocations are incentive compatible when agents' ambiguous beliefs are unrestricted. Accordingly, De Castro et al. (2017a,b) implement all Pareto efficient allocations. Under the private value assumption, Wolitzky (2016) establishes a necessary condition for the existence of an efficient, individually rational, and weak budget-balanced mechanism. In an environment with multi-dimensional and interdependent values, Song (2016) quantifies the amount of ambiguity that is necessary and sometimes sufficient for efficient mechanism design. We do not assume exogenous ambiguity in agents' beliefs, which is the biggest difference between the above papers and our work.

Bose and Renou (2014) and Di Tillio et al. (2017) contrast the above works in that ambiguity is endogenously engineered by the mechanism designer. Before the allocation stage, Bose and Renou (2014) let the mechanism designer communicate with agents via an ambiguous device, which generates ambiguous beliefs. Their paper characterizes social choice functions that are implementable under this method. Our paper is different from Bose and Renou (2014), as we do not need multiple beliefs.

Di Tillio et al. (2017) consider the problem of revenue maximization in a private value and independent belief environment. The seller commits to a simple mechanism, i.e., an allocation and transfer rule, but informs agents of a set of simple mechanisms. As all the simple mechanisms generate the same expected revenue (imposed by the Consistency

condition), agents do not know the exact rule and thus make decisions based on the worst-case scenario. Compared to the standard Bayesian mechanism, their ambiguous approach yields a higher expected revenue.

In the current paper, ambiguity is engineered in a similar way to Di Tillio et al. (2017). However, instead of studying how ambiguous mechanisms improve second-best revenues under independent beliefs, the current paper studies when the first-best outcome in surplus extraction or implementation can be achieved without restricting attention to independent beliefs. As mentioned before, we fix an efficient allocation rule and only allow for ambiguity in transfer rules, but in Di Tillio et al. (2017)'s mechanism both allocation and transfer rules are ambiguous. This restriction is compatible with Di Tillio et al. (2017)'s Consistency condition, because each transfer rule satisfies interim individual rationality and extracts full surplus (or is interim individually rational and ex-post budget-balanced in the implementation problem). To see that allowing for ambiguity in allocation rules may result in a failure of full surplus extraction or implementation, we consider a finite-type environment where the total surplus is maximized by a unique allocation rule for example. If it is common knowledge that the mechanism designer's objective is to extract full surplus or implement efficient outcomes, any other allocation with lower surplus levels are non-credible to the agents, and thus should not be used in the mechanism. In Di Tillio et al. (2017)'s framework with independent beliefs and finitely many types, ambiguity in allocation rules plays an important role to achieve incentive compatibility, and therefore, full surplus extraction cannot be achieved. But with continuous types, their approach works for full surplus extraction, and thus we focus on environments with finitely many types.

The essential factor that enables us to achieve the first-best outcome in finite type environment is the correlation in agents' beliefs. In fact, we show correlated beliefs are necessary and sufficient for full surplus extraction and implementing efficient allocations. Correlation also results in different constructions of mechanisms between Di Tillio et al. (2017) and the current paper: in the main section of our paper (Section 5.2), we only need two transfer rules, while the number of simple mechanisms in their paper depends on the cardinality of the type space.

## 2 Asymmetric information environment

We study the asymmetric information environment given by

$$\mathcal{E} = \{I, A, (\Theta_i, u_i)_{i=1}^N, p\},$$

where

- $I = \{1, \dots, N\}$  is the finite set of agents; assume  $N \geq 2$ ;
- $A$  is the set of **feasible outcomes**;
- let  $\theta_i \in \Theta_i$  be agent  $i$ 's **type**; denote  $\times_{i \in I} \Theta_i$  by  $\Theta$ ,  $\times_{j \in I, j \neq i} \Theta_j$  by  $\Theta_{-i}$ , and  $\times_{k \in I, k \neq i, j} \Theta_k$  by  $\Theta_{-i, j}$ ; let  $|\Theta_i|$  represent the cardinality of  $\Theta_i$ , where we assume  $2 \leq |\Theta_i| < \infty$ ;<sup>5</sup>
- each agent  $i$  has a quasi-linear **utility function**  $u_i(a, \theta) + b$ , where  $a \in A$  is a feasible outcome,  $b \in \mathbb{R}$  is a monetary transfer, and  $\theta \in \Theta$  is the realized type profile;
- $p$  is a probability distribution on  $\Theta$ , representing agents' **common prior**; let  $p(\theta_i)$  and  $p(\theta_i, \theta_j)$  represent the marginal distribution of  $p$  on  $\theta_i$  and  $(\theta_i, \theta_j)$  respectively; when agent  $i$  has type  $\theta_i$ , her **belief** is derived from Bayesian updating  $p$ , i.e., others have type profile  $\theta_{-i} \in \Theta_{-i}$  with probability  $p_i(\theta_{-i}|\theta_i)$ ; for agent  $j \neq i$  and type  $\theta_j$  we let  $p_i(\theta_j|\theta_i)$  denote the marginal belief of  $p_i(\cdot|\theta_i) \equiv (p_i(\theta_{-i}|\theta_i))_{\theta_{-i} \in \Theta_{-i}}$  on type  $\theta_j$ .

The environment  $\mathcal{E}$  is common knowledge between the mechanism designer and the agents.

We impose the following assumption throughout the paper unless otherwise specified.

**Assumption 2.1:** For all  $i, j \in I$  with  $i \neq j$ , and  $(\theta_i, \theta_j) \in \Theta_i \times \Theta_j$ , assume  $p(\theta_i, \theta_j) > 0$ .

An **allocation rule**  $q : \Theta \rightarrow A$  is a plan to assign a feasible outcome contingent on agents' realized type profile. An allocation rule  $q$  is said to be ex-post **efficient** if  $\sum_{i \in I} u_i(q(\theta), \theta) \geq$

$\sum_{i \in I} u_i(q'(\theta), \theta)$  for all  $q' : \Theta \rightarrow A$  and  $\theta \in \Theta$ .

### 3 A motivating example

In this example, we look at a common prior  $p$  such that the standard Bayesian mechanism design approach can neither guarantee full surplus extraction nor implementation of an efficient allocation via an interim individually rational and ex-post budget-balanced mechanism. However, when the mechanism designer is allowed to be ambiguous about which transfer rule she adopts, we show that both goals can be achieved.

We assume there are two agents and each agent has three types. The common prior  $p \in \Delta(\Theta)$  is defined below.

---

<sup>5</sup>The assumption that  $|\Theta_i| \geq 2$  for all  $i$  is imposed for simplicity of notation. When at least two agents satisfy this cardinality condition, i.e., when at least two agents have private information, all theorems of this paper hold. See Appendix A.2 for more details.

$p$	$\theta_2^1$	$\theta_2^2$	$\theta_2^3$
$\theta_1^1$	$\frac{1}{18}$	$\frac{2}{18}$	$\frac{3}{18}$
$\theta_1^2$	$\frac{3}{18}$	$\frac{2}{18}$	$\frac{1}{18}$
$\theta_1^3$	$\frac{2}{18}$	$\frac{2}{18}$	$\frac{2}{18}$

### 3.1 Full surplus extraction

The belief of  $\theta_1^3$  is a convex combination of  $\theta_1^1$  and  $\theta_1^2$ . Therefore, the Convex Independence condition of Crémer and McLean (1988) fails. We briefly sketch their argument to see why full surplus extraction is impossible via a Bayesian mechanism in an auction with private values satisfying  $\theta_1^1 > \theta_1^2 > \theta_1^3 > \theta_2 > 0$  for all  $\theta_2 \in \Theta_2$ . Suppose by way of contradiction that a transfer rule to agents,  $\phi = (\phi_1, \phi_2) : \Theta \rightarrow \mathbb{R}^2$ , extracts full surplus. To maximize social surplus, the good should always be allocated to agent 1. As agent 1 obtains zero surplus at every type, incentive compatibility implies:

$$\begin{aligned} IC(\theta_1^1 \theta_1^3) & \quad 0 \geq \theta_1^1 + \frac{1}{6}\phi_1(\theta_1^3, \theta_2^1) + \frac{2}{6}\phi_1(\theta_1^3, \theta_2^2) + \frac{3}{6}\phi_1(\theta_1^3, \theta_2^3), \\ IC(\theta_1^2 \theta_1^3) & \quad 0 \geq \theta_1^2 + \frac{3}{6}\phi_1(\theta_1^3, \theta_2^1) + \frac{2}{6}\phi_1(\theta_1^3, \theta_2^2) + \frac{1}{6}\phi_1(\theta_1^3, \theta_2^3). \end{aligned}$$

Averaging them yields  $0 \geq \frac{1}{2}\theta_1^1 + \frac{1}{2}\theta_1^2 + \frac{2}{6}\phi_1(\theta_1^3, \theta_2^1) + \frac{2}{6}\phi_1(\theta_1^3, \theta_2^2) + \frac{2}{6}\phi_1(\theta_1^3, \theta_2^3) > \theta_1^3 + \frac{2}{6}\phi_1(\theta_1^3, \theta_2^1) + \frac{2}{6}\phi_1(\theta_1^3, \theta_2^2) + \frac{2}{6}\phi_1(\theta_1^3, \theta_2^3)$ , a contradiction, as type- $\theta_1^3$  agent 3 should have non-negative payoff.

Next, we see how full surplus extraction can be guaranteed when ambiguous transfers are allowed. Suppose the mechanism designer writes down either  $\phi^1$  or  $\phi^2$  on a piece of paper, which is the true transfer rule. She hides the paper, only tells agents that  $\phi^1$  and  $\phi^2$  are the two potential rules, and announces that agent 1 will get the good. After agents report their types, the mechanism designer reveals the paper with the true rule, makes transfers accordingly, and allocates the good.

Let the transfer rules to agents be  $\phi^1 = (\phi_1^1, \phi_2^1)$  and  $\phi^2 = (\phi_1^2, \phi_2^2)$ :

$$\phi_i^1(\theta_1, \theta_2) = \begin{cases} -\theta_1 + c\psi(\theta_1, \theta_2), & \text{if } i = 1, \\ -c\psi(\theta_1, \theta_2), & \text{if } i = 2, \end{cases} \quad \phi_i^2(\theta_1, \theta_2) = \begin{cases} -\theta_1 - c\psi(\theta_1, \theta_2), & \text{if } i = 1, \\ c\psi(\theta_1, \theta_2), & \text{if } i = 2, \end{cases}$$

where  $c \geq 2(\theta_1^1 - \theta_1^3)$  and the function  $\psi : \Theta \rightarrow \mathbb{R}$  is defined below.

$\psi$	$\theta_2^1$	$\theta_2^2$	$\theta_2^3$
$\theta_1^1$	6	0	-2
$\theta_1^2$	-2	3	0
$\theta_1^3$	0	-3	3

Notice that when a type- $\bar{\theta}_i$  agent  $i$  truthfully reports her type,  $\psi(\bar{\theta}_i, \cdot)$  has an expected value of zero under the belief  $p_i(\cdot|\bar{\theta}_i)$ . However, when she misreports  $\hat{\theta}_i \neq \bar{\theta}_i$ ,  $\psi(\hat{\theta}_i, \cdot)$  has a non-zero expected value. Therefore, among the expected values of  $c\psi(\hat{\theta}_i, \cdot)$  and  $-c\psi(\hat{\theta}_i, \cdot)$ , the lower one is negative. This feature gives ambiguity-averse agents the incentive to truthfully reveal their types.

To formally prove that full surplus extraction can be achieved, we first notice that both  $\phi^1$  and  $\phi^2$  give the mechanism designer the expected total surplus,  $\frac{1}{3}(\theta_1^1 + \theta_1^2 + \theta_1^3)$ , and that both agents obtain expected payoff of zero.

Then we check the incentive compatibility condition. When type- $\bar{\theta}_2$  agent 2 misreports  $\hat{\theta}_2 \neq \bar{\theta}_2$ , her expected payoffs according to  $\phi^1$  and  $\phi^2$  are  $\pm \sum_{\theta_1 \in \Theta_1} c\psi(\theta_1, \hat{\theta}_2)p_2(\theta_1|\bar{\theta}_2)$ . The worst case is non-positive, which is not better than truthfully revealing. When type- $\bar{\theta}_1$  agent 1 misreports  $\hat{\theta}_1 \neq \bar{\theta}_1$ , her worst case expected payoff is  $\min\{\bar{\theta}_1 - \hat{\theta}_1 \pm c \sum_{\theta_2 \in \Theta_2} \psi(\hat{\theta}_1, \theta_2)p_1(\theta_2|\bar{\theta}_1)\} \leq \bar{\theta}_1 - \hat{\theta}_1$ . Therefore, any ‘‘upward’’ misreport is not profitable. It remains to verify the three ‘‘downward’’ incentive compatibility constraints:

$$\begin{aligned} IC(\theta_1^1\theta_1^2) & 0 \geq \theta_1^1 - \theta_1^2 - c|\frac{1}{6} \times (-2) + \frac{2}{6} \times 3 + \frac{3}{6} \times 0| = \theta_1^1 - \theta_1^2 - \frac{2}{3}c, \\ IC(\theta_1^1\theta_1^3) & 0 \geq \theta_1^1 - \theta_1^3 - c|\frac{1}{6} \times 0 + \frac{2}{6} \times (-3) + \frac{3}{6} \times 3| = \theta_1^1 - \theta_1^3 - \frac{1}{2}c, \\ IC(\theta_2^2\theta_2^3) & 0 \geq \theta_1^2 - \theta_1^3 - c|\frac{3}{6} \times 0 + \frac{2}{6} \times (-3) + \frac{1}{6} \times 3| = \theta_1^2 - \theta_1^3 - \frac{1}{2}c. \end{aligned}$$

They hold because  $c \geq 2(\theta_1^1 - \theta_1^3)$  and  $\theta_1^1 > \theta_1^2 > \theta_1^3$ . Therefore, in this example, full surplus can be extracted via ambiguous transfers.

Intuitively, with multiple transfer rules, an agent’s worst-case expected payoffs of different misreports are achieved by distinct transfers. Compared to Bayesian mechanisms, we do not need one transfer rule to satisfy all incentive compatibility constraints. Hence, full surplus can be extracted under a weaker condition than Convex Independence.

Notice that for each  $i$  and  $\bar{\theta}_i \neq \hat{\theta}_i$ ,  $p_i(\cdot|\bar{\theta}_i) \neq p_i(\cdot|\hat{\theta}_i)$ , i.e.,  $i$ ’s beliefs under distinct types are different. This fact plays an essential role for ambiguous transfers to work. To see this, consider an alternative common prior  $\tilde{p}$  satisfying  $\tilde{p}_1(\cdot|\theta_1^1) = \tilde{p}_1(\cdot|\theta_1^2)$  and suppose by way of contradiction that full surplus extraction is guaranteed by a set of ambiguous transfers  $\tilde{\Phi}$ . Then by truthfully revealing, every agent, obtains zero expected payoff, in particular, type- $\theta_1^2$  agent 1 has expected payoff  $0 = \theta_1^2 + \inf_{\tilde{\phi} \in \tilde{\Phi}} \sum_{\theta_2 \in \Theta_2} \tilde{\phi}_1(\theta_1^2, \theta_2)\tilde{p}_1(\theta_2|\theta_1^2)$ . By misreporting, every agent obtains non-positive expected payoff, in particular, type- $\theta_1^1$  agent 1 has the following expected payoff of misreporting  $\theta_1^2$ ,  $\theta_1^1 + \inf_{\tilde{\phi} \in \tilde{\Phi}} \sum_{\theta_2 \in \Theta_2} \tilde{\phi}_1(\theta_1^1, \theta_2)\tilde{p}_1(\theta_2|\theta_1^1) \leq 0$ . As  $\tilde{p}_1(\cdot|\theta_1^1) = \tilde{p}_1(\cdot|\theta_1^2)$ , the two expressions imply  $\theta_1^1 \leq \theta_1^2$ , a contradiction.

## 3.2 Implementation

The common prior  $p$  satisfies neither the Convex Independence condition nor the Identifiability condition of Kosenok and Severinov (2008). Therefore, one can follow their approach to construct utility functions such that an efficient allocation rule is not implementable. Agents 1 and 2 face a feasible set of alternatives  $A = \{x_0, x_1, x_2\}$ . The outcome  $x_0$  gives both agents zero payoffs at all type profiles. The payoffs given by allocation rules  $x_1$  and  $x_2$  are presented below, where the first component denotes agent 1's payoff and the second denotes 2's. We assume  $0 < 3a < B$ .

$x_1$	$\theta_2^1$	$\theta_2^2$	$\theta_2^3$	$x_2$	$\theta_2^1$	$\theta_2^2$	$\theta_2^3$
$\theta_1^1$	$a, 0$	$a, a$	$a, a$	$\theta_1^1$	$a, a$	$a - 2B, a + B$	$a, 0$
$\theta_1^2$	$a, 0$	$a, a$	$a, a$	$\theta_1^2$	$a, a$	$a - 2B, a + B$	$a, 0$
$\theta_1^3$	$a, 0$	$a, a$	$a, a$	$\theta_1^3$	$a, a$	$a - 2B, a + B$	$a, 0$

The efficient allocation rule is  $q(\theta_1, \theta_2^1) = x_2$  and  $q(\theta_1, \theta_2^2) = q(\theta_1, \theta_2^3) = x_1$  for all  $\theta_1 \in \Theta_1$ . To see  $q$  is not implementable via an interim individually rational and ex-post budget-balanced Bayesian mechanism, we suppose by way of contradiction that there is a transfer rule  $\phi : \Theta \rightarrow \mathbb{R}^N$  implementing  $q$ . Adding all the twelve incentive compatibility constraints and taking into account ex-post budget balance yield  $3a \geq B$ , a contradiction.

Next, we construct ambiguous transfers to implement  $q$ . Let  $\phi^1$  and  $\phi^2$  be the two potential transfer rules defined by  $\phi_1^1(\theta) = \phi_2^2(\theta) = c\psi(\theta)$  and  $\phi_2^1(\theta) = \phi_1^2(\theta) = -c\psi(\theta)$  for all  $\theta \in \Theta$ , where  $c \geq 0.75B$  and  $\psi$  is defined in the previous subsection. Note that both  $\phi^1$  and  $\phi^2$  are ex-post budget-balanced.

Type- $\bar{\theta}_i$  agent  $i$ 's interim individual rationality holds, because (1)  $u_i(q(\bar{\theta}_i, \theta_{-i}), (\bar{\theta}_i, \theta_{-i})) = a$  for all  $i \in I$  and  $(\bar{\theta}_i, \theta_{-i}) \in \Theta$  and (2)  $c\psi_i(\bar{\theta}_i, \cdot)$ 's expected value is 0. When type- $\bar{\theta}_i$  agent misreports  $\hat{\theta}_i \neq \bar{\theta}_i$ , her worst-case expected payoff satisfies

$$\min \sum_{\theta_j \in \Theta_j} \left( u_i(q(\hat{\theta}_i, \theta_j), (\bar{\theta}_i, \theta_j)) \pm \phi_i(\hat{\theta}_i, \theta_j) \right) p_i(\theta_j | \bar{\theta}_i) \leq \sum_{\theta_j \in \Theta_j} \left( u_i(q(\hat{\theta}_i, \theta_j), (\bar{\theta}_i, \theta_j)) \right) p_i(\theta_j | \bar{\theta}_i).$$

When  $(\bar{\theta}_i, \hat{\theta}_i) \neq (\theta_2^2, \theta_2^1)$ , such a misreport gives the agent the same or a lower payoff, and therefore, the above inequality implies that  $IC(\bar{\theta}_i, \hat{\theta}_i)$  holds. For  $IC(\theta_2^2, \theta_2^1)$ , such a misreport gives  $i$  worst-case expected payoff is  $a + B - c \left[ \frac{2}{6} \times 6 + \frac{2}{6} \times (-2) + \frac{2}{6} \times 0 \right] = a + B - \frac{4}{3}c \leq a$ . Therefore, the interim individually rational and ex-post budget-balanced mechanism with ambiguous transfers implements  $q$ .

Again, the fact that beliefs are different for distinct types plays an essential role. To see this, consider an alternative common prior  $\tilde{p}$  satisfying  $\tilde{p}_2(\cdot | \theta_2^1) = \tilde{p}_2(\cdot | \theta_2^2)$ . Suppose by way of

contradiction that the interim individually rational and ex-post budget-balanced ambiguous transfers  $\tilde{\Phi}$  implement  $q$ . Then the following inequalities hold:

$$\begin{aligned}
IC(\theta_2^1 \theta_2^2) \quad & a + \min_{\tilde{\phi} \in \tilde{\Phi}} \sum_{\theta_1 \in \Theta_1} \tilde{\phi}_2(\theta_1, \theta_2^1) \tilde{p}_2(\theta_1 | \theta_2^1) \geq \min_{\tilde{\phi} \in \tilde{\Phi}} \sum_{\theta_1 \in \Theta_1} \tilde{\phi}_2(\theta_1, \theta_2^2) \tilde{p}_2(\theta_1 | \theta_2^1), \\
IC(\theta_2^2 \theta_2^1) \quad & a + \min_{\tilde{\phi} \in \tilde{\Phi}} \sum_{\theta_1 \in \Theta_1} \tilde{\phi}_2(\theta_1, \theta_2^2) \tilde{p}_2(\theta_1 | \theta_2^2) \geq a + B + \min_{\tilde{\phi} \in \tilde{\Phi}} \sum_{\theta_1 \in \Theta_1} \tilde{\phi}_2(\theta_1, \theta_2^1) \tilde{p}_2(\theta_1 | \theta_2^2).
\end{aligned}$$

As  $\tilde{p}_2(\cdot | \theta_2^1) = \tilde{p}_2(\cdot | \theta_2^2)$ , summing the two expressions gives  $2a \geq a + B$ , a contradiction.

## 4 Mechanism with ambiguous transfers

In this section, we formalize the mechanism adopted in the motivating example.

**Definition 4.1:** *A mechanism with ambiguous transfers is a triplet  $\mathcal{M} = (M, \tilde{q}, \tilde{\Phi})$ , where  $M = \times_{i \in I} M_i$  is the message space,  $\tilde{q} : M \rightarrow A$  is a message-contingent allocation rule, and  $\tilde{\Phi}$  is a set of message-contingent transfer rules with a generic element  $\tilde{\phi} : M \rightarrow \mathbb{R}^n$ . We call the set  $\tilde{\Phi}$  **ambiguous transfers**.*

The mechanism designer commits to the allocation rule  $\tilde{q}$  and an arbitrary transfer rule  $\tilde{\phi} \in \tilde{\Phi}$ . Before reporting messages, agents are informed of the set of transfers  $\tilde{\Phi}$  and the allocation rule  $\tilde{q}$ , but not  $\tilde{\phi}$ , the secretly chosen transfer rule. After agents report their messages, the mechanism designer reveals  $\tilde{\phi}$ . Then allocations and transfers are made according to  $\tilde{q}$  and  $\tilde{\phi}$ .

In this mechanism, agents face both risk and uncertainty. They merely know the distribution of others' private information, which we interpret as the risk. Their limited knowledge on the exact transfer rule leads to a layer of uncertainty. For each transfer rule, agents compute their expected payoffs based on beliefs generated by the common prior. As agents only know a set of potential transfer rules  $\tilde{\Phi}$ , following the spirit of Gilboa and Schmeidler (1989)'s maxmin expected utility (MEU), we assume that agents make decisions based on the worst-case transfer rule.

A **strategy** of agent  $i$  is a mapping  $\sigma_i : \Theta_i \rightarrow M_i$ . Like most mechanism design works with ambiguity aversion (e.g., Wolitzky (2016), Di Tillio et al. (2017)), we restrict attention to pure strategies. When there is no ambiguity, the restriction is without loss of generality. When there is ambiguity, depending on how the payoff of playing a mixed strategy is formalized, the restriction could be with or without loss of generality.<sup>6</sup> An **equilibrium** of the mechanism  $\mathcal{M} = (M, \tilde{q}, \tilde{\Phi})$  is a strategy profile  $\sigma = (\sigma_i)_{i \in I}$  such that

---

<sup>6</sup>See Wolitzky (2016) for more details.

$$\begin{aligned} & \inf_{\tilde{\phi} \in \tilde{\Phi}} \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(\tilde{q}(\sigma(\theta_i, \theta_{-i})), (\theta_i, \theta_{-i})) + \tilde{\phi}_i(\sigma(\theta_i, \theta_{-i}))] p_i(\theta_{-i} | \theta_i) \\ & \geq \inf_{\tilde{\phi} \in \tilde{\Phi}} \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(\tilde{q}(\sigma'_i(\theta_i), \sigma_{-i}(\theta_{-i})), (\theta_i, \theta_{-i})) + \tilde{\phi}_i(\sigma'_i(\theta_i), \sigma_{-i}(\theta_{-i}))] p_i(\theta_{-i} | \theta_i) \end{aligned}$$

for all  $i \in I$ ,  $\theta_i \in \Theta_i$ , and  $\sigma'_i : \Theta_i \rightarrow M_i$ .

This paper studies two related but different objectives. One is full surplus extraction by a revenue maximizing mechanism designer, and the other is implementation of an efficient allocation rule via an interim individually rational and ex-post budget-balanced mechanism.

A mechanism with ambiguous transfers  $\mathcal{M} = (M, \tilde{q}, \tilde{\Phi})$  **extracts full surplus** if there exists an equilibrium  $\sigma$  such that

$$-\sum_{\theta \in \Theta} \sum_{i \in I} \tilde{\phi}_i(\sigma(\theta)) p(\theta) = \max_{\hat{q}: \Theta \rightarrow A} \sum_{\theta \in \Theta} \sum_{i \in I} u_i(\hat{q}(\theta), \theta) p(\theta), \forall \tilde{\phi} \in \tilde{\Phi}. \quad (1)$$

The requirement that every transfer rule achieves the same ex-ante revenue follows from Di Tillio et al. (2017)'s Consistency condition, i.e., any transfer rule with a lower expected revenue is non-credible to buyers and thus should not be included in  $\tilde{\Phi}$ . To maximize total surplus, the mechanism designer chooses an ex-post efficient allocation rule  $\hat{q}$ .

A mechanism with ambiguous transfers  $\mathcal{M} = (M, \tilde{q}, \tilde{\Phi})$  (partially) **implements** the efficient allocation rule  $q$ , if there exists an equilibrium  $\sigma$  such that  $\tilde{q}(\sigma(\theta)) = q(\theta)$  for all  $\theta \in \Theta$ .

If for each agent  $i \in I$ , we have  $M_i = \Theta_i$ , i.e.,  $M = \Theta$ , then  $\mathcal{M}$  is said to be a **direct mechanism**. We omit the message space  $\Theta$  in direct mechanisms. A direct mechanism  $(q, \Phi)$  satisfies interim **incentive compatibility** if  $\inf_{\phi \in \Phi} \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\theta_i, \theta_{-i}), (\theta_i, \theta_{-i})) + \phi_i(\theta_i, \theta_{-i})] p_i(\theta_{-i} | \theta_i) \geq \inf_{\phi \in \Phi} \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\theta'_i, \theta_{-i}), (\theta_i, \theta_{-i})) + \phi_i(\theta'_i, \theta_{-i})] p_i(\theta_{-i} | \theta_i)$  for all  $i \in I$ ,  $\theta_i, \theta'_i \in \Theta_i$ . Lemma 4.1 (on revelation principle) implies that it is without loss of generality to focus on incentive compatible direct mechanisms.

**Lemma 4.1:** *Full surplus extraction can be achieved by a mechanism with ambiguous transfers if and only if there is an incentive compatible direct mechanism with ambiguous transfers  $(q, \Phi)$  that extracts the full surplus. An allocation rule  $q$  is implementable via a mechanism with ambiguous transfers if and only if there exists a set of ambiguous transfers  $\Phi$  such that the direct mechanism with ambiguous transfers  $(q, \Phi)$  is incentive compatible.*

Throughout this paper, the outside option  $x_0$  is normalized to give all agents zero payoffs at all type profiles. The direct mechanism with ambiguous transfers  $(q, \Phi)$  satisfies interim **individual rationality** if for all  $i \in I$  and  $\theta_i \in \Theta_i$ ,  $\inf_{\phi \in \Phi} \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\theta_i, \theta_{-i}), (\theta_i, \theta_{-i})) + \phi_i(\theta_i, \theta_{-i})] p_i(\theta_{-i} | \theta_i) \geq 0$ .

$\phi_i(\theta_i, \theta_{-i})]p_i(\theta_{-i}|\theta_i) \geq 0$ . For both full surplus extraction and implementation, we require that the mechanism is interim individually rational so that agents participate voluntarily.

The direct mechanism with ambiguous transfers  $(q, \Phi)$  satisfies ex-post **budget balance** if for all  $\phi \in \Phi$  and  $\theta \in \Theta$ ,  $\sum_{i \in I} \phi_i(\theta) = 0$ . To implement an efficient allocation rule  $q$ , we also require the mechanism is ex-post budget-balanced so that outside money is not needed to finance the efficient outcome. Budget balance is not required for the problem of full surplus extraction, because the mechanism designer collects the full surplus.

## 5 Necessary and sufficient condition

Our necessary and sufficient condition, the Beliefs Determine Preferences property, is introduced by Neeman (2004). It requires that an agent with different types should have distinct beliefs.

**Definition 5.1:** *The **Beliefs Determine Preferences (BDP)** property holds for agent  $i$  if there does not exist  $\bar{\theta}_i, \hat{\theta}_i \in \Theta_i$  with  $\bar{\theta}_i \neq \hat{\theta}_i$  such that*

$$p_i(\theta_{-i}|\bar{\theta}_i) = p_i(\theta_{-i}|\hat{\theta}_i), \forall \theta_{-i} \in \Theta_{-i}.$$

The following subsections present the necessary and sufficient condition for full surplus extraction and for implementation of an efficient allocation under any utility functions. The BDP property plays the key role in both results.

### 5.1 Full surplus extraction

**Theorem 5.1:** *Given a common prior  $p$ , full surplus extraction can be achieved via an interim individually rational mechanism with ambiguous transfers under any profile of utility functions if and only if the BDP property holds for all agents.*

In the Appendix, the proof starts with converting the original problem into finding incentive compatible ambiguous transfers such that every interim individual rationality constraint binds.

The necessity part is proved through constructing utility functions such that full surplus extraction cannot be achieved when the BDP property fails for some agent.

We prove the sufficiency part by constructing a mechanism consisting of two transfer rules. Although there are mechanisms with more transfers that extract the full surplus, to be consistent with the spirit of minimal mechanisms as in Di Tillio et al. (2017), we only

present the one with two rules. The construction is decomposed into several lemmas, which are useful for both full surplus extraction and implementation. Lemma A.1 shows that for each  $i \in I$  and  $\bar{\theta}_i, \hat{\theta}_i \in \Theta_i$  satisfying  $\bar{\theta}_i \neq \hat{\theta}_i$ , there exists a budget-balanced transfer rule  $\psi^{\bar{\theta}_i, \hat{\theta}_i}$  with zero expected values to all truthfully reporting agents, such that  $i$  achieves a negative expected value when she lies from  $\bar{\theta}_i$  to  $\hat{\theta}_i$ . This step is proven via Fredholm's theorem of the alternative. As  $\psi^{\bar{\theta}_i, \hat{\theta}_i}$  only needs to satisfy one incentive compatibility constraint, its existence is guaranteed by the BDP property instead of the stronger Convex Independence condition. Lemmas A.2 and A.3 construct a linear combination of transfer rules  $(\psi^{\bar{\theta}_i, \hat{\theta}_i})_{i \in I, \bar{\theta}_i, \hat{\theta}_i \in \Theta_i, \bar{\theta}_i \neq \hat{\theta}_i}$ , denoted by  $\psi$ , such that  $\psi$  is ex-post budget-balanced, gives all truth-telling agents zero expected values, and gives all misreporting agents non-zero ones. Pick an ex-post efficient allocation rule  $q$  and let  $\eta_i(\theta) = -u_i(q(\theta), \theta)$  for all  $i \in I$  and  $\theta \in \Theta$ . Let the set of ambiguous transfers for agent  $i$  be  $\Phi_i = \{\eta_i + c\psi_i, \eta_i - c\psi_i\}$ . As  $\eta_i$  transfers agent  $i$ 's entire surplus to the mechanism designer and  $\psi_i$  has zero expected value, every interim individual rationality constraint binds. As  $\psi_i$  has non-zero expected value whenever  $i$  misreports, with a sufficiently large  $c$ , the worse expected utility derived from  $\eta_i + c\psi_i$  and  $\eta_i - c\psi_i$  is negative. Thus, incentive compatibility can be achieved.

We remark that in the construction of ambiguous transfers, budget balance of  $\psi$  is not necessary for full surplus extraction. However, requiring budget balance of  $\psi$  allows us to use the same lemmas to study both full surplus extraction and implementation. In addition, we achieve ex-post full surplus extraction. Namely, if the mechanism designer wishes to equate the ex-post revenue and ex-post total surplus, our method still works.

When  $N \geq 2$  and  $|\Theta_i| \geq 2$  for all  $i$ , our necessary and sufficient condition holds for almost every common prior  $p \in \Delta(\Theta)$ .<sup>7</sup>

The necessary and sufficient condition for full surplus extraction under Bayesian mechanism, the Convex Independence condition, requires that for every agent  $i$  and type  $\theta_i$ ,  $p_i(\cdot|\theta_i)$  is not in the convex hull of  $(p_i(\cdot|\theta'_i))_{\theta'_i \neq \theta_i}$ . It holds generically when  $N \geq 2$  and  $2 \leq |\Theta_i| \leq |\Theta_{-i}|$  for all  $i$ . The BDP property is weaker than Convex Independence in two aspects. Firstly, the BDP property can address some linear cases of correlation that are ruled out by Convex Independence, and secondly, the BDP property holds generically even if one agent has too many types compared to others, but in this case Convex Independence fails for sure. When the BDP property holds for all agents but the Convex Independence fails for someone, mechanisms with ambiguous transfers perform strictly better than Bayesian mechanisms in full surplus extraction.

---

<sup>7</sup>If agents without private information are included in  $I$  (see Appendix A.2), the BDP property holds generically for all agents if there exists  $i, j \in I$  with  $i \neq j$  such that  $|\Theta_i|, |\Theta_j| \geq 2$ .

## 5.2 Implementation

**Theorem 5.2:** *Given a common prior  $p$ , an ex-post efficient allocation rule  $q$  is implementable via an interim individually rational and ex-post budget-balanced mechanism with ambiguous transfers under any profile of utility functions if and only if the BDP property holds for all agents.*

When the BDP property fails, we construct utility functions such that an efficient allocation rule is not implementable. We thus prove the necessity part of this theorem. For the sufficiency part, recall that we constructed budget-balanced transfer rule  $\psi$  that gives all truth-telling agents zero expected values and all misreporting agents non-zero ones. Pick any ex-post budget-balanced and interim individually rational transfer rule  $\eta$ . Let the set of ambiguous transfers be  $\Phi = \{\eta + c\psi, \eta - c\psi\}$ . Incentive compatibility can be achieved by choosing a sufficiently large  $c$ .

We remark that efficiency of  $q$  does not play a role in the proof. Actually, by combining our proof with that of Kosenok and Severinov (2008), Theorem 5.2 can be extended to implement any ex-ante **socially rational** allocation rule  $q$ , i.e.,  $q$  satisfying  $\sum_{\theta \in \Theta} \sum_{i \in I} u_i(q(\theta), \theta)p(\theta) \geq 0$  rather than just efficient ones.

Kosenok and Severinov (2008) prove that the conditions of Convex Independence and Identifiability are necessary and sufficient for implementing all efficient or all ex-ante socially rational allocation rules via interim individually rational and ex-post budget-balanced Bayesian mechanisms. The Identifiability condition is generic when  $N = 3$  and there exists  $i \in I$  such that  $|\Theta_i| \geq 3$  or  $N > 3$ . In a budget balanced Bayesian mechanism without Identifiability condition, some agent  $i$  may have the incentive to misreport in a way that makes the truthful report of some  $j \neq i$  appear untruthful because by budget balance  $i$  can benefit from  $j$ 's negative expected transfer. However, when the set of ambiguous transfers  $\Phi$  is used,  $i$  does not have such an incentive, because  $i$  is ambiguous about whether misreport of  $j$  would result in a positive expected transfer to  $j$  or a negative expected transfer. Hence, with ambiguous transfers, we can relax the Identifiability condition.

As the BDP property is weaker than the Convex Independence condition, our ambiguous transfers require a weaker condition than Bayesian mechanisms. The difference between our condition and that of Kosenok and Severinov (2008) characterizes when ambiguous transfers perform strictly better than Bayesian mechanisms in implementation of all efficient or ex-ante socially rational allocation rules. In particular, as Convex Independence and Identifiability never hold simultaneously in two-agent settings but the BDP property holds generically, ambiguous transfers provide a solution for the generic impossibility of two-agent

individually rational, budget-balanced and efficient mechanism design.

## 6 Extension

### 6.1 Implementation under private value environments

When proving the necessity part of Theorem 5.2, we construct a profile of interdependent value utility functions. Some may wonder if the BDP property is necessary for implementation under private value environments. We will show at least  $N - 1$  agents satisfying the BDP property is necessary and sufficient for ex-post efficient, interim individually rational, and ex-post budget-balanced implementation under all private value utility functions. We will also demonstrate that the condition is strictly weaker than the one needed for Bayesian implementation under private value environments.

A utility function  $u_i$  is said to have **private value** if  $u_i(a, (\theta_i, \theta_{-i})) = u_i(a, (\theta_i, \theta'_{-i}))$  for all  $\theta_i \in \Theta_i$ ,  $\theta_{-i}, \theta'_{-i} \in \Theta_{-i}$ , and  $a \in A$ . We denote  $u_i(a, (\theta_i, \theta_{-i}))$  by  $u_i(a, \theta_i)$  in this case.

**Theorem 6.1:** *Given a common prior  $p$ , an ex-post efficient allocation rule  $q$  is implementable via an interim individually rational and ex-post budget-balanced mechanism with ambiguous transfers under any profile of private value utility functions if and only if the BDP property holds for at least  $N - 1$  agents.*

We prove the necessity part by construction again, but the utility functions have private values. For the sufficiency part, we first construct transfers such that  $N - 1$  agents are incentive compatible. Then by allocating all the surplus to the remaining agent and aligning her incentives with the mechanism designer, the agent will also report truthfully in the private value environment, i.e., when all agents have private values.

Recall in Theorem 5.2, efficiency of the allocation rule  $q$  does not play any role, and thus one can implement inefficient but ex-ante socially rational allocation rules if all agents satisfy the BDP property. However, when only  $N - 1$  agents satisfy the BDP property, efficiency of  $q$  plays a role in this proof. This is because we let the agent whose BDP property fails be a budget breaker. Example A.1 in the Appendix illustrates that an inefficient allocation rule may not be implementable if just  $N - 1$  agents satisfy the BDP property.

To compare ambiguous transfers with Bayesian mechanisms, we present the following necessary condition for Bayesian implementation under private value environments.

**Proposition 6.1:** *Given a common prior  $p$ , if any ex-post efficient allocation rule  $q$  is implementable via an interim individually rational and ex-post budget-balanced Bayesian mech-*

*anism under any profile of private value utility functions, then the Convex Independence condition holds for at least  $N - 1$  agents.*

The necessary and sufficient condition of Theorem 6.1 is strictly weaker than the necessary condition of Proposition 6.1.<sup>8</sup> Hence, ambiguous transfers perform strictly better than Bayesian mechanisms in implementing efficient allocation rules under private value environments.

## 6.2 No common prior

This subsection adopts Aumann (1976)'s agreeing to disagree framework to study ambiguous transfers. Namely, we relax the assumption that beliefs are generated by a common prior but still assume common knowledge of their structure.<sup>9</sup> We provide sufficient conditions under which efficient allocations are implementable via an interim individually rational and ex-post budget-balanced mechanism with ambiguous transfers. We also demonstrate with examples that ambiguous transfers can implement Bayesian non-implementable allocations.

In Bayesian mechanism design literature, Bergemann et al. (2012), Smith (2010), and Börgers et al. (2015) have documented results related to ex-post efficiency maximization under agreeing to disagree framework. Without requiring individual rationality and budget balance, Bergemann et al. (2012) show that the BDP property is sufficient for Bayesian implementation of efficient allocations, but the current paper requires interim individual rationality and ex-post budget balance. Smith (2010) compares the welfare of two different mechanisms on public good provision. Börgers et al. (2015) provides a sufficient condition on when agents' interim payoffs can be arbitrarily increased, given there is an incentive compatible mechanism. Different from Smith (2010) and Börgers et al. (2015), the current paper provides a general condition on when the first-best efficiency is implementable.

In this subsection,  $p_i(\cdot|\theta_i)$  still represents the belief of type- $\theta_i$  agent  $i$ , although the beliefs are not generated by a common prior, i.e., there does not exist  $p \in \Delta(\Theta)$  with  $p(\theta_i) > 0$  for all  $\theta_i \in \Theta_i$  such that every  $p_i(\cdot|\theta_i)$  is obtained by Bayesian updating  $p$ .

Without a common prior between the mechanism designer and all agents, full surplus extraction can still be guaranteed via ambiguous transfers when the BDP property holds for all agents. However, full surplus extraction is no longer equivalent to revenue maximization. By utilizing the lack of common prior between the mechanism designer and agents, the

---

<sup>8</sup>The necessary condition of Proposition 6.1 is not sufficient for Bayesian implementation under private value environments. By strengthening it with the Identifiability condition, we can adapt the argument of Kosenok and Severinov (2008) to give a sufficiency result under private value environments.

<sup>9</sup>See Morris (1995) for a review of the justifications of modeling with and without a common prior.

mechanism designer can arbitrarily increase ex-ante revenue. Therefore, we do not study this problem in this section.

Common prior plays an important role when we prove the Theorems 5.2 and 6.1. In Example 6.1, the BDP property holds for all agents, but without a common prior, an efficient allocation rule is not implementable via ambiguous transfers.

**Example 6.1:** Consider a two-agent two-type model with beliefs  $p_1(\theta_2^1|\theta_1^1) = 0.3$ ,  $p_1(\theta_2^1|\theta_1^2) = 0.2$ ,  $p_2(\theta_1^1|\theta_2^1) = 0.3$ , and  $p_2(\theta_1^1|\theta_2^2) = 0.25$ , which are not generated by a common prior. The BDP property is satisfied by both agents. The feasible set of alternatives, the payoffs, and the efficient allocation rule are identical to the motivating example on implementation, except that (1) we no longer have types  $\theta_1^3$  and  $\theta_2^3$  and (2) we assume  $0 < 145a < 21B$ .

Suppose by way of contradiction that there is an interim individually rational and ex-post budget-balanced mechanism with ambiguous transfers  $\mathcal{M} = (q, \Phi)$  implementing the efficient allocation rule  $q$ . As  $IC(\theta_2^2\theta_2^1)$  holds, for any  $\epsilon$ , there exists an interim individually rational transfer rule  $\phi \in \Phi$  (interpreted as the payment from agent 1 to 2) such that:

$$a + 0.25\phi(\theta_1^1, \theta_2^2) + 0.75\phi(\theta_1^2, \theta_2^2) + \epsilon \geq a + B + 0.25\phi(\theta_1^1, \theta_2^1) + 0.75\phi(\theta_1^2, \theta_2^1).$$

Multiply  $IR(\theta_1^1), IR(\theta_1^2), IR(\theta_2^1), IR(\theta_2^2)$ , and  $IC(\theta_2^2\theta_2^1)$  by 40, 105, 75, 70, and 42 respectively, add them up, and let  $\epsilon$  go to zero, we obtain that  $290a \geq 42B$ , a contradiction. Therefore,  $q$  is not implementable via ambiguous transfers.

Despite the failure of Theorem 5.2 without a common prior, a sufficient condition on when efficient allocations are implementable is still feasible. We start with replacing Assumption 2.1 with the following one throughout this subsection because without a common prior, the notation  $p(\theta_i, \theta_j)$  is not well defined.

**Assumption 6.1:** For each  $i, j \in I$ ,  $i \neq j$ , and  $(\theta_i, \theta_j) \in \Theta_i \times \Theta_j$ , assume  $p_i(\theta_j|\theta_i) > 0$ .

Below we introduce a condition called the No Common Prior\* property, which strengthens the assumption that agents' beliefs are not generated by a common prior. For all  $i \neq j$ ,  $\theta_i$ , and  $\theta_j$ , we abuse notation by let  $p_j(\theta_i, \cdot|\theta_j)$  be the vector  $(p_j(\theta_i, \theta_{-i-j}|\theta_j))_{\theta_{-i-j} \in \Theta_{-i-j}}$  when  $N \geq 3$ , and be  $p_j(\theta_i|\theta_j)$  when  $N = 2$ .

**Definition 6.1:** Agent  $i$  satisfies the **No Common Prior\*** (NCP\*) property if there do not exist different types  $\bar{\theta}_i \neq \hat{\theta}_i$  such that

1. there exists  $\mu \in \Delta(\Theta)$  such that  $\mu(\theta_j) > 0$  and  $\mu(\theta_{-j}|\theta_j) = p_j(\theta_{-j}|\theta_j)$  for all  $(j, \theta_j) \neq (i, \hat{\theta}_i)$ ,

2. there exists  $\bar{C} > 0$  and  $\hat{C} > 1$  such that  $\hat{C}p_i(\theta_j, \cdot|\hat{\theta}_i) = p_i(\theta_j, \cdot|\bar{\theta}_i) + \bar{C}\frac{p_i(\theta_j|\bar{\theta}_i)}{p_j(\bar{\theta}_i|\theta_j)}p_j(\hat{\theta}_i, \cdot|\theta_j)$  for any  $j \neq i$  and  $\theta_j$ .

When there is a common prior over  $\Theta$ , one can show the NCP\* property is equivalent to the BDP property. Without a common prior over  $\Theta$ , the statement of NCP\* property cannot be simplified, but it is very weak. For example, if  $N \geq 3$  and there are agents  $i \neq j$  and types  $\bar{\theta}_i \neq \hat{\theta}_i$ ,  $\bar{\theta}_j \neq \hat{\theta}_j$ , such that the probability distributions over  $\Theta_{-i-j}$  satisfy  $p_i(\cdot|\bar{\theta}_i\bar{\theta}_j) \neq p_j(\cdot|\bar{\theta}_i\bar{\theta}_j)$  and  $p_i(\cdot|\hat{\theta}_i\hat{\theta}_j) \neq p_j(\cdot|\hat{\theta}_i\hat{\theta}_j)$ , then the NCP\* property holds for all  $i \in I$ .

In Example 6.1, the NCP\* property fails for agent 1, as  $(i, \bar{\theta}_i, \hat{\theta}_i) = (1, \theta_1^2, \theta_1^1)$  satisfies the two conditions in the definition. The first condition holds by letting  $\mu(\theta_1^1, \theta_2^1) = \frac{9}{142}$ ,  $\mu(\theta_1^1, \theta_2^2) = \frac{28}{142}$ ,  $\mu(\theta_1^2, \theta_2^1) = \frac{21}{142}$ , and  $\mu(\theta_1^2, \theta_2^2) = \frac{84}{142}$ . The second condition can be written as  $\hat{C}\frac{p_i(\theta_j|\hat{\theta}_i)}{p_i(\theta_j|\bar{\theta}_i)} = 1 + \bar{C}\frac{p_j(\hat{\theta}_i|\theta_j)}{p_j(\bar{\theta}_i|\theta_j)}$ . By setting  $\hat{C} = \frac{25}{14}$  and  $\bar{C} = \frac{8}{3}$ , we see that  $\frac{25}{14}(\frac{0.3}{0.25}, \frac{0.7}{0.75}) = (1, 1) + \frac{8}{3}(\frac{0.3}{0.7}, \frac{0.2}{0.8})$ .

In the following theorem, we provide a sufficient condition for implementation via ambiguous transfers when there is no common prior.

**Theorem 6.2:** *Given beliefs  $(p_i(\cdot|\theta_i))_{i \in I, \theta_i \in \Theta_i}$  that are not generated by a common prior, if the BDP and NCP\* properties hold for all agents, then an ex-post efficient allocation rule  $q$  is implementable via an interim individually rational and ex-post budget-balanced mechanism with ambiguous transfers under any profile of utility functions.*

Similar to Theorem 5.2, efficiency of  $q$  does not play a role in this proof. We can actually implement any  $q$  such that  $\sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\theta_i, \theta_{-i}), (\theta_i, \theta_{-i})) + \eta_i(\theta_i, \theta_{-i})]p_i(\theta_{-i}|\theta_i) \geq 0$  for all  $i \in I$ ,  $\theta_i \in \Theta_i$  and some ex-post budget-balanced  $\eta$ .<sup>10</sup>

The example below shows that there are cases when ambiguous transfers perform better than Bayesian mechanisms.

**Example 6.2:** *In this example without a common, the efficient allocation rule  $q$  is not Bayesian implementable, but is implementable via ambiguous transfers.*

*Consider the following beliefs that are not be generated by a common prior:*

$p_1(\tilde{\theta}_2 \tilde{\theta}_1)$	$\theta_2^1$	$\theta_2^2$
$\theta_1^1$	$\frac{1}{3}$	$\frac{2}{3}$
$\theta_1^2$	$\frac{1}{2}$	$\frac{1}{2}$
$\theta_1^3$	$\frac{2}{3}$	$\frac{1}{3}$

$p_2(\tilde{\theta}_1 \tilde{\theta}_2)$	$\theta_2^1$	$\theta_2^2$
$\theta_1^1$	$\frac{7}{28}$	$\frac{13}{28}$
$\theta_1^2$	$\frac{12}{28}$	$\frac{12}{28}$
$\theta_1^3$	$\frac{9}{28}$	$\frac{3}{28}$

<sup>10</sup>With common prior, Kosenok and Severinov (2008) has proved the equivalence between this property and ex-ante social rationality.

The feasible set of alternatives, payoffs, and the efficient allocation rule are identical to those in the motivating example of implementation, except that the type space is smaller here and  $0 < 9a < B$  is imposed. Suppose by way of contradiction that there exists payment from agent 1 to 2, denoted by  $\phi$ , that implements  $q$ . By multiplying  $IR(\theta_1^1)$ ,  $IR(\theta_1^2)$ ,  $IR(\theta_1^3)$ ,  $IC(\theta_1^1\theta_1^2)$ ,  $IC(\theta_1^1\theta_1^3)$ ,  $IC(\theta_1^2\theta_1^1)$ ,  $IC(\theta_1^3\theta_1^2)$ ,  $IR(\theta_2^1)$ ,  $IR(\theta_2^2)$ ,  $IC(\theta_2^1\theta_2^2)$ , and  $IC(\theta_2^2\theta_2^1)$  by 3, 8, 3, 3, 3, 4, 3, 7, 7, 3.5, and 3.5, and summing up, we obtain  $0 \geq 3.5B - 31.5a$ , a contradiction.

To see  $q$  is implementable via ambiguous transfers, by Theorem 6.2 it is sufficient to check both agents satisfy the BDP and NCP\* properties. The BDP property holds clearly. To verify the NCP\* property for agent 1, consider  $(i, \bar{\theta}_i, \hat{\theta}_i) = (1, \theta_1^1, \theta_1^2)$ . The second condition does not hold because there does not exist  $\bar{C} > 0$  and  $\hat{C} > 1$  such that  $\hat{C}(\frac{13}{7}, 1, \frac{1}{3}) = (1, 1, 1) + \bar{C}(2, 1, 0.5)$ . A symmetric argument applies to  $(i, \bar{\theta}_i, \hat{\theta}_i) = (1, \theta_1^2, \theta_1^1)$ . Agent 2 satisfies the NCP\* property because for each pair  $(\bar{\theta}_2, \hat{\theta}_2)$ , the first condition in the NCP\* property fails.

In a private value environment without a common prior, we have the following sufficient condition for implementation of efficient allocations.

**Theorem 6.3:** *Given beliefs  $(p_i(\cdot|\theta_i))_{i \in I, \theta_i \in \Theta_i}$  that are not generated by a common prior, if there do not exist  $i \neq j$  such that the BDP property fails for  $i$  and the NCP\* property fails for  $j$ , then any ex-post efficient allocation rule  $q$  is implementable via an interim individually rational and ex-post budget-balanced mechanism with ambiguous transfers under any profile of private value utility functions.*

The sufficient conditions of Theorem 6.3 is weak. Recall the strengthening of the NCP\* property after Definition 6.1. Corollary 6.1 follows directly from the above theorem.

**Corollary 6.1:** *When  $N \geq 3$  and there are agents  $i \neq j$  and types  $\bar{\theta}_i \neq \hat{\theta}_i$ ,  $\bar{\theta}_j \neq \hat{\theta}_j$ , such that the probability distributions over  $\Theta_{-i-j}$  satisfy  $p_i(\cdot|\bar{\theta}_i\bar{\theta}_j) \neq p_j(\cdot|\bar{\theta}_i\bar{\theta}_j)$  and  $p_i(\cdot|\hat{\theta}_i\hat{\theta}_j) \neq p_j(\cdot|\hat{\theta}_i\hat{\theta}_j)$ , then any ex-post efficient allocation rule  $q$  is implementable via an interim individually rational and ex-post budget-balanced mechanism with ambiguous transfers under any profile of private value utility functions.*

This result may be surprising, because the sufficient condition is merely a slight strengthening of the fact that there is no common prior.

Corollary 6.2 provides another simple sufficient condition to guarantee implementation of efficient allocations under private value environments.

**Corollary 6.2:** *Given beliefs  $(p_i(\cdot|\theta_i))_{i \in I, \theta_i \in \Theta_i}$  that are not generated by a common prior, if the BDP property holds for all agents, then any ex-post efficient allocation rule  $q$  is imple-*

mentable via an interim individually rational and ex-post budget-balanced mechanism with ambiguous transfers under any profile of private value utility functions.

This corollary is based on Lemma A.4, which says if the BDP property holds for all agents, then the NCP\* property holds for at least N-1 agents. The details are in the Appendix.

The example below shows that there are cases when ambiguous transfers perform better than Bayesian mechanisms.

**Example 6.3:** *In this example of bilateral trading, the efficient allocation rule  $q$  is not Bayesian implementable, but it is implementable via ambiguous transfers.*

*Agent 1 is the buyer, and 2 is the seller. Outcomes in  $A = \{x_0, x_1\}$  are feasible, where  $x_0$  represents no trade. The payoffs of  $x_1$ , trading, for both agents are given below.*

$x_1$	$\theta_2^1$	$\theta_2^2$
$\theta_1^1$	4, -3.5	4, -0.5
$\theta_1^2$	1, -3.5	1, -0.5

*The efficient allocation rule satisfies  $q(\theta_1^2, \theta_2^1) = x_0$  and  $q(\theta) = x_1$  for all other  $\theta$ . We adopt the same beliefs as Example 6.1, and thus there is no common prior.*

*To see  $q$  is not Bayesian implementable, suppose by way of contradiction that there exists an interim individually rational and ex-post budget-balanced Bayesian mechanism that implements  $q$ . Let  $\phi$  denote the payment from agent 1 to 2. Multiply  $IC(\theta_1^1\theta_2^2)$ ,  $IR(\theta_1^2)$ ,  $IC(\theta_1^2\theta_2^1)$ ,  $IR(\theta_2^1)$  and  $IC(\theta_2^2\theta_2^1)$  by 4, 10, 1, 10, and 8 respectively, and then add them up. We obtain  $-0.9 \geq 0$ , which is a contradiction. Therefore,  $q$  is not Bayesian implementable.*

*However, from Corollary 6.2, we know  $q$  is implementable via ambiguous transfers.*

### 6.3 Other ambiguity aversion preferences

To check the robustness of our result, we look at alternative preferences of ambiguity aversion in this subsection. One is the  $\alpha$ -maxmin expected utility ( $\alpha$ -MEU) as in Ghirardato and Marinacci (2002), and the other is the smooth ambiguity aversion preferences of Klibanoff et al. (2005). This section shows that the mechanism designer can benefit from generating ambiguity in transfers, even if agents' preferences differ from Gilboa and Schmeidler (1989).

Ghirardato and Marinacci (2002) introduce the  $\alpha$ -MEU, which is a generalization of the MEU. Under an environment described in Section 2, a type- $\theta_i$  agent  $i$  with  **$\alpha$ -maxmin expected utility** has the following interim utility level from participating and reporting truthfully when  $\Phi$  is the set of ambiguous transfers:

$$\sum_{\theta_{-i} \in \Theta_{-i}} u_i(q(\theta_i, \theta_{-i}), (\theta_i, \theta_{-i})) p_i(\theta_{-i} | \theta_i) + \alpha \inf_{\phi \in \Phi} \sum_{\theta_{-i} \in \Theta_{-i}} \phi_i(\theta_i, \theta_{-i}) p_i(\theta_{-i} | \theta_i) \\ + (1 - \alpha) \sup_{\phi \in \Phi} \sum_{\theta_{-i} \in \Theta_{-i}} \phi_i(\theta_i, \theta_{-i}) p_i(\theta_{-i} | \theta_i),$$

where  $\alpha \in [0, 1]$ . An agent is said to be ambiguity-averse if  $\alpha > 0.5$ . All previous sections adopt the MEU preferences, which correspond to the case  $\alpha = 1$ .

Under the  $\alpha$ -MEU preferences with  $\alpha > 0.5$ , Theorem 5.2, as well as the sufficiency part of Theorems 5.1 and 6.1, still holds. The construction of ambiguous transfers are the same except for choosing a different multiplier  $c$ .

An agent  $i$  with **smooth ambiguity aversion** has a utility function of

$$\int_{\pi \in \Delta(\Phi)} v \left( \int_{\phi \in \Phi} \left( \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\theta_i, \theta_{-i}), (\theta_i, \theta_{-i})) + \phi_i(\theta_i, \theta_{-i})] p_i(\theta_{-i} | \theta_i) \right) d\pi \right) d\mu,$$

where

- for each distribution  $\pi \in \Delta(\Phi)$ ,  $\pi(\phi)$  measures the subjective density that  $\phi$  is the true transfer rule chosen by the mechanism designer;
- for each distribution  $\mu \in \Delta(\Delta(\Phi))$ ,  $\mu(\pi)$  measures the subjective density that  $\pi \in \Delta(\Phi)$  is the right density function the mechanism designer uses to choose the transfer rule;
- $v : R \rightarrow R$  is a strictly increasing function that characterizes ambiguity attitude, where a strictly concave  $v$  implies ambiguity aversion.

To see ambiguous transfers help under smooth ambiguity aversion preferences, we demonstrate with the motivating example on implementation. Let  $v$  be a strictly increasing and strictly concave function. Consider the same transfers as  $\phi^1$  and  $\phi^2$  except for a potentially different multiplier  $c$ . Then it is easy to verify individual rationality and budget balance. A generic element of  $\Delta(\Phi)$  is a Bernoulli distribution between  $\phi^1$  and  $\phi^2$ . Let  $\mu$  be the uniform distribution over  $\Delta(\Phi)$  for example. As an illustration, we check  $IC(\theta_2^2 \theta_2^1)$ . Truth-telling always gives agent 2 an expected utility of

$$\int_0^1 v(\mu a + (1 - \mu)a) d\mu = v(a).$$

By misreporting from  $\theta_2^2$  to  $\theta_2^1$ , agent 2 gets an interim utility of

$$\int_0^1 v(\mu(a + B + \frac{3}{4}c) + (1 - \mu)(a + B - \frac{3}{4}c)) d\mu.$$

For  $v$  sufficiently concave or  $c$  sufficiently large, the above expression has a value no more than  $v(a)$ , implying that truth-telling is incentive compatible. One can verify other incentive compatibility constraints as well.

## 7 Conclusion

This paper introduces ambiguous transfers to study full surplus extraction and implementation of an efficient allocation rule via an individually rational and budget-balanced mechanism. We show that the BDP property is necessary and sufficient for both problems, which is weaker than the necessary and sufficient condition for full surplus extraction and implementation via Bayesian mechanisms. Hence, ambiguous transfers can go beyond Bayesian mechanisms. The BDP property holds generically when there are at least two agents. In particular, under two-agent settings, the BDP property offers a solution to overcome the negative results on bilateral trading problems generically.

## A Appendix

### A.1 Proofs and examples

**Proof of Lemma 4.1.** It is sufficient to prove the “only if” direction.

Suppose a mechanism with ambiguous transfers  $\mathcal{M} = (M, \tilde{q}, \tilde{\Phi})$  extracts surplus, then there exists an equilibrium  $\sigma$  such that

$$-\sum_{\theta \in \Theta} \sum_{i \in I} \tilde{\phi}_i(\sigma(\theta))p(\theta) = \max_{\hat{q}: \Theta \rightarrow A} \sum_{\theta \in \Theta} \sum_{i \in I} u_i(\hat{q}(\theta), \theta)p(\theta), \forall \tilde{\phi} \in \tilde{\Phi}.$$

Define  $q(\theta) = \tilde{q}(\sigma(\theta))$  for all  $\theta \in \Theta$ . For each  $\tilde{\phi} \in \tilde{\Phi}$ , define  $\phi : \Theta \rightarrow \mathbb{R}^n$  by  $\phi = \tilde{\phi} \circ \sigma$ , and denote the collection of all  $\phi$  by  $\Phi$ .

Suppose a mechanism with ambiguous transfers  $\mathcal{M} = (M, \tilde{q}, \tilde{\Phi})$  implements  $q$ . Then there exists an equilibrium  $\sigma$  such that  $\tilde{q}(\sigma(\theta)) = q(\theta)$  for all  $\theta \in \Theta$ . For each  $\tilde{\phi} \in \tilde{\Phi}$ , define  $\phi : \Theta \rightarrow \mathbb{R}^n$  by  $\phi = \tilde{\phi} \circ \sigma$ , and denote the collection of all  $\phi$  by  $\Phi$ .

For both cases, we prove that the direct mechanism with ambiguous transfers  $\mathcal{M}' = (q, \Phi)$  is incentive compatible. To see this, for all  $i \in I$ ,  $\theta_i, \theta'_i \in \Theta_i$ ,

$$\begin{aligned} & \inf_{\phi \in \Phi} \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\theta'_i, \theta_{-i}), (\theta_i, \theta_{-i})) + \phi_i(\theta'_i, \theta_{-i})]p_i(\theta_{-i}|\theta_i) \\ &= \inf_{\tilde{\phi} \in \tilde{\Phi}} \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(\tilde{q}(\sigma(\theta'_i, \theta_{-i})), (\theta_i, \theta_{-i})) + \tilde{\phi}_i(\sigma(\theta'_i, \theta_{-i}))]p_i(\theta_{-i}|\theta_i) \\ &\leq \inf_{\tilde{\phi} \in \tilde{\Phi}} \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(\tilde{q}(\sigma(\theta_i, \theta_{-i})), (\theta_i, \theta_{-i})) + \tilde{\phi}_i(\sigma(\theta_i, \theta_{-i}))]p_i(\theta_{-i}|\theta_i) \\ &= \inf_{\phi \in \Phi} \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\theta_i, \theta_{-i}), (\theta_i, \theta_{-i})) + \phi_i(\theta_i, \theta_{-i})]p_i(\theta_{-i}|\theta_i), \end{aligned}$$

where the inequality comes from the fact that  $\sigma_i(\theta'_i) \in M_i$  can be viewed as a message sent by  $i$  under a constant strategy. Therefore, truth-telling is an equilibrium of  $\mathcal{M}'$ .  $\square$

We present three lemmas before proving Theorems 5.1 and 5.2. With the BDP property, Lemma A.1 implies that for each incentive compatibility constraint, there exists an individually rational and budget-balanced transfer rule satisfying this constraint. Lemmas A.2 and A.3 establish the existence of an individually rational and budget-balanced transfer rule that gives all truth-telling agents zero expected values and all misreporting agents non-zero ones.

**Lemma A.1:** *The BDP property holds for agent  $i$  if and only if for all  $\bar{\theta}_i, \hat{\theta}_i \in \Theta_i$  with  $\bar{\theta}_i \neq \hat{\theta}_i$ , there exists  $\psi^{\bar{\theta}_i, \hat{\theta}_i} : \Theta \rightarrow \mathbb{R}^n$  such that,*

1.  $\sum_{j \in I} \psi_j^{\bar{\theta}_i, \hat{\theta}_i}(\theta) = 0$  for all  $\theta \in \Theta$ ;
2.  $\sum_{\theta_{-j} \in \Theta_{-j}} \psi_j^{\bar{\theta}_i, \hat{\theta}_i}(\theta_j, \theta_{-j}) p_j(\theta_{-j} | \theta_j) = 0$  for all  $j \in I$  and  $\theta_j \in \Theta_j$ ;
3.  $\sum_{\theta_{-i} \in \Theta_{-i}} \psi_i^{\bar{\theta}_i, \hat{\theta}_i}(\hat{\theta}_i, \theta_{-i}) p_i(\theta_{-i} | \bar{\theta}_i) < 0$ .

*Proof.* We start with defining vectors  $e_\theta$  for all  $\theta \in \Theta$  and  $p_{\theta_j, \theta'_j}$  for all  $j \in I, \theta_j, \theta'_j \in \Theta_j$ . Each of the vectors has  $N \times |\Theta|$  dimensions, and each dimension corresponds to an agent and a type profile. For each  $\theta \in \Theta$ , let all elements of  $e_\theta$  that correspond to the type profile  $\theta$  be 1 and everywhere else be 0. For each  $j \in I$  and  $\theta_j, \theta'_j \in \Theta_j$ , let elements of  $p_{\theta_j, \theta'_j}$  that correspond to the agent  $j$  and some type profile  $(\theta'_j, \theta_{-j})$  be  $p_j(\theta_{-j} | \theta_j)$  for all  $\theta_{-j} \in \Theta_{-j}$ . Everywhere else of  $p_{\theta_j, \theta'_j}$  is 0.<sup>11</sup>

**Sufficiency.** Suppose by way of contradiction that there exists  $\bar{\theta}_i, \hat{\theta}_i \in \Theta_i$  with  $\bar{\theta}_i \neq \hat{\theta}_i$ , such that no  $\psi^{\bar{\theta}_i, \hat{\theta}_i}$  satisfies the three conditions. By Fredholm's theorem of the alternative, there exist coefficients  $(a_{\theta_j})_{j \in I, \theta_j \in \Theta_j}$  and  $(b_\theta)_{\theta \in \Theta}$  such that

$$p_{\bar{\theta}_i, \hat{\theta}_i} = \sum_{j \in I} \sum_{\theta_j \in \Theta_j} a_{\theta_j} p_{\theta_j, \theta_j} + \sum_{\theta \in \Theta} b_\theta e_\theta. \quad (2)$$

Fix any agent  $j \neq i$ . All elements of  $p_{\bar{\theta}_i, \hat{\theta}_i}$  that correspond to agent  $j$  are zero. All those corresponding to agent  $i$  and  $\bar{\theta}_i$  are zero, too. Those corresponding to agent  $i$  and  $\hat{\theta}_i$  may not be zero. The three observations, along with expression (2), imply that

$$0 = a_{\theta_j} p_j(\theta_i, \theta_{-i-j} | \theta_j) + b_{\theta_i, \theta_j, \theta_{-i-j}}, \forall \theta_i, \theta_j, \theta_{-i-j}, \quad (3)$$

<sup>11</sup>As an illustration, we look at a two-agent example with  $\Theta$  being  $((\theta_1^1, \theta_2^1), (\theta_1^1, \theta_2^2), (\theta_1^2, \theta_2^1), (\theta_1^2, \theta_2^2))$ . For each  $e_\theta$  or  $p_{\theta_j, \theta'_j}$ , any of its first four dimensions corresponds to agent 1 and a type profile. Any of its last four dimensions correspond to agent 2 and a type profile. Then for example,  $e_{(\theta_1^2, \theta_2^2)} = (0, 0, 1, 0, 0, 0, 1, 0)$ , and  $p_{\theta_2^2, \theta_1^1} = (0, 0, 0, 0, p_2(\theta_1^1 | \theta_2^2), 0, p_2(\theta_1^2 | \theta_2^2), 0)$ .

$$0 = a_{\bar{\theta}_i} p_i(\theta_j, \theta_{-i-j} | \bar{\theta}_i) + b_{\bar{\theta}_i, \theta_j, \theta_{-i-j}}, \forall \theta_j, \theta_{-i-j}, \quad (4)$$

$$p_i(\theta_j, \theta_{-i-j} | \bar{\theta}_i) = a_{\hat{\theta}_i} p_i(\theta_j, \theta_{-i-j} | \hat{\theta}_i) + b_{\hat{\theta}_i, \theta_j, \theta_{-i-j}}, \forall \theta_j, \theta_{-i-j}. \quad (5)$$

By choosing  $\theta_i = \bar{\theta}_i$  in expression (3) and cancelling  $b_{\bar{\theta}_i, \theta_j, \theta_{-i-j}}$  in expressions (3) and (4), we have  $a_{\theta_j} p_j(\bar{\theta}_i, \theta_{-i-j} | \theta_j) = a_{\bar{\theta}_i} p_i(\theta_j, \theta_{-i-j} | \bar{\theta}_i)$ . Summing across all  $\theta_{-i-j} \in \Theta_{-i-j}$  when  $N \geq 3$  or ignoring any  $\theta_{-i-j}$  when  $N = 2$  yields  $a_{\theta_j} p_j(\bar{\theta}_i | \theta_j) = a_{\bar{\theta}_i} p_i(\theta_j | \bar{\theta}_i)$ . From Bayes' rule, we further know  $a_{\theta_j} = a_{\bar{\theta}_i} \frac{p_j(\theta_j)}{p_i(\bar{\theta}_i)}$  for all  $\theta_j \in \Theta_j$ .

By choosing  $\theta_i = \hat{\theta}_i$  in expression (3) and plugging in  $a_{\theta_j}$  derived in the previous paragraph, we know  $b_{\hat{\theta}_i, \theta_j, \theta_{-i-j}} = -a_{\bar{\theta}_i} \frac{p_j(\theta_j)}{p_i(\bar{\theta}_i)} p_j(\hat{\theta}_i, \theta_{-i-j} | \theta_j) = -a_{\bar{\theta}_i} \frac{p_i(\hat{\theta}_i)}{p_i(\bar{\theta}_i)} p_i(\theta_j, \theta_{-i-j} | \hat{\theta}_i)$  for all  $\theta_j, \theta_{-i-j}$ .

Plugging  $b_{\hat{\theta}_i, \theta_j, \theta_{-i-j}}$  into expression (5) yields  $p_i(\theta_j, \theta_{-i-j} | \bar{\theta}_i) = (a_{\bar{\theta}_i} - a_{\bar{\theta}_i} \frac{p_i(\hat{\theta}_i)}{p_i(\bar{\theta}_i)}) p_i(\theta_j, \theta_{-i-j} | \hat{\theta}_i)$  for all  $\theta_j, \theta_{-i-j}$ . Hence,  $a_{\hat{\theta}_i} - a_{\bar{\theta}_i} \frac{p_i(\hat{\theta}_i)}{p_i(\bar{\theta}_i)} = 1$  and  $p_i(\cdot | \bar{\theta}_i) = p_i(\cdot | \hat{\theta}_i)$ , a contradiction.

**Necessity.** Suppose the BDP property fails for agent  $i$ , i.e., there exists  $\bar{\theta}_i \neq \hat{\theta}_i$  such that  $p_i(\cdot | \bar{\theta}_i) = p_i(\cdot | \hat{\theta}_i)$ . For all  $\psi^{\bar{\theta}_i, \hat{\theta}_i}$  such that  $\sum_{\theta_{-i} \in \Theta_{-i}} \psi_i^{\bar{\theta}_i, \hat{\theta}_i}(\hat{\theta}_i, \theta_{-i}) p_i(\theta_{-i} | \hat{\theta}_i) = 0$ , we have

$\sum_{\theta_{-i} \in \Theta_{-i}} \psi_i^{\bar{\theta}_i, \hat{\theta}_i}(\hat{\theta}_i, \theta_{-i}) p_i(\theta_{-i} | \bar{\theta}_i) = 0$ . Hence, there does not exist  $\psi^{\bar{\theta}_i, \hat{\theta}_i}$  such that the three conditions stated in the lemma hold.  $\square$

**Lemma A.2:** For any positive integer  $K$  and any matrix  $X_{K \times K}$  whose diagonal elements are all negative, there exists  $\lambda \in \mathbb{R}_+^K \setminus \{\mathbf{0}\}$  such that  $\sum_{\tilde{k}=1}^K x_{k\tilde{k}} \lambda_{\tilde{k}} \neq 0$  for all  $k \in \{1, \dots, K\}$ .

*Proof.* We prove the result by induction.

First, let  $K = 1$ . Pick an arbitrary  $\lambda_1 > 0$ . As  $x_{11} < 0$ , the statement holds for 1.

Suppose the statement holds for  $K-1$ , where  $K \geq 2$ . Now we consider an arbitrary  $X_{K \times K}$  with negative diagonal elements. By the supposition for the northwest  $K-1$  by  $K-1$  block, there exists  $(\lambda_1, \dots, \lambda_{K-1}) \in \mathbb{R}_+^{K-1} \setminus \{\mathbf{0}\}$  such that  $\sum_{\tilde{k}=1}^{K-1} x_{k\tilde{k}} \lambda_{\tilde{k}} \neq 0$  for all  $k \in \{1, \dots, K-1\}$ .

Case 1. Suppose  $\sum_{\tilde{k}=1}^{K-1} x_{K\tilde{k}} \lambda_{\tilde{k}} \neq 0$ . Let  $\lambda_K = 0$ , and thus the statement holds for  $K$ .

Case 2. Suppose  $\sum_{\tilde{k}=1}^{K-1} x_{K\tilde{k}} \lambda_{\tilde{k}} = 0$  and  $x_{Kk_0} \lambda_{k_0} \neq 0$  for some  $k_0 \in \{1, \dots, K-1\}$ . Let

$(\lambda'_1, \dots, \lambda'_{K-1}) = (\lambda_1, \dots, \lambda_{k_0-1}, \lambda_{k_0} + \epsilon, \lambda_{k_0+1}, \dots, \lambda_{K-1})$  for  $\epsilon > 0$ . Then  $\sum_{\tilde{k}=1}^{K-1} x_{K\tilde{k}} \lambda'_{\tilde{k}} \neq 0$ . When

$\epsilon$  is sufficiently close to zero,  $\sum_{\tilde{k}=1}^{K-1} x_{k\tilde{k}} \lambda'_{\tilde{k}} \neq 0$  for all  $k \in \{1, \dots, K-1\}$ . Therefore, we can replace  $(\lambda_1, \dots, \lambda_{K-1})$  with  $(\lambda'_1, \dots, \lambda'_{K-1})$  and go back to Case 1.

Case 3. Suppose  $x_{K\tilde{k}} \lambda_{\tilde{k}} = 0$  for all  $\tilde{k} \in \{1, \dots, K-1\}$ . Let  $\lambda_K > 0$  and  $\lambda_K \neq -\frac{\sum_{\tilde{k}=1}^{K-1} x_{k\tilde{k}} \lambda_{\tilde{k}}}{x_{kK}}$  for all  $k \in \{1, \dots, K-1\}$  with  $x_{kK} \neq 0$ . Then the statement holds for  $K$ .  $\square$

**Lemma A.3:** *If the BDP property holds for all agents, then there exists  $\psi : \Theta \rightarrow \mathbb{R}^n$  such that*

1.  $\sum_{i \in I} \psi_i(\theta) = 0$  for all  $\theta \in \Theta$ ;
2.  $\sum_{\theta_{-i} \in \Theta_{-i}} \psi_i(\theta_i, \theta_{-i}) p_i(\theta_{-i} | \theta_i) = 0$  for all  $i \in I$  and  $\theta_i \in \Theta_i$ ;
3.  $\sum_{\theta_{-i} \in \Theta_{-i}} \psi_i(\hat{\theta}_i, \theta_{-i}) p_i(\theta_{-i} | \bar{\theta}_i) \neq 0$  for all  $i \in I$  and  $\bar{\theta}_i, \hat{\theta}_i \in \Theta_i$  with  $\bar{\theta}_i \neq \hat{\theta}_i$ .

*Proof.* Let  $K$  be the cardinality of  $\mathcal{K} = \{(\bar{\theta}_i, \hat{\theta}_i) | i \in I, \bar{\theta}_i, \hat{\theta}_i \in \Theta_i, \bar{\theta}_i \neq \hat{\theta}_i\}$ . Let  $f : \mathcal{K} \rightarrow \{1, \dots, K\}$  be a one to one mapping, which allows us to index the elements of  $\mathcal{K}$ .

For all  $k, \tilde{k} \in \{1, \dots, K\}$  ( $k, \tilde{k}$  may be equal), where  $f^{-1}(k) = (\bar{\theta}_i, \hat{\theta}_i)$  and  $f^{-1}(\tilde{k}) = (\tilde{\theta}_j, \tilde{\theta}_j)$ , we define  $x_{k\tilde{k}} = \sum_{\theta_{-i} \in \Theta_{-i}} \psi_i^{\tilde{\theta}_j, \tilde{\theta}_j}(\hat{\theta}_i, \theta_{-i}) p_i(\theta_{-i} | \bar{\theta}_i)$ , where each  $\psi_i^{\tilde{\theta}_j, \tilde{\theta}_j}$  is defined and proved to exist

in Lemma A.1. By the third property of  $\psi_i^{\tilde{\theta}_j, \tilde{\theta}_j}$ , we know  $x_{k\tilde{k}} < 0$  for all  $\tilde{k} \in \{1, \dots, K\}$ .

From Lemma A.2, there exists  $\lambda \in \mathbb{R}_+^K \setminus \{\mathbf{0}\}$  such that  $\sum_{\tilde{k}=1}^K x_{k\tilde{k}} \lambda_{\tilde{k}} \neq 0$  for all  $k \in \{1, \dots, K\}$ .

This implies that for all  $(\bar{\theta}_i, \hat{\theta}_i) \in \mathcal{K}$ ,

$$\sum_{\tilde{k}=1}^K \left[ \sum_{\theta_{-i} \in \Theta_{-i}} \psi_i^{f^{-1}(\tilde{k})}(\hat{\theta}_i, \theta_{-i}) p_i(\theta_{-i} | \bar{\theta}_i) \right] \lambda_{\tilde{k}} = \sum_{\theta_{-i} \in \Theta_{-i}} \left[ \sum_{\tilde{k}=1}^K \lambda_{\tilde{k}} \psi_i^{f^{-1}(\tilde{k})}(\hat{\theta}_i, \theta_{-i}) \right] p_i(\theta_{-i} | \bar{\theta}_i) \neq 0.$$

Define  $\psi_i = \sum_{\tilde{k}=1}^K \lambda_{\tilde{k}} \psi_i^{f^{-1}(\tilde{k})}$  for all  $i \in I$ . Then  $\psi$  satisfies the third requirement of this lemma. The other two requirements are trivial because  $\psi$  is a linear combination of transfers satisfying the two equations.  $\square$

**Proof of Theorem 5.1.** We first claim that an interim individually rational mechanism with ambiguous transfers  $(q, \Phi)$  extracts full surplus if and only if  $q$  is ex-post efficient and  $\sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\theta), \theta) + \phi_i(\theta)] p_i(\theta_{-i} | \theta_i) = 0$  for all  $i \in I$ ,  $\theta_i \in \Theta_i$ , and  $\phi_i \in \Phi_i$ . The “if” direction

is clear given expression (1). To see the “only if” direction, suppose  $q$  is inefficient or there exists  $i \in I$ ,  $\theta_i \in \Theta_i$ , and  $\phi \in \Phi$  such that  $\sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\theta), \theta) + \phi_i(\theta)] p_i(\theta_{-i} | \theta_i) > 0$ . By individual rationality, the fact that  $q$  is a common prior, and Assumption 2.1,

$$-\sum_{\theta \in \Theta} \sum_{j \in I} \phi_j(\sigma(\theta)) p(\theta) \leq \sum_{\theta \in \Theta} \sum_{j \in I} u_j(q(\theta), \theta) p(\theta) \leq \max_{\hat{q}: \Theta \rightarrow A} \sum_{\theta \in \Theta} \sum_{j \in I} u_j(\hat{q}(\theta), \theta) p(\theta), \quad (6)$$

and strict inequality holds for at least one of the inequalities, contradicting expression (1).

Subsequently, we prove the **necessity** of the BDP property for full surplus extraction. Suppose there exists  $i \in I$  and  $\bar{\theta}_i, \hat{\theta}_i \in \Theta$  with  $\bar{\theta}_i \neq \hat{\theta}_i$  such that  $p_i(\cdot | \bar{\theta}_i) = p_i(\cdot | \hat{\theta}_i)$  and surplus extraction can be guaranteed. Consider a private value auction environment with one dimensional valuations satisfying  $\bar{\theta}_i > \hat{\theta}_i > \theta_j$  for  $(j, \theta_j) \neq (i, \bar{\theta}_i)$  and  $(i, \hat{\theta}_i)$ . Note that full surplus extraction requires  $i$  obtain the good. Then the argument in the first paragraph and interim incentive compatibility require that

$$\inf_{\phi \in \Phi} \sum_{\theta_{-i} \in \Theta_{-i}} (\bar{\theta}_i + \phi_i(\bar{\theta}_i, \theta_{-i})) p_i(\theta_{-i} | \bar{\theta}_i) = 0 \geq \inf_{\phi \in \Phi} \sum_{\theta_{-i} \in \Theta_{-i}} (\hat{\theta}_i + \phi_i(\hat{\theta}_i, \theta_{-i})) p_i(\theta_{-i} | \bar{\theta}_i).$$

The above inequality,  $p_i(\cdot | \bar{\theta}_i) = p_i(\cdot | \hat{\theta}_i)$ , and the fact that  $\bar{\theta}_i > \hat{\theta}_i$  imply

$$\inf_{\phi \in \Phi} \sum_{\theta_{-i} \in \Theta_{-i}} (\hat{\theta}_i + \phi_i(\hat{\theta}_i, \theta_{-i})) p_i(\theta_{-i} | \hat{\theta}_i) < 0,$$

which violates individual rationality of type- $\hat{\theta}_i$  agent  $i$ .

To demonstrate the **sufficiency** of the BDP property, pick an arbitrary ex-post efficient allocation rule  $q$ . For each  $i \in I$ , define  $\Phi_i = \{-\eta_i + c\psi_i, -\eta_i - c\psi_i\}$ , where  $\psi_i$  is defined and proved to exist in Lemma A.3,  $\eta_i(\theta) = u_i(q(\theta), \theta)$  for all  $\theta \in \Theta$ , and  $c$  is no less than

$$\max_{\substack{i, \bar{\theta}_i, \hat{\theta}_i \in \Theta_i, \\ \bar{\theta}_i \neq \hat{\theta}_i}} \frac{\sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\hat{\theta}_i, \theta_{-i}), (\bar{\theta}_i, \theta_{-i})) - u_i(q(\hat{\theta}_i, \theta_{-i}), (\hat{\theta}_i, \theta_{-i}))] p_i(\theta_{-i} | \bar{\theta}_i)}{|\sum_{\theta_{-i} \in \Theta_{-i}} \psi_i(\hat{\theta}_i, \theta_{-i}) p_i(\theta_{-i} | \bar{\theta}_i)|}.$$

Let the set of ambiguous transfers be  $\Phi = \times_{i \in I} \Phi_i$ . All interim individual rationality constraints bind because  $-\eta_i$  extracts agent  $i$ 's surplus and  $c\psi_i$  has expected value of zero. To check incentive compatibility, notice that the choice of  $c$  gives agents non-positive worst-case expected payoffs when they misreport. Hence,  $\Phi = \times_{i \in I} \Phi_i$  extracts full surplus.  $\square$

**Proof of Theorem 5.2. Necessity.** Suppose there exists  $i \in I$ ,  $\bar{\theta}_i, \hat{\theta}_i \in \Theta$  such that  $p_i(\cdot | \bar{\theta}_i) = p_i(\cdot | \hat{\theta}_i)$ . We will establish the existence of a profile of utility functions and an efficient allocation rule  $q$  such that  $q$  cannot be implemented via an individually rational and budget-balanced mechanism with ambiguous transfers.

Consider an adaptation of the utility functions constructed by Kosenok and Severinov (2008). Let  $A = \{x_0, x_1, x_2\}$ , where all agents' payoffs of consuming the outside option  $x_0$  are zero. The payoffs for agent  $i$  and all  $j \neq i$  to consume  $x_1$  and  $x_2$  are given below with  $0 < a < B$ .

	$u_i(x_1, (\theta_i, \theta_j))$	$u_j(x_1, (\theta_i, \theta_j))$	$u_i(x_2, (\theta_i, \theta_j))$	$u_j(x_2, (\theta_i, \theta_j))$
$\theta_i = \bar{\theta}_i$	a	a	a+B	a-2B
$\theta_i = \hat{\theta}_i$	0	a	a	a
$\theta_i \neq \bar{\theta}_i, \hat{\theta}_i$	a	a	0	a

The efficient allocation rule is  $q(\theta) = x_2$  if  $\theta_i = \hat{\theta}_i$  and  $q(\theta) = x_1$  elsewhere.

Suppose by way of contradiction that there exists an interim individually rational and ex-post budget-balanced mechanism with ambiguous transfers implementing  $q$ . Denote the set of transfers by  $\Phi$ . Then from  $IC(\bar{\theta}_i, \hat{\theta}_i)$  and  $IC(\hat{\theta}_i, \bar{\theta}_i)$ ,

$$\begin{aligned}
a + \inf_{\phi \in \Phi} \sum_{\theta_{-i} \in \Theta_{-i}} \phi_i(\bar{\theta}_i, \theta_{-i}) p_i(\theta_{-i} | \bar{\theta}_i) &\geq a + B + \inf_{\phi \in \Phi} \sum_{\theta_{-i} \in \Theta_{-i}} \phi_i(\hat{\theta}_i, \theta_{-i}) p_i(\theta_{-i} | \bar{\theta}_i), \\
a + \inf_{\phi \in \Phi} \sum_{\theta_{-i} \in \Theta_{-i}} \phi_i(\hat{\theta}_i, \theta_{-i}) p_i(\theta_{-i} | \hat{\theta}_i) &\geq 0 + \inf_{\phi \in \Phi} \sum_{\theta_{-i} \in \Theta_{-i}} \phi_i(\bar{\theta}_i, \theta_{-i}) p_i(\theta_{-i} | \hat{\theta}_i).
\end{aligned}$$

Recall that  $p_i(\cdot | \bar{\theta}_i) = p_i(\cdot | \hat{\theta}_i)$ . Adding the above two inequalities gives  $2a \geq a + B$ , a contradiction. Therefore,  $q$  is not implementable via an interim individually rational and ex-post budget-balanced mechanism with ambiguous transfers.

**Sufficiency.** We pick an arbitrary interim individually rational and ex-post budget-balanced transfer rule  $\eta : \Theta \rightarrow \mathbb{R}^n$ . According to Lemma A.3, there exists a budget-balanced transfer rule  $\psi$  such that for all  $i \in I$ , truthfully revealing gives  $i$  expected values of zero and misreporting gives her non-zero ones.

Pick any  $c$  that is no less than

$$\max_{\substack{i \in I, \bar{\theta}_i, \hat{\theta}_i \in \Theta_i, \\ \bar{\theta}_i \neq \hat{\theta}_i}} \frac{\sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\hat{\theta}_i, \theta_{-i}), (\bar{\theta}_i, \theta_{-i})) + \eta_i(\hat{\theta}_i, \theta_{-i}) - u_i(q(\bar{\theta}_i, \theta_{-i}), (\bar{\theta}_i, \theta_{-i})) - \eta_i(\bar{\theta}_i, \theta_{-i})] p_i(\theta_{-i} | \bar{\theta}_i)}{|\sum_{\theta_{-i} \in \Theta_{-i}} \psi_i(\hat{\theta}_i, \theta_{-i}) p_i(\theta_{-i} | \bar{\theta}_i)|},$$

where  $c$  exists because the denominator is positive. Let  $\mathcal{M}$  be  $(\Theta, q, \{\eta + c\psi, \eta - c\psi\})$ .

Interim individual rationality of  $\mathcal{M}$  comes from the fact that  $\eta$  is individually rational and that  $\psi$  gives all truth-telling agents expected values of zero. For all  $i \in I$  and  $\bar{\theta}_i, \hat{\theta}_i \in \Theta_i$  with  $\bar{\theta}_i \neq \hat{\theta}_i$ , the choice of  $c$  indicates that

$$\min \left\{ \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\hat{\theta}_i, \theta_{-i}), (\bar{\theta}_i, \theta_{-i})) + \eta_i(\hat{\theta}_i, \theta_{-i}) \pm c\psi_i(\hat{\theta}_i, \theta_{-i})] p_i(\theta_{-i} | \bar{\theta}_i) \right\}$$

$$\leq \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\bar{\theta}_i, \theta_{-i}), (\bar{\theta}_i, \theta_{-i})) + \eta_i(\bar{\theta}_i, \theta_{-i})] p_i(\theta_{-i} | \bar{\theta}_i),$$

which further implies interim incentive compatibility of  $\mathcal{M}$ . Ex-post budget balance of  $\mathcal{M}$  follows from budget balance of  $\eta$  and  $\psi$ . Therefore,  $\mathcal{M}$  is an individually rational and budget-balanced mechanism with ambiguous transfers that implements  $q$ .  $\square$

**Proof of Theorem 6.1. Necessity.** By relabeling the indices, we assume without loss of generality that agent 1 has distinct beliefs under  $\theta_1^1$  and  $\theta_1^2$ , that agent 2 has distinct beliefs under  $\theta_2^1$  and  $\theta_2^2$ , and that  $L_2 \geq L_1$ . Let  $\theta_1^m$ ,  $\theta_2^n$ , and  $\theta_{-1,2}$  be generic elements of  $\Theta_1$ ,  $\Theta_2$ , and  $\Theta_{-1,2}$ . We ignore  $\theta_{-1,2}$  if  $N = 2$ . Now we construct a profile of private value utility functions such that an efficient outcome is not implementable. This would establish the necessity of the condition that at least  $N - 1$  agents satisfy the BDP property.

Suppose agent 1 owns a unit of private good and all other agents are potential buyers. Let  $\theta_i$  represent agent  $i$ 's private value of trading, where  $\theta_2^1 > -\theta_1^1 > \theta_2^2 > -\theta_1^2 > \dots > \theta_2^{L_1} > -\theta_1^{L_1}$  and  $-\theta_1^{L_1} > \theta_i > 0$  for all other agent-type combinations. No trade gives all agents zero payoffs. The efficient allocation rule  $q$  is that agent 1 should trade with agent 2 if  $\theta_1 + \theta_2 > 0$  and not trade otherwise.

Suppose by way of contradiction that an individually rational and budget-balanced mechanism with ambiguous transfers, denoted by  $\mathcal{M} = (q, \Phi)$ , implements  $q$ . By individual rationality, for all  $i \in I$  and  $\theta_i$ , type- $\theta_i$  agent  $i$ 's worst-case expected utility from participation is  $U_{\theta_i} \geq 0$ . Hence, for all  $\phi \in \Phi$ , for all  $i \in I$  and  $\theta_i \in \Theta_i$ ,

$$\sum_{\theta_{-i} \in \Theta_{-i}} p_i(\theta_{-i} | \theta_i) \phi_i(\theta_i, \theta_{-i}) \geq U_{\theta_i} - \sum_{\theta_{-i} \in \Theta_{-i}} u_i(q(\theta_i, \theta_{-i}), (\theta_i, \theta_{-i})) p_i(\theta_{-i} | \theta_i). \quad (7)$$

Multiply each of the inequalities by  $p_i(\theta_i)$  and sum across all  $i$  and  $\theta_i$ . By ex-post budget balance, the left-hand side of the aggregated inequality is zero and the right-hand side,

$$\sum_m p_1(\theta_1^m) \left( -\theta_1^m \sum_{n \leq m} p_1(\theta_2^n | \theta_1^m) \right) + \sum_n p_2(\theta_2^n) \left( -\theta_2^n \sum_{m \geq n} p_2(\theta_1^m | \theta_2^n) \right) + \sum_{i \in I} \sum_{\theta_i \in \Theta_i} p_i(\theta_i) U_{\theta_i}, \quad (8)$$

is non-positive. From  $IC(\theta_1^2 \theta_1^1)$  and  $IC(\theta_2^1 \theta_2^2)$ , for all  $\epsilon > 0$ , there exists  $\phi^1, \phi^2 \in \Phi$  satisfying

$$\begin{aligned} IC(\theta_1^2 \theta_1^1) & - \sum_{n, \theta_{-1,2}} p_1(\theta_2^n, \theta_{-1,2} | \theta_1^2) \phi_1^1(\theta_1^1, \theta_2^n, \theta_{-1,2}) + \epsilon \geq -U_1^2 + \theta_1^2 \sum_{n \leq 1} p_1(\theta_2^n | \theta_1^2), \\ IC(\theta_2^1 \theta_2^2) & - \sum_{m, \theta_{-1,2}} p_2(\theta_1^m, \theta_{-1,2} | \theta_2^1) \phi_2^2(\theta_1^m, \theta_2^2, \theta_{-1,2}) + \epsilon \geq -U_2^1 + \theta_2^1 \sum_{m \geq 2} p_2(\theta_1^m | \theta_2^1). \end{aligned}$$

In view of the assumption that  $p_1(\cdot | \theta_1^1) = p_1(\cdot | \theta_1^2)$  and  $p_2(\cdot | \theta_2^1) = p_2(\cdot | \theta_2^2)$ , by adding  $IC(\theta_2^2 \theta_1^1)$  and (7) with  $\theta_i = \theta_1^1$  and  $\phi = \phi^1$ , adding  $IC(\theta_2^1 \theta_2^2)$  and (7) with  $\theta_i = \theta_1^1$  with  $\theta_i = \theta_2^2$  and

$\phi = \phi^2$ , and letting  $\epsilon$  go to zero, we obtain that

$$U_1^2 \geq U_1^1 + (\theta_1^2 - \theta_1^1) \sum_{n \leq 1} p_1(\theta_2^n | \theta_1^1), \quad U_2^1 \geq U_2^2 + (\theta_2^1 - \theta_2^2) \sum_{m \geq 2} p_2(\theta_1^m | \theta_2^1).$$

By plugging the above two inequalities into expression (8), we have that (8) is no less than

$$\begin{aligned} & \sum_m p_1(\theta_1^m) \left( -\theta_1^m \sum_{n \leq m} p_1(\theta_2^n | \theta_1^m) \right) + p_1(\theta_1^2) (\theta_1^2 - \theta_1^1) \sum_{n \leq 1} p_1(\theta_2^n | \theta_1^1) \\ & \quad + \sum_n p_2(\theta_2^n) \left( -\theta_2^n \sum_{m \geq n} p_2(\theta_1^m | \theta_2^n) \right) + p_2(\theta_2^1) (\theta_2^1 - \theta_2^2) \sum_{m \geq 2} p_2(\theta_1^m | \theta_2^1). \end{aligned} \quad (9)$$

In the above expression, the coefficients of  $\theta_1^1$  and  $\theta_2^1$  are

$$\begin{aligned} & -p_1(\theta_1^1) \sum_{n \leq 1} p_1(\theta_2^n | \theta_1^1) - p_1(\theta_1^2) \sum_{n \leq 1} p_1(\theta_2^n | \theta_1^1) = -(p_1(\theta_1^1) + p_1(\theta_1^2)) \frac{p_{1,2}(\theta_1^1, \theta_2^1)}{p_1(\theta_1^1)} < -p_{1,2}(\theta_1^1, \theta_2^1), \\ & -p_2(\theta_2^1) \sum_{m \geq 1} p_2(\theta_1^m | \theta_2^1) + p_2(\theta_2^2) \sum_{m \geq 2} p_2(\theta_1^m | \theta_2^2) = -p_2(\theta_2^1) \frac{p_{1,2}(\theta_1^1, \theta_2^1)}{p_2(\theta_2^1)} = -p_{1,2}(\theta_1^1, \theta_2^1), \end{aligned}$$

where the strict inequality follows from Assumption 2.1. Let  $\theta_1^1$  and  $\theta_2^1$  be sufficiently close in absolute value and all other values  $\theta_i$  be close to zero. Then expression (9) is positive, contradicting  $0 \geq (8) \geq (9)$ . Therefore,  $q$  cannot be implemented via an individually rational and budget-balanced mechanism with ambiguous transfers.

**Sufficiency.** When all agents satisfy the BDP property, the sufficiency part is proven by Theorem 5.2. When there is exactly one agent,  $i$ , whose BDP property fails, following Lemmas A.1 through A.3, one can prove that there exists  $\psi : \Theta \rightarrow \mathbb{R}^n$  such that

1.  $\sum_{j \in I} \psi_j(\theta) = 0$  for all  $\theta \in \Theta$ ;
2.  $\sum_{\theta_{-j} \in \Theta_{-j}} \psi_j(\theta_j, \theta_{-j}) p_j(\theta_{-j} | \theta_j) = 0$  for all  $j \in I$  and  $\theta_j \in \Theta_j$ ;
3.  $\sum_{\theta_{-j} \in \Theta_{-j}} \psi_j(\hat{\theta}_j, \theta_{-j}) p_j(\theta_{-j} | \bar{\theta}_j) \neq 0$  for all  $j \neq i$  and  $\bar{\theta}_j, \hat{\theta}_j \in \Theta_j$  satisfying  $\bar{\theta}_j \neq \hat{\theta}_j$ ,

Notice that the third statement is different from the one in Lemma A.3, as agent  $i$  in this theorem has identical beliefs under different types.

We construct a mechanism where agent  $i$  obtains all the surplus by truthfully reporting. For all  $\theta \in \Theta$  and  $j \in I$  with  $j \neq i$ , let  $\eta_j(\theta) = -u_j(q(\theta), \theta_j)$ , and  $\eta_i(\theta) = -\sum_{j \neq i} \eta_j(\theta)$ .

Pick any  $c$  that is no less than

$$\max_{\substack{j \neq i, \hat{\theta}_j, \bar{\theta}_j \in \Theta_j, \\ \bar{\theta}_j \neq \hat{\theta}_j}} \frac{\sum_{\theta_{-j} \in \Theta_{-j}} [u_j(q(\hat{\theta}_j, \theta_{-j}), \theta_j) + \eta_j(\hat{\theta}_j, \theta_{-j})] p_j(\theta_{-j} | \bar{\theta}_j)}{|\sum_{\theta_{-j} \in \Theta_{-j}} \psi_j(\hat{\theta}_j, \theta_{-j}) p_j(\theta_{-j} | \bar{\theta}_j)|}.$$

Let the set of ambiguous transfers be  $\Phi = \{\eta + c\psi, \eta - c\psi\}$ , which is interim individually rational and ex-post budget-balanced. The choice of  $\eta$ ,  $\psi$ , and  $c$  implies that for any agent  $j \neq i$  with type  $\bar{\theta}_j$ , truthfully reporting gives her zero worst-case expected payoff while lying gives her non-positive ones. Therefore,  $j$ 's incentive compatibility constraints are satisfied.

For type- $\bar{\theta}_i$  agent  $i$ , the argument below verifies her incentive compatibility constraints:

$$\begin{aligned}
& \min\left\{ \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\bar{\theta}_i, \theta_{-i}), \bar{\theta}_i) + \sum_{j \neq i} u_j(q(\bar{\theta}_i, \theta_{-i}), \theta_j) \pm c\psi_i(\bar{\theta}_i, \theta_{-i})] p_i(\theta_{-i} | \bar{\theta}_i) \right\} \\
&= \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\bar{\theta}_i, \theta_{-i}), \bar{\theta}_i) + \sum_{j \neq i} u_j(q(\bar{\theta}_i, \theta_{-i}), \theta_j)] p_i(\theta_{-i} | \bar{\theta}_i) \\
&\geq \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\hat{\theta}_i, \theta_{-i}), \bar{\theta}_i) + \sum_{j \neq i} u_j(q(\hat{\theta}_i, \theta_{-i}), \theta_j)] p_i(\theta_{-i} | \bar{\theta}_i) \\
&\geq \min\left\{ \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\hat{\theta}_i, \theta_{-i}), \bar{\theta}_i) + \sum_{j \neq i} u_j(q(\hat{\theta}_i, \theta_{-i}), \theta_j) \pm c\psi_i(\hat{\theta}_i, \theta_{-i})] p_i(\theta_{-i} | \bar{\theta}_i) \right\},
\end{aligned}$$

where the equality comes from the second property of  $\psi$ , the first inequality comes from ex-post efficiency of  $q$ , and the second inequality comes from the minimization operation.

Therefore, the individually rational and budget-balanced mechanism with ambiguous transfers implements  $q$ .  $\square$

**Example A.1:** *This example demonstrates that if  $N - 1$  agents satisfy the BDP property, under private value environments, an inefficient allocation rule may not be implementable via an individually rational and budget-balanced mechanism with ambiguous transfers.*

*Consider a common prior  $p$  defined by  $p(\theta_1^3, \theta_2^2) = 2/7$ , and  $p(\theta) = 1/7$  for all other  $\theta$ . Let the set of feasible allocations be  $A = \{x_0, x_1, x_2\}$ . Recall that  $x_0$ , the outside option, gives both agents zero payoffs. The payoffs of  $x_1$  and  $x_2$  are presented below.*

$x_1$	$\theta_2^1$	$\theta_2^2$
$\theta_1^1$	0,0	0,0
$\theta_1^2$	2,0	2,0
$\theta_1^3$	0,0	0,0

$x_2$	$\theta_2^1$	$\theta_2^2$
$\theta_1^1$	2,0	2,0
$\theta_1^2$	0,0	0,0
$\theta_1^3$	0,0	0,0

*Consider an allocation rule  $q(\theta) = x_2$  if  $\theta_1 = \theta_1^2$ , and  $q(\theta) = x_1$  elsewhere. Suppose by way of contradiction that  $q$  is implemented by a mechanism with ambiguous transfers  $\mathcal{M} = (\Theta, q, \Phi)$ , where each  $\phi \in \Phi$  is a transfer or payment from agent 1 to 2. Let  $U_1^1$  denote type- $\theta_1^1$  agent 1's worst-case expected payoff from participation. Similarly, let  $U_1^2$  denote type- $\theta_1^2$  agent 1's payoff.*

*As  $IR(\theta_1^1)$  and  $IC(\theta_1^2\theta_1^1)$  hold, for any  $\epsilon > 0$ , there exists  $\phi^1 \in \Phi$  such that*

$$IR(\theta_1^1) - 0.5\phi^1(\theta_1^1, \theta_2^1) - 0.5\phi^1(\theta_1^1, \theta_2^2) \geq U_1^1,$$

$$IC(\theta_1^2\theta_1^1) \quad U_1^2 + \epsilon \geq 2 - 0.5\phi^1(\theta_1^1, \theta_2^1) - 0.5\phi^1(\theta_1^1, \theta_2^2).$$

Similarly, by  $IR(\theta_1^2)$  and  $IC(\theta_1^1\theta_1^2)$ , for any  $\epsilon > 0$ , there exists  $\phi^2 \in \Phi$  such that

$$\begin{aligned} IR(\theta_1^2) & \quad -0.5\phi^2(\theta_1^2, \theta_2^1) - 0.5\phi^2(\theta_1^2, \theta_2^2) \geq U_1^2, \\ IC(\theta_1^1\theta_1^2) & \quad U_1^1 + \epsilon \geq 2 - 0.5\phi^2(\theta_1^2, \theta_2^1) - 0.5\phi^2(\theta_1^2, \theta_2^2). \end{aligned}$$

We add the above inequalities pairwise and let  $\epsilon$  go to zero. Thus we have  $U_1^2 \geq 2 + U_1^1$  and  $U_1^1 \geq 2 + U_1^2$ . These two expressions imply  $0 \geq 4$ , which is a contradiction.

**Proof of Proposition 6.1.** By relabelling the indices, we assume without loss of generality there are  $(\beta_{\theta_1})_{\theta_1 \neq \theta_1^1}, (\beta_{\theta_2})_{\theta_2 \neq \theta_2^2} \in \Delta$  such that  $p_1(\cdot|\theta_1^1) = \sum_{\theta_1 \neq \theta_1^1} \beta_{\theta_1} p_1(\cdot|\theta_1)$  and  $p_2(\cdot|\theta_2^2) =$

$\sum_{\theta_2 \neq \theta_2^2} \beta_{\theta_2} p_2(\cdot|\theta_2)$ , and

$$\frac{\beta_{\theta_2^1}}{p_2(\theta_2^1)} \geq \frac{\beta_{\theta_2}}{p_2(\theta_2)}, \forall \theta_2 \neq \theta_2^1, \theta_2^2. \quad (10)$$

Suppose agent 1 owns a unit of private good and all others are potential buyers. For each  $i \in I$ , let  $\theta_i$  be agent  $i$ 's private value of trading, where  $\theta_2^1 > -\theta_1^1 > \theta_2^2 > -\theta_1^2 > \dots > \theta_2^{L_1} > -\theta_1^{L_1} > \theta_i$  for all other  $\theta_i$ . No trade gives all agents zero payoffs. The efficient allocation rule  $q$  is that agent 1 should trade with 2 if  $\theta_1 + \theta_2 > 0$  and not trade otherwise. Subsequently, we will prove that  $q$  is not implementable, which proves the necessity of the condition.

Suppose by way of contradiction there exists an individually rational and budget-balanced Bayesian transfer  $\phi$  that implements  $q$ . Then by individual rationality and incentive compatibility, for all  $i \in I$ ,  $\bar{\theta}_i \neq \hat{\theta}_i$ , the following inequalities hold:

$$\begin{aligned} IR(\bar{\theta}_i) & \quad \sum_{\theta_{-i}} p_i(\theta_{-i}|\bar{\theta}_i) \phi_i(\bar{\theta}_i, \theta_{-i}) \geq - \sum_{\theta_{-i}} u_i(q(\bar{\theta}_i, \theta_{-i}), \bar{\theta}_i) p_i(\theta_{-i}|\bar{\theta}_i), \\ IC(\bar{\theta}_i \hat{\theta}_i) & \quad \sum_{\theta_{-i}} p_i(\theta_{-i}|\bar{\theta}_i) \phi_i(\bar{\theta}_i, \theta_{-i}) - \sum_{\theta_{-i}} p_i(\theta_{-i}|\hat{\theta}_i) \phi_i(\hat{\theta}_i, \theta_{-i}) \\ & \quad \geq - \sum_{\theta_{-i}} u_i(q(\bar{\theta}_i, \theta_{-i}), \bar{\theta}_i) p_i(\theta_{-i}|\bar{\theta}_i) - \sum_{\theta_{-i}} u_i(q(\hat{\theta}_i, \theta_{-i}), \bar{\theta}_i) p_i(\theta_{-i}|\bar{\theta}_i). \end{aligned}$$

We choose a constant  $\delta > 0$  sufficiently large such that

$$\frac{\delta \beta_{\theta_2^1} p_2(\theta_2^1)}{p_2(\theta_2^1)} \geq \frac{\beta_{\theta_1} p_1(\theta_1^1)}{p_1(\theta_1^1)}, \forall \theta_1 \neq \theta_1^1. \quad (11)$$

and then denote the left-hand-side term by  $\gamma$ . Now we compute the weighted sum of the above individual rationality and incentive compatibility constraints where (1) the weight of  $IR(\theta_1^1)$  is  $p_1(\theta_1^1)(\gamma + 1)$ , (2) for each  $\theta_1 \neq \theta_1^1$  the weight of  $IR(\theta_1)$  is  $p_1(\theta_1)\gamma - \beta_{\theta_1} p_1(\theta_1^1)$ ,

(3) the weight of  $IR(\theta_2^2)$  is  $p_2(\theta_2^2)(\gamma + \delta)$ , (4) for each  $\theta_2 \neq \theta_2^2$  the weight of  $IR(\theta_2)$  is  $p_2(\theta_2)\gamma - \delta\beta_{\theta_2}p_2(\theta_2^2)$ , (5) for each  $i \neq 1, 2$  and  $\theta_i \in \Theta_i$  the weight of  $IR(\theta_i)$  is  $p_i(\theta_i)\gamma$ , (6) for each  $\theta_1 \neq \theta_1^1$  the weight of  $IC(\theta_1\theta_1^1)$  is  $p_1(\theta_1^1)\beta_{\theta_1}$ , (7) for each  $\theta_2 \neq \theta_2^2$  the weight of  $IC(\theta_2\theta_2^2)$  is  $\delta\beta_{\theta_2}p(\theta_2^2)$ , (8) every other inequality has weight zero. From expressions (10) and (11), we know all the weights are non-negative.

Ex-post budget balance cancels all terms containing transfers in the weighted sum and thus the left-hand side is zero. On the right-hand side, the coefficients of  $\theta_1^1$  and  $\theta_2^1$  are  $-(\gamma + 1)p(\theta_1^1\theta_2^1)$  and  $-\gamma p(\theta_1^1\theta_2^1)$  respectively. Therefore, by choosing  $\theta_1^1$  and  $\theta_2^1$  sufficiently close in absolute value and all other  $\theta_i$  close to zero, the right-hand side of the weighted sum is positive, a contradiction.  $\square$

**Lemma A.4:** *Given the belief system  $(p_i(\cdot|\theta_i))_{i \in I, \theta_i \in \Theta_i}$ , the BDP and NCP\* properties hold for agent  $i$  if for all  $\bar{\theta}_i, \hat{\theta}_i \in \Theta_i$  with  $\bar{\theta}_i \neq \hat{\theta}_i$ , there exists  $\psi^{\bar{\theta}_i, \hat{\theta}_i} : \Theta \rightarrow \mathbb{R}^n$  such that,*

1.  $\sum_{j \in I} \psi_j^{\bar{\theta}_i, \hat{\theta}_i}(\theta) = 0$  for all  $\theta \in \Theta$ ;
2.  $\sum_{\theta_{-j} \in \Theta_{-j}} \psi_j^{\bar{\theta}_i, \hat{\theta}_i}(\theta_j, \theta_{-j}) p_j(\theta_{-j}|\theta_j) \geq 0$  for all  $j \in I, \theta_j \in \Theta_j$ ;
3.  $\sum_{\theta_{-i} \in \Theta_{-i}} \psi_i^{\bar{\theta}_i, \hat{\theta}_i}(\hat{\theta}_i, \theta_{-i}) p_i(\theta_{-i}|\hat{\theta}_i) < 0$ .

*Proof.* By Motzkin's theorem of the alternative, the above system has a solution if and only if there do not exist coefficients  $(b_\theta)_{\theta \in \Theta}$  and  $(a_{\theta_j})_{j \in I, \theta_j \in \Theta_j} \geq \mathbf{0}$  such that

$$p_{\bar{\theta}_i, \hat{\theta}_i} = \sum_{j \in I} \sum_{\theta_j \in \Theta_j} a_{\theta_j} p_{\theta_j} - \sum_{\theta \in \Theta} b_\theta e_\theta. \quad (12)$$

To prove this lemma, we will subsequently establish that there are coefficients  $(b_\theta)_{\theta \in \Theta}$  and  $(a_{\theta_j})_{j \in I, \theta_j \in \Theta_j} \geq \mathbf{0}$ , such that expression (12) hold if and only if  $p_i(\cdot|\bar{\theta}_i) = p_i(\cdot|\hat{\theta}_i)$  or both conditions in NCP\* property are satisfied by  $(i, \bar{\theta}_i, \hat{\theta}_i)$ .

To prove the “**only if**” direction, suppose there are coefficients  $(b_\theta)_{\theta \in \Theta}$  and  $(a_{\theta_j})_{j \in I, \theta_j \in \Theta_j} \geq \mathbf{0}$ , such that expression (12) holds. Then,

$$a_{\theta_i} p_i(\theta_j \theta_{-i-j}|\theta_i) = b_{\theta_i, \theta_j \theta_{-i-j}}, \forall \theta_i \neq \hat{\theta}_i, j \neq i, \theta_j, \theta_{-i-j}, \quad (13)$$

$$a_{\hat{\theta}_i} p_i(\theta_j \theta_{-i-j}|\hat{\theta}_i) - p_i(\theta_j \theta_{-i-j}|\bar{\theta}_i) = b_{\hat{\theta}_i, \theta_j \theta_{-i-j}}, \forall j \neq i, \theta_j, \theta_{-i-j}, \quad (14)$$

$$a_{\theta_j} p_j(\theta_i \theta_{-i-j}|\theta_j) = b_{\theta_i, \theta_j \theta_{-i-j}}, \forall \theta_i, j \neq i, \theta_j, \theta_{-i-j}. \quad (15)$$

We remark that thoutghout the proof, if  $N = 2$ , we ignore any term  $\theta_{-i-j}$  to avoid introducing additional notation. By canceling  $b_{\bar{\theta}_i, \theta_j \theta_{-i-j}}$  in (13) and (15), we also have  $a_{\bar{\theta}_i} \geq 0$ .

**Case 1.** Suppose  $a_{\tilde{\theta}_i} = 0$  for some  $\tilde{\theta}_i \neq \hat{\theta}_i$ . The argument below shows that  $a_{\hat{\theta}_i} = 1$ ,  $a_{\theta_j} = 0$  for all  $(j, \theta_j) \neq (i, \hat{\theta}_i)$ ,  $b_{\theta_i, \theta_j, \theta_{-i-j}} = 0$  for all  $\theta_i, \theta_j$ , and  $\theta_{-i-j}$ , and  $p_i(\cdot | \bar{\theta}_i) = p_i(\cdot | \hat{\theta}_i)$ .

Canceling  $b_{\tilde{\theta}_i, \theta_j, \theta_{-i-j}}$  in (13) and (15) yields  $0 = a_{\tilde{\theta}_i} p_i(\theta_j \theta_{-i-j} | \tilde{\theta}_i) = a_{\theta_j} p_j(\tilde{\theta}_i \theta_{-i-j} | \theta_j)$  for all  $j \neq i, \theta_j, \theta_{-i-j}$ . From Assumption 6.1, it must be the case that  $a_{\theta_j} = 0$  for all  $j \neq i$  and  $\theta_j$ .

By expression (15), last paragraph implies  $b_{\theta_i, \theta_j, \theta_{-i-j}} = 0$  for all  $\theta_i, \theta_j$ , and  $\theta_{-i-j}$ . From expression (13), we further know  $a_{\theta_i} = 0$  for all  $\theta_i \neq \hat{\theta}_i$ .

By canceling  $b_{\hat{\theta}_i, \theta_j, \theta_{-i-j}}$  in (14) and (15), we have  $a_{\hat{\theta}_i} p_i(\theta_j \theta_{-i-j} | \hat{\theta}_i) - p_i(\theta_j \theta_{-i-j} | \bar{\theta}_i) = a_{\theta_j} p_j(\hat{\theta}_i \theta_{-i-j} | \theta_j) = 0$  for all  $\theta_j$  and  $\theta_{-i-j}$ . Summing the equation across all  $\theta_j$  and  $\theta_{-i-j}$ , we get  $a_{\hat{\theta}_i} = 1$  and thus  $p_i(\cdot | \bar{\theta}_i) = p_i(\cdot | \hat{\theta}_i)$ .

**Case 2.** Suppose  $a_{\theta_i} > 0$  for all  $\theta_i \neq \hat{\theta}_i$ . Following the symmetric argument of the previous case, we know  $a_{\hat{\theta}_i} > 1$  and  $a_{\theta_j} > 0$  for all  $(j, \theta_j) \neq (i, \hat{\theta}_i)$ . Subsequently, we will establish that  $(i, \bar{\theta}_i, \hat{\theta}_i)$  satisfies both conditions in the NCP\* property so that the property fails.

Define  $\mu \in \Delta(\Theta)$  by  $\mu(\theta) = \frac{b_\theta}{\sum_{\bar{\theta} \in \Theta} b_{\bar{\theta}}}$  for all  $\theta \in \Theta$ . Then from expressions (13) and (15), we know  $\mu(\cdot | \theta_j) = p_j(\cdot | \theta_j)$  and  $\mu(\theta_j) = \frac{a_{\theta_j}}{\sum_{\bar{\theta} \in \Theta} b_{\bar{\theta}}} > 0$  for all  $(j, \theta_j) \neq (i, \hat{\theta}_i)$ . Hence, the first condition holds in the NCP\* property. By canceling  $b_{\hat{\theta}_i, \theta_j, \theta_{-i-j}}$  in expressions (14) and (15), we have  $a_{\hat{\theta}_i} p_i(\theta_j, \cdot | \hat{\theta}_i) = p_i(\theta_j, \cdot | \bar{\theta}_i) + a_{\theta_j} p_j(\hat{\theta}_i, \cdot | \theta_j)$  for all  $j \neq i$  and  $\theta_j$ , where  $a_{\theta_j} = \mu(\theta_j) \sum_{\bar{\theta} \in \Theta} b_{\bar{\theta}} = \mu(\bar{\theta}_i) \frac{\mu(\theta_j | \bar{\theta}_i)}{\mu(\theta_i | \bar{\theta}_i)} \sum_{\bar{\theta} \in \Theta} b_{\bar{\theta}} = a_{\bar{\theta}_i} \frac{p_i(\theta_j | \bar{\theta}_i)}{p_j(\theta_i | \bar{\theta}_i)}$ . Recall  $a_{\bar{\theta}_i} > 0$  and  $a_{\hat{\theta}_i} > 1$ . Thus by defining  $\bar{C} = a_{\bar{\theta}_i}$  and  $\hat{C} = a_{\hat{\theta}_i}$ , the second condition also holds in the NCP\* property.  $\square$

**Proof of Theorem 6.2.** Suppose the BDP and NCP\* properties hold for all agents. According to Lemma A.4, for all  $i \in I$  and  $\bar{\theta}_i, \hat{\theta}_i \in \Theta_i$  with  $\bar{\theta}_i \neq \hat{\theta}_i$ , there exists  $\psi^{\bar{\theta}_i, \hat{\theta}_i} : \Theta \rightarrow \mathbb{R}^n$ , such that the three requirements are satisfied.

Let  $\eta$  be any interim individually rational and ex-post budget-balanced transfer rule. Define  $\Phi = \{\eta, \eta + c\psi^{\bar{\theta}_j, \hat{\theta}_j} : j \in I, \bar{\theta}_j, \hat{\theta}_j \in \Theta_j, \bar{\theta}_j \neq \hat{\theta}_j\}$ , where  $c$  is sufficiently large such that for all  $j \in I$  and  $\bar{\theta}_j, \hat{\theta}_j \in \Theta_j$  with  $\bar{\theta}_j \neq \hat{\theta}_j$ , the following term is negative:

$$\sum_{\theta_{-j} \in \Theta_{-j}} [u_j(q(\hat{\theta}_j, \theta_{-j}), (\bar{\theta}_j, \theta_{-j})) - u_j(q(\bar{\theta}_j, \theta_{-j}), (\bar{\theta}_j, \theta_{-j})) + \eta_j(\hat{\theta}_j, \theta_{-j}) - \eta_j(\bar{\theta}_j, \theta_{-j}) + c\psi_j^{\bar{\theta}_j, \hat{\theta}_j}(\hat{\theta}_j, \theta_{-j})] p_j(\theta_{-j} | \bar{\theta}_j).$$

For any type- $\bar{\theta}_i$  agent  $i$ , the inequality below shows that misreporting  $\hat{\theta}_i$  is not profitable:

$$\begin{aligned} & \min_{\phi \in \Phi} \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\bar{\theta}_i, \theta_{-i}), (\bar{\theta}_i, \theta_{-i})) + \eta(\bar{\theta}_i, \theta_{-i}) + \phi_i(\bar{\theta}_i, \theta_{-i})] p_i(\theta_{-i} | \bar{\theta}_i) \\ &= \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\bar{\theta}_i, \theta_{-i}), (\bar{\theta}_i, \theta_{-i})) + \eta(\bar{\theta}_i, \theta_{-i})] p_i(\theta_{-i} | \bar{\theta}_i) \\ &\geq \min_{\phi \in \Phi} \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\hat{\theta}_i, \theta_{-i}), (\bar{\theta}_i, \theta_{-i})) + \eta(\hat{\theta}_i, \theta_{-i}) + \phi_i(\hat{\theta}_i, \theta_{-i})] p_i(\theta_{-i} | \bar{\theta}_i), \end{aligned}$$

where the equality follows from the second and third requirements of Lemma A.4, and the inequality comes from the choice of  $c$ . Interim individual rationality and ex-post budget balance follow from corresponding properties of  $\eta$  and each  $\phi \in \Phi$ .  $\square$

**Lemma A.5:** *Given beliefs  $(p_i(\cdot|\theta_i))_{i \in I, \theta_i \in \Theta_i}$  that are not generated by a common prior, if the BDP property holds for all agents, then the NCP\* property holds for at least  $N-1$  agents.*

*Proof.* Suppose by way of contradiction that the BDP property holds for all agents, and that  $(i, \bar{\theta}_i, \hat{\theta}_i \neq \bar{\theta}_i)$  and  $(j \neq i, \bar{\theta}_j, \hat{\theta}_j \neq \bar{\theta}_j)$  satisfy the two conditions in the NCP\* property so that the property fails for  $i$  and  $j$ . From the two-case argument of Lemma A.4, there exist coefficients  $(a_{\theta_k})_{k \in I, \theta_k \in \Theta_k} > \mathbf{0}$  where  $a_{\hat{\theta}_i} > 1$ ,  $(b_{\theta})_{\theta \in \Theta}$ ,  $(c_{\theta_k})_{k \in I, \theta_k \in \Theta_k} > \mathbf{0}$  where  $c_{\hat{\theta}_j} > 1$ , and  $(d_{\theta})_{\theta \in \Theta}$  such that  $p_{\bar{\theta}_i, \hat{\theta}_i} = \sum_{k \in I} \sum_{\theta_k \in \Theta_k} a_{\theta_k} p_{\theta_k, \theta_k} - \sum_{\theta \in \Theta} b_{\theta} e_{\theta}$ , and  $p_{\bar{\theta}_j, \hat{\theta}_j} = \sum_{k \in I} \sum_{\theta_k \in \Theta_k} c_{\theta_k} p_{\theta_k, \theta_k} - \sum_{\theta \in \Theta} d_{\theta} e_{\theta}$ . Thus, the following equations hold. Note that we ignore  $\theta_{-i-j}$  if  $N = 2$ .

$$\begin{aligned} a_{\theta_i} p_i(\theta_j, \theta_{-i-j} | \theta_i) &= b_{\theta_i, \theta_j, \theta_{-i-j}}, \forall \theta_i \neq \hat{\theta}_i, \forall \theta_j, \theta_{-i-j}, \\ a_{\hat{\theta}_i} p_i(\theta_j, \theta_{-i-j} | \hat{\theta}_i) - p_i(\theta_j, \theta_{-i-j} | \bar{\theta}_i) &= b_{\hat{\theta}_i, \theta_j, \theta_{-i-j}}, \forall \theta_j, \theta_{-i-j} \\ a_{\theta_j} p_j(\theta_i, \theta_{-i-j} | \theta_j) &= b_{\theta_i, \theta_j, \theta_{-i-j}}, \forall \theta_i, \theta_j, \theta_{-i-j}, \\ c_{\theta_i} p_i(\theta_j, \theta_{-i-j} | \theta_i) &= d_{\theta_i, \theta_j, \theta_{-i-j}}, \forall \theta_i, \theta_j, \theta_{-i-j}, \\ c_{\theta_j} p_j(\theta_i, \theta_{-i-j} | \theta_j) &= d_{\theta_i, \theta_j, \theta_{-i-j}}, \forall \theta_j \neq \hat{\theta}_j, \forall \theta_i, \theta_{-i-j}, \\ c_{\hat{\theta}_j} p_j(\theta_i, \theta_{-i-j} | \hat{\theta}_j) - p_j(\theta_i, \theta_{-i-j} | \bar{\theta}_j) &= d_{\theta_i, \hat{\theta}_j, \theta_{-i-j}}, \forall \theta_i, \theta_{-i-j}. \end{aligned}$$

Canceling all  $b_{\theta_i, \theta_j, \theta_{-i-j}}$ ,  $d_{\theta_i, \theta_j, \theta_{-i-j}}$ , and  $p_j(\theta_i, \theta_{-i-j} | \theta_j)$  in above equations yields:

$$\begin{aligned} \frac{a_{\theta_i} p_i(\theta_j, \theta_{-i-j} | \theta_i)}{a_{\theta_j}} &= \frac{c_{\theta_i} p_i(\theta_j, \theta_{-i-j} | \theta_i)}{c_{\theta_j}}, \forall \theta_i \neq \hat{\theta}_i, \theta_j \neq \hat{\theta}_j, \text{ and } \theta_{-i-j}, \\ \frac{a_{\hat{\theta}_i} p_i(\theta_j, \theta_{-i-j} | \hat{\theta}_i) - p_i(\theta_j, \theta_{-i-j} | \bar{\theta}_i)}{a_{\theta_j}} &= \frac{c_{\hat{\theta}_i} p_i(\theta_j, \theta_{-i-j} | \hat{\theta}_i)}{c_{\theta_j}}, \forall \theta_j \neq \hat{\theta}_j, \forall \theta_{-i-j}, \end{aligned} \quad (16)$$

$$\begin{aligned} \frac{a_{\theta_i} p_i(\hat{\theta}_j \theta_{-i-j} | \theta_i)}{a_{\hat{\theta}_j}} &= \frac{c_{\theta_i} p_i(\hat{\theta}_j \theta_{-i-j} | \theta_i)}{c_{\hat{\theta}_j}} + \frac{c_{\theta_i} p_i(\bar{\theta}_j \theta_{-i-j} | \theta_i)}{c_{\hat{\theta}_j} c_{\bar{\theta}_j}}, \forall \theta_i \neq \hat{\theta}_i, \forall \theta_{-i-j}, \\ \frac{a_{\hat{\theta}_i} p_i(\hat{\theta}_j \theta_{-i-j} | \hat{\theta}_i) - p_i(\hat{\theta}_j \theta_{-i-j} | \bar{\theta}_i)}{a_{\hat{\theta}_j}} &= \frac{c_{\hat{\theta}_i} p_i(\hat{\theta}_j \theta_{-i-j} | \hat{\theta}_i)}{c_{\hat{\theta}_j}} + \frac{c_{\hat{\theta}_i} p_i(\bar{\theta}_j \theta_{-i-j} | \hat{\theta}_i)}{c_{\hat{\theta}_j} c_{\bar{\theta}_j}}, \forall \theta_{-i-j}. \end{aligned} \quad (17)$$

By summing each of the four equation across all  $\theta_{-i-j} \in \Theta_{-i-j}$ , we have

$$a_{\theta_i} : c_{\theta_i} = a_{\theta_j} : c_{\theta_j} = (a_{\hat{\theta}_i} - 1) : c_{\hat{\theta}_i} = a_{\hat{\theta}_j} : \frac{c_{\hat{\theta}_j} c_{\bar{\theta}_j}}{c_{\bar{\theta}_j} + 1}, \forall \theta_i \neq \hat{\theta}_i, \theta_j \neq \hat{\theta}_j.$$

Plugging this relationship back into (16) and (17) yields  $p_i(\theta_j, \theta_{-i-j} | \bar{\theta}_i) = p_i(\theta_j, \theta_{-i-j} | \hat{\theta}_i)$  for all  $\theta_j \neq \hat{\theta}_j$  and  $\theta_{-i-j}$ , and  $p_i(\hat{\theta}_j \theta_{-i-j} | \hat{\theta}_i) = p_i(\hat{\theta}_j \theta_{-i-j} | \bar{\theta}_i)$ . Hence,  $p_i(\cdot | \bar{\theta}_i) = p_i(\cdot | \hat{\theta}_i)$ , a contradiction.  $\square$

**Lemma A.6:** *Let  $q$  be an efficient allocation rule under a private value environment. For any  $i \in I$ ,  $\tilde{\Theta}_i \subseteq \Theta_i$  with  $|\tilde{\Theta}_i| \geq 2$ , and distribution  $\pi \in \Delta(\Theta_{-i})$ , there exist values  $\{U_{\theta_i}\}_{\theta_i \in \tilde{\Theta}_i} \geq \mathbf{0}$  such that  $U_{\theta_i} - U_{\theta'_i} \geq \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\theta'_i, \theta_{-i}), \theta_i) - u_i(q(\theta_i, \theta_{-i}), \theta_i)] \pi(\theta_{-i})$  for all  $\theta_i, \theta'_i \in \tilde{\Theta}_i$ .*

*Proof.* Let a loop be a sequence  $(\theta_i^1, \theta_i^2, \dots, \theta_i^K)$  in  $\tilde{\Theta}_i$  with length  $K \geq 2$  and  $\theta_i^1 = \theta_i^K$ . As  $q$  is ex-post efficient,  $u_i(q(\theta_i^{k+1}, \theta_{-i}), \theta_i^{k+1}) + \sum_{j \neq i} u_j(q(\theta_i^{k+1}, \theta_{-i}), \theta_j) \geq u_i(q(\theta_i^k, \theta_{-i}), \theta_i^k) + \sum_{j \neq i} u_j(q(\theta_i^k, \theta_{-i}), \theta_j)$  for all  $k = 1, \dots, K-1$  and  $\theta_{-j} \in \Theta_{-j}$ . Summing the inequalities across  $k = 1, \dots, K-1$ , we obtain that  $\sum_{k=1}^{K-1} [u_i(q(\theta_i^{k+1}, \theta_{-i}), \theta_i^{k+1}) - u_i(q(\theta_i^k, \theta_{-i}), \theta_i^k)] \leq 0$ . This is the ‘‘cyclical monotonicity’’ condition in the literature.

From Theorem 1 of Rochet (1987), we know cyclical monotonicity is sufficient for the existence of ex-post incentive compatible transfers. Namely, for each  $\theta_{-i} \in \Theta_{-i}$  and loop  $(\theta_i^1, \theta_i^2, \dots, \theta_i^K)$  in  $\tilde{\Theta}_i$ , if  $\sum_{k=1}^{K-1} [u_i(q(\theta_i^{k+1}, \theta_{-i}), \theta_i^{k+1}) - u_i(q(\theta_i^k, \theta_{-i}), \theta_i^k)] \leq 0$ , then there exists a transfer rule  $t_i(\cdot, \theta_{-i})$  such that  $u_i(q(\theta_i, \theta_{-i}), \theta_i) + t_i(\theta_i, \theta_{-i}) \geq u_i(q(\theta'_i, \theta_{-i}), \theta_i) + t_i(\theta'_i, \theta_{-i})$  for  $\theta_i, \theta'_i \in \tilde{\Theta}_i$ . Let  $C$  be a constant sufficiently large such that  $\tilde{t}_i(\theta) = t_i(\theta) + C \geq 0$  for all  $\theta \in \tilde{\Theta}_i \times \Theta_{-i}$ . By defining  $U_{\theta_i} = \sum_{\theta_{-i} \in \Theta_{-i}} \tilde{t}_i(\theta_i, \theta_{-i}) \pi_i(\theta_{-i})$  for all  $\theta_i \in \tilde{\Theta}_i$ , we have established the desired result.  $\square$

**Proof of Theorem 6.3.** Suppose there do not exist  $i \neq j$  such that the BDP property fails for  $i$  and the NCP\* property fails for  $j$ . Then either of the following is true. Case 1: there are at least  $N-1$  agents satisfying both the BDP and NCP\* properties. Note by Lemma A.5, a special situation in this case is that all agents satisfy the BDP property. Case 2: all agents satisfy the NCP\* property.

**Case 1.** Suppose there are at least  $N-1$  agents satisfying both the BDP and NCP\* properties. By Lemma A.4, there is  $I' \subseteq I$  with  $|I'| \geq N-1$  such that for all  $i \in I'$  and  $\bar{\theta}_i \neq \hat{\theta}_i$ , there exists  $\psi^{\bar{\theta}_i, \hat{\theta}_i} : \Theta \rightarrow \mathbb{R}^n$ , such that the four requirements are satisfied.

Pick an agent  $i \in I$ , where  $\{i\} = I \setminus I'$  if  $I \setminus I'$  is a singleton, and  $i \in I$  is arbitrary if  $I \setminus I' = \emptyset$ . As in Theorem 6.1, let  $\eta$  be an interim individually rational and ex-post budget-balanced transfer rule such that agent  $i$  obtains all the surplus. Define  $\Phi = \{\eta\} \cup \{\eta + c\psi^{\bar{\theta}_j, \hat{\theta}_j} : j \in I \text{ with } j \neq i, \bar{\theta}_j, \hat{\theta}_j \in \Theta_j \text{ with } \bar{\theta}_j \neq \hat{\theta}_j\}$ , where  $c$  is sufficiently large such that for all  $j \in I$  with  $j \neq i$  and  $\bar{\theta}_j, \hat{\theta}_j \in \Theta_j$  with  $\bar{\theta}_j \neq \hat{\theta}_j$ ,

$$0 \geq \sum_{\theta_{-j} \in \Theta_{-j}} [u_j(q(\hat{\theta}_j, \theta_{-j}), \bar{\theta}_j) - u_j(q(\hat{\theta}_j, \theta_{-j}), \hat{\theta}_j) + c\psi_j^{\bar{\theta}_j, \hat{\theta}_j}(\hat{\theta}_j, \theta_{-j})] p_j(\theta_{-j} | \bar{\theta}_j).$$

For agent  $j \neq i$  with type  $\theta_j$ , truthfully reporting gives her a worst-case expected utility level of zero because her worst transfer rule,  $\eta$ , extracts all her surplus. Thus,  $j$ 's individual rationality condition binds. The choice of  $c$  implies that misreporting cannot give her a positive worst-case expected utility. Therefore, her incentive compatibility condition hold.

The type- $\bar{\theta}_i$  agent  $i$  obtains a worst-case expected utility of  $\min_{\phi \in \Phi} \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\bar{\theta}_i, \theta_{-i}), \bar{\theta}_i) + \phi_i(\bar{\theta}_i, \theta_{-i})] p_i(\theta_{-i} | \bar{\theta}_i) = \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\bar{\theta}_i, \theta_{-i}), \bar{\theta}_i) + \sum_{j \neq i} u_j(q(\bar{\theta}_i, \theta_{-i}), \theta_j)] p_i(\theta_{-i} | \bar{\theta}_i) \geq 0$  by truthfully reporting. By ex-post efficiency of  $q$ , it is weakly higher than  $\sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\hat{\theta}_i, \theta_{-i}), \bar{\theta}_i) + \sum_{j \neq i} u_j(q(\hat{\theta}_i, \theta_{-i}), \theta_j)] p_i(\theta_{-i} | \bar{\theta}_i)$  for all  $\hat{\theta}_i \neq \bar{\theta}_i$ , an upper bound of the worst-case expected payoff of misreporting  $\hat{\theta}_i$ ,  $\min_{\phi \in \Phi} \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\hat{\theta}_i, \theta_{-i}), \bar{\theta}_i) + \phi_i(\hat{\theta}_i, \theta_{-i})] p_i(\theta_{-i} | \bar{\theta}_i)$ . Individual rationality and ex-post budget balance are easy to verify. Therefore, the individually rational and budget-balanced mechanism with ambiguous transfers implements  $q$ .

**Case 2.** Suppose all agents satisfy the NCP\* property. For any  $j \in I$ , let  $\mathcal{P}_j$  be the partition of  $\Theta_j$  such that  $p_j(\cdot | \theta_j) = p_j(\cdot | \theta'_j)$  if and only if  $\theta_j, \theta'_j$  are in the same  $\tilde{\Theta}_j \in \mathcal{P}_j$ . For each  $\tilde{\Theta}_j$  with  $|\tilde{\Theta}_j| \geq 2$  and  $\theta_j \in \tilde{\Theta}_j$ , define  $U_{\theta_j}$  according to Lemma A.6. For a singleton  $\tilde{\Theta}_j \in \mathcal{P}_j$  and  $\{\theta_j\} = \tilde{\Theta}_j$ , define  $U_{\theta_j} = 0$ .

We will demonstrate that for each  $i$  and  $\bar{\theta}_i \neq \hat{\theta}_i$ , the following system has a solution  $\phi^{\bar{\theta}_i \hat{\theta}_i}$ .

$$\begin{aligned} \sum_{\theta_{-j} \in \Theta_{-j}} \phi_i^{\bar{\theta}_i \hat{\theta}_i}(\bar{\theta}_i, \theta_{-i}) p_i(\theta_{-i} | \bar{\theta}_i) &= U_{\bar{\theta}_i} - \sum_{\theta_{-i} \in \Theta_{-i}} u_i(q(\bar{\theta}_i, \theta_{-i}), \theta_i) p_i(\theta_{-i} | \bar{\theta}_i), \\ \sum_{\theta_{-j} \in \Theta_{-j}} \phi_j^{\bar{\theta}_i \hat{\theta}_i}(\theta_j, \theta_{-j}) p_j(\theta_{-j} | \theta_j) &\geq U_{\theta_j} - \sum_{\theta_{-j} \in \Theta_{-j}} u_j(q(\theta_j, \theta_{-j}), \theta_j) p_j(\theta_{-j} | \theta_j), \forall (j, \theta_j) \neq (i, \bar{\theta}_i), \\ - \sum_{j \in I} \phi_j^{\bar{\theta}_i \hat{\theta}_i}(\theta) &= 0, \forall \theta \in \Theta, \\ - \sum_{\theta_{-i} \in \Theta_{-i}} \phi_i^{\bar{\theta}_i \hat{\theta}_i}(\hat{\theta}_i, \theta_{-i}) p_i(\theta_{-i} | \bar{\theta}_i) &\geq -U_{\hat{\theta}_i} + \sum_{\theta_{-i} \in \Theta_{-i}} u_i(q(\hat{\theta}_i, \theta_{-i}), \theta_i) p_i(\theta_{-i} | \bar{\theta}_i). \end{aligned}$$

Suppose by way of contradiction that the system does not have a solution. By a theorem of the alternative, there exist coefficients  $a_{\bar{\theta}_i}, (a_{\theta_j})_{(j, \theta_j) \neq (i, \bar{\theta}_i)} \geq \mathbf{0}, (b_\theta)_{\theta \in \Theta}$ , and  $\gamma_{\bar{\theta}_i \hat{\theta}_i} \geq 0$  such that the weighted sum of left-hand sides of the expressions are cancelled and the weighted sum of right-hand sides is positive.

Suppose  $\gamma_{\bar{\theta}_i \hat{\theta}_i} = 0$ . Following the argument of Lemma A.4, we know  $(a_{\theta_j})_{j \in I, \theta_j \in \Theta_j} > \mathbf{0}$  and  $(b_\theta)_{\theta \in \Theta} \not\geq \mathbf{0}$ . Define  $\mu(\theta) = \frac{b_\theta}{\sum_{\bar{\theta} \in \Theta} b_{\bar{\theta}}}$  for all  $\theta$ , which is a common prior, a contradiction.

If  $\gamma_{\bar{\theta}_i \hat{\theta}_i} > 0$ , in view of Lemma A.4 and the fact that the NCP\* property holds for all agents, we know: (1)  $p_i(\cdot | \bar{\theta}_i) = p_i(\cdot | \hat{\theta}_i)$ , (2) among all the coefficients,  $a_{\hat{\theta}_i} = \gamma_{\bar{\theta}_i \hat{\theta}_i} > 0$

and everything else is zero. According to Lemma A.6, the choice of  $U_{\bar{\theta}_i}$  and  $U_{\hat{\theta}_i}$  satisfies  $U_{\hat{\theta}_i} - U_{\bar{\theta}_i} + \sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\hat{\theta}_i, \theta_{-i}), \bar{\theta}_i) - u_i(q(\bar{\theta}_i, \theta_{-i}), \bar{\theta}_i)] p_i(\theta_{-i} | \bar{\theta}_i) \leq 0$ . Hence, the weighted sum of the right-hand side is nonpositive, a contradiction.

Therefore, for each  $i$ ,  $\bar{\theta}_i \neq \hat{\theta}_i$ , the system has a solution  $\phi^{\bar{\theta}_i \hat{\theta}_i}$ . Let the set of ambiguous transfers be  $\Phi = \{\phi^{\bar{\theta}_i \hat{\theta}_i}, \forall i, \bar{\theta}_i, \hat{\theta}_i \in \Theta_i, \bar{\theta}_i \neq \hat{\theta}_i\}$ . The interim individually rational and ex-post budget-balanced mechanism with ambiguous transfers implements  $q$ .  $\square$

**Example A.2:** *This example shows that an inefficient allocation rule may not be implementable via an interim individually rational and ex-post budget-balanced mechanism with ambiguous transfers in a private value environment without a common prior.*

Consider the same feasible outcomes and payoffs as in Example A.1 except we do not have  $\theta_1^3$  and  $\theta_2^3$  here. The beliefs satisfy  $p_1(\theta_2^1 | \theta_1^1) = 0.6$ ,  $p_1(\theta_2^1 | \theta_1^2) = 0.4$ ,  $p_2(\theta_1^1 | \theta_2^1) = 0.75$ , and  $p_2(\theta_1^1 | \theta_2^2) = 0.25$ . Let an inefficient allocation rule  $q$  be  $q(\theta) = x_1$  if  $\theta_1 = \theta_1^1$  and  $q(\theta) = x_2$  if  $\theta_1 = \theta_1^2$ . Suppose by way of contradiction that there exists a mechanism with ambiguous transfers  $\mathcal{M} = (\Theta, q, \Phi)$  implementing  $q$ . Hence, for all  $\epsilon > 0$ , there exist transfer rules  $(\phi^k)_{k \in \{1,2\}}$  (payments from agent 1 to 2) and utility levels  $(U_{\theta_i})_{i \in N, \theta_i \in \Theta_i} \geq \mathbf{0}$ , such that

$$\begin{aligned} IR(\theta_1^1) & 0 - 0.6\phi^1(\theta_1^1, \theta_2^1) - 0.4\phi^1(\theta_1^1, \theta_2^2) \geq U_1^1, \\ IR(\theta_1^2) & 0 - 0.4\phi^1(\theta_1^2, \theta_2^1) - 0.6\phi^1(\theta_1^2, \theta_2^2) \geq U_1^2, \\ IR(\theta_2^1) & 0 + 0.75\phi^1(\theta_1^1, \theta_2^1) + 0.25\phi^1(\theta_1^2, \theta_2^1) \geq U_2^1, \\ IR(\theta_2^2) & 0 + 0.25\phi^1(\theta_1^1, \theta_2^2) + 0.75\phi^1(\theta_1^2, \theta_2^2) \geq U_2^2, \\ IC(\theta_1^1 \theta_1^2) & U_1^1 + \epsilon \geq 2 - 0.6\phi^1(\theta_1^2, \theta_2^1) - 0.4\phi^1(\theta_1^2, \theta_2^2), \end{aligned}$$

and

$$\begin{aligned} IR(\theta_1^1) & 0 - 0.6\phi^2(\theta_1^1, \theta_2^1) - 0.4\phi^2(\theta_1^1, \theta_2^2) \geq U_1^1, \\ IR(\theta_1^2) & 0 - 0.4\phi^2(\theta_1^2, \theta_2^1) - 0.6\phi^2(\theta_1^2, \theta_2^2) \geq U_1^2, \\ IR(\theta_2^1) & 0 + 0.75\phi^2(\theta_1^1, \theta_2^1) + 0.25\phi^2(\theta_1^2, \theta_2^1) \geq U_2^1, \\ IR(\theta_2^2) & 0 + 0.25\phi^2(\theta_1^1, \theta_2^2) + 0.75\phi^2(\theta_1^2, \theta_2^2) \geq U_2^2, \\ IC(\theta_1^2 \theta_1^1) & U_1^2 + \epsilon \geq 2 - 0.4\phi^2(\theta_1^1, \theta_2^1) - 0.6\phi^2(\theta_1^1, \theta_2^2). \end{aligned}$$

Multiply each inequality in the first group by 2.5, 8, 2, 4, and 4.5 respectively, multiply the second group by 8, 2.5, 4, 2, and 4.5 respectively, add them up, and let  $\epsilon$  go to zero. We obtain  $0 \geq 6U_1^1 + 6U_1^2 + 6U_2^1 + 6U_2^2 + 18 \geq 18$ , which is a contradiction. Hence,  $q$  is not implementable via ambiguous transfers.

## A.2 Including agents without private information

In this subsection, we relax the assumption that  $|\Theta_i| \geq 2$  for all  $i \in I$ . Denote the set of all agents with at least two types by  $\tilde{I}$ , which has a cardinality of  $\tilde{N}$ . An agent in  $I \setminus \tilde{I}$  has only one type and thus she cannot lie. We claim that all theorems of this paper hold if  $\tilde{N} \geq 2$ , i.e., at least two agents have private information.

To see why including agents without private information may be interesting, consider two consumers with unknown values paying for producing a costly public project. In this example  $\tilde{I} = \{1, 2\}$  and  $I = \{1, 2, 3\}$ , where 3 is interpreted as a producer whose payoff (profit) is the payments of 1 and 2 minus the cost of production. By efficiency and budget balance, two consumers' aggregated utility from the project minus the cost of production should be maximized.

We demonstrate the modification needed for Theorem 6.3 as an example. In Lemma A.4, we replace all  $I$  with  $\tilde{I}$  and all  $N$  with  $\tilde{N}$ . Let  $\eta$  be a transfer rule such that agent  $i$  obtains all the surplus of  $N$  agents, where  $\{i\} = \tilde{I} \setminus I'$  if  $\tilde{I} \setminus I'$  is a singleton and  $i \in I$  can be arbitrary if  $\tilde{I} \setminus I' = \emptyset$ . For any  $j \in \tilde{I}$  with  $j \neq i$  and  $\bar{\theta}_j \neq \hat{\theta}_j$ , let  $\psi_k^{\bar{\theta}_j \hat{\theta}_j}(\theta) = 0$  for all  $\theta$  and  $k \in I \setminus \tilde{I}$ . Then one can follow Theorem 6.3 to construct ambiguous transfers. Incentive compatibility of agents in  $\tilde{I}$  is achieved in the same way as the original proof. We obtain incentive compatibility of all other agents for free as each of them has only one type. Individual rationality and budget balance follow from the respective properties of  $\eta$  and  $\psi^{\bar{\theta}_j \hat{\theta}_j}$  for all  $j \in \tilde{I}$  with  $j \neq i$  and  $\bar{\theta}_j, \hat{\theta}_j \in \Theta_j$  with  $\bar{\theta}_j \neq \hat{\theta}_j$ .

## References

- Aoyagi, M. (1998). Correlated types and Bayesian incentive compatible mechanisms with budget balance. *Journal of Economic Theory*, 79(1):142–151.
- Aumann, R. J. (1976). Agreeing to disagree. *The Annals of Statistics*, 4(6):1236–1239.
- Barelli, P. (2009). On the genericity of full surplus extraction in mechanism design. *Journal of Economic Theory*, 144(3):1320–1332.
- Bergemann, D., Morris, S., and Takahashi, S. (2012). Efficient auctions and interdependent types. *American Economic Review: Papers and Proceedings*, 102(3):319–324.
- Bodoh-Creed, A. L. (2012). Ambiguous beliefs and mechanism design. *Games and Economic Behavior*, 75(2):518–537.

- Börger, T., Krämer, D., and Strausz, R. (2015). *An introduction to the theory of mechanism design*. Oxford University Press, USA.
- Bose, S. and Daripa, A. (2009). A dynamic mechanism and surplus extraction under ambiguity. *Journal of Economic Theory*, 144(5):2084–2114.
- Bose, S., Ozdenoren, E., and Pape, A. (2006). Optimal auctions with ambiguity. *Theoretical Economics*, 1(4):411–438.
- Bose, S. and Renou, L. (2014). Mechanism design with ambiguous communication devices. *Econometrica*, 82(5):1853–1872.
- Chen, Y.-C. and Xiong, S. (2011). The genericity of beliefs-determine-preferences models revisited. *Journal of Economic Theory*, 146(2):751–761.
- Chen, Y.-C. and Xiong, S. (2013). Genericity and robustness of full surplus extraction. *Econometrica*, 81(2):825–847.
- Chung, K.-S. (1999). A note on Matsushima’s regularity condition. *Journal of Economic Theory*, 87(2):429–433.
- Clarke, E. H. (1971). Multipart pricing of public goods. *Public Choice*, 11(1):17–33.
- Crémer, J. and McLean, R. (1985). Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. *Econometrica*, 53(2):345–361.
- Crémer, J. and McLean, R. P. (1988). Full extraction of the surplus in Bayesian and dominant strategy auctions. *Econometrica*, 56(6):1247–1257.
- Dasgupta, P. and Maskin, E. (2000). Efficient auctions. *The Quarterly Journal of Economics*, 115(2):341–388.
- d’Aspremont, C., Crémer, J., and Gérard-Varet, L.-A. (2004). Balanced Bayesian mechanisms. *Journal of Economic Theory*, 115(2):385–396.
- d’Aspremont, C. and Gérard-Varet, L.-A. (1979). Incentives and incomplete information. *Journal of Public Economics*, 11(1):25–45.
- De Castro, L. I., Liu, Z., and Yannelis, N. C. (2017a). Ambiguous implementation: the partition model. *Economic Theory*, 63(1):233–261.
- De Castro, L. I., Liu, Z., and Yannelis, N. C. (2017b). Implementation under ambiguity. *Games and Economic Behavior*, 101:20–33.

- De Castro, L. I. and Yannelis, N. C. (2009). Ambiguity aversion solves the conflict between efficiency and incentive compatibility. Working paper.
- Di Tillio, A., Kos, N., and Messner, M. (2017). The design of ambiguous mechanisms. *The Review of Economic Studies*, 84(1):237–274.
- Ghirardato, P. and Marinacci, M. (2002). Ambiguity made precise: A comparative foundation. *Journal of Economic Theory*, 102(2):251–289.
- Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153.
- Groves, T. (1973). Incentives in teams. *Econometrica*, 41(4):617–631.
- Heifetz, A. and Neeman, Z. (2006). On the generic (im) possibility of full surplus extraction in mechanism design. *Econometrica*, 74(1):213–233.
- Jehiel, P. and Moldovanu, B. (2001). Efficient design with interdependent valuations. *Econometrica*, 69(5):1237–1259.
- Klibanoff, P., Marinacci, M., and Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica*, 73(6):1849–1892.
- Kosenok, G. and Severinov, S. (2008). Individually rational, budget-balanced mechanisms and allocation of surplus. *Journal of Economic Theory*, 140(1):126–161.
- Liu, H. (2014). Efficient dynamic mechanisms in environments with interdependent valuations. Working paper.
- Matsushima, H. (1991). Incentive compatible mechanisms with full transferability. *Journal of Economic Theory*, 54(1):198–203.
- Matsushima, H. (2007). Mechanism design with side payments: Individual rationality and iterative dominance. *Journal of Economic Theory*, 133(1):1–30.
- McAfee, R. P. and Reny, P. J. (1992). Correlated information and mechanism design. *Econometrica*, 60(2):395–421.
- McLean, R. and Postlewaite, A. (2002). Informational size and incentive compatibility. *Econometrica*, 70(6):2421–2453.
- McLean, R. and Postlewaite, A. (2004). Informational size and efficient auctions. *The Review of Economic Studies*, 71(3):809–827.

- McLean, R. P. and Postlewaite, A. (2003a). Informational size and incentive compatibility with aggregate uncertainty. *Games and Economic Behavior*, 45(2):410–433.
- McLean, R. P. and Postlewaite, A. (2003b). Informational size, incentive compatibility, and the core of a game with incomplete information. *Games and Economic Behavior*, 45(1):222–241.
- McLean, R. P. and Postlewaite, A. (2015). Implementation with interdependent valuations. *Theoretical Economics*, 10(3):923–952.
- Morris, S. (1995). The common prior assumption in economic theory. *Economics and Philosophy*, 11(2):227–253.
- Myerson, R. B. and Satterthwaite, M. A. (1983). Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 29(2):265–281.
- Neeman, Z. (2004). The relevance of private information in mechanism design. *Journal of Economic Theory*, 117(1):55–77.
- Noda, S. (2015). Full surplus extraction and costless information revelation in dynamic environments. Working paper.
- Rochet, J.-C. (1987). A necessary and sufficient condition for rationalizability in a quasi-linear context. *Journal of mathematical Economics*, 16(2):191–200.
- Smith, D. (2010). A prior free efficiency comparison of mechanisms for the public good problem. Working paper.
- Song, Y. (2016). Efficient implementation with interdependent valuations and maxmin agents. Working paper.
- Sun, Y. and Yannelis, N. C. (2007). Perfect competition in asymmetric information economies: compatibility of efficiency and incentives. *Journal of Economic Theory*, 134(1):175–194.
- Sun, Y. and Yannelis, N. C. (2008). Ex ante efficiency implies incentive compatibility. *Economic Theory*, 36(1):35–55.
- Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance*, 16(1):8–37.
- Wolitzky, A. (2016). Mechanism design with maxmin agents: theory and an application to bilateral trade. *Theoretical Economics*, 11:971–1004.