Research Paper Series

No. 20

Implementability and the Axioms of Bargaining Theory

Hideki Mizukami† and Takuma Wakayama‡

March 2007

†Faculty of Economics, University of Toyama, 3190 Gofuku, Toyama, 930-8555, JAPAN; mizukami@eco.u-toyama.ac.jp
‡Faculty of Urban Liberal Arts, Tokyo Metropolitan University, 1-1 Minami-Osawa, Hachioji, Tokyo 192-0397, JAPAN; t_waka@center.tmu.ac.jp

Implementability and the Axioms of Bargaining Theory*

HIDEKI MIZUKAMI
Faculty of Economics
University of Toyama
mizukami@eco.u-toyama.ac.jp

TAKUMA WAKAYAMA†
Faculty of Urban Liberal Arts
Tokyo Metropolitan University
t_waka@center.tmu.ac.jp

March 6, 2007

Abstract

This paper considers problems where agents bargain over their shares of a divisible commodity. We explore the relationships between dominant strategy implementability and the axioms of bargaining solutions. We demonstrate that a rule inducing a bargaining solution is implementable and satisfies an auxiliary condition if and only if the bargaining solution induced by the rule satisfies efficiency, strong monotonicity, and scale invariance, provided that there are two agents. Further, we show that the three axioms mentioned above are still sufficient for a bargaining solution to be induced by an implementable rule even if there are three or more agents.

Keywords: Dominant strategy implementation, Efficiency, Scale invariance, Strong monotonicity, Welfarism.

JEL Classification Numbers: C78, D78, C71, C72, D71.

*We are grateful to Masaki Aoyagi, Atsushi Kajii, Tatsuyoshi Saijo, Tadashi Sekiguchi, Ken-ichi Shimomura, Koji Takamiya, and Naoki Yoshihara for helpful comments and suggestions. We also thank Chiaki Hara, Haruo Imai, Ryo-ichi Nagahisa, Akira Okada, Shigehiro Serizawa, and participants of the 2003 Annual Meeting of the Japanese Economic Association in Oita, the 8th International Meeting of the Society for Social Choice and Welfare in Istanbul, and the seminars at Kyoto University and Osaka University.

†**Corresponding author:** Faculty of Urban Liberal Arts, Tokyo Metropolitan University, 1-1 Minami-Osawa, Hachioji, Tokyo 192-0397, JAPAN.

1 Introduction

We consider problems where agents bargain over their shares of a divisible commodity. Examples of such problems are the share of profit between the employer and the labor union, and the distribution of property. In such situations, if all agents agree on a feasible outcome, then they can arrive at it; otherwise they arrive at the predetermined outcome or nothing. To avoid disagreements, they would decide to follow a recommendation made by an impartial arbitrator, for example, the Central Arbitration Committee presiding over labor disputes, or a judge over civil trials.

Axiomatic bargaining theory initiated by Nash (1950) theoretically deals with the situations described above. Axiomatic bargaining theory provides numerous "bargaining solutions," for instance, the Nash solution (Nash, 1950), the Kalai–Smorodinsky solution (Kalai and Smorodinsky, 1975), and the egalitarian solution (Kalai, 1977). As pointed out by Raiffa (1953), a bargaining solution, which associates a profile of welfare levels with each bargaining problem, can be interpreted as the recommendation of the arbitrator.

In the context of axiomatic bargaining theory, a bargaining problem consists of a utility possibility set and an element of it, known as the "disagreement point." Both the utility possibility set and the disagreement point are generated by the agents' (private) "types" such as agents' preferences. Since agents' types are usually unknown to the arbitrator, selfish agents may have an incentive to gain by manipulating the bargaining solution through generating a different bargaining problem by misrepresenting their types. Thus, the arbitrator should deal with the problem of incentive: how does the arbitrator make the agents voluntarily reveal their *true* types, which she is not aware of. In other words, the arbitrator faces the problem of constructing a mechanism, the equilibrium outcome of which is the outcome obtained at the true type profile. If it is possible for the arbitrator to construct such a mechanism, then the solution is said to be "implementable." Thus, it is important to focus attention on the implementability of bargaining solutions.

However, the notion of implementability is usually defined on the commodity space, whereas bargaining solutions are defined on the utility space. Hence, we cannot directly apply the notion of implementability to bargaining solutions. To overcome this difficulty, in this paper, we explicitly consider an "allocation rule," which associates an outcome with each private type profile.[1] Moreover, we introduce the following concept: roughly speaking, a bargaining solution is *induced* by an allocation rule if there exists an allocation rule such that for each type profile and each outcome chosen by the allocation rule, the outcome attains the list of welfare levels chosen by the bargaining solution. Thus, in order to examine "the implementability of bargaining solutions," this paper investigates the relationships between implementable allocation rules and bargaining solutions induced by the allocation rules.

A large number of studies have been conducted on implementability in the context of bargaining. It is well-known that some of the bargaining solutions are induced by subgame perfect implementable allocation rules. Moulin (1984) and Howard (1992) set up mechanisms that subgame perfect implement the allocation rules inducing the Kalai–Smorodinsky and Nash solutions, respectively. Miyagawa (2002) provides a "simple" mechanism that subgame perfect implements a wide class of allocation rules inducing bargaining solutions such as the Nash and Kalai–Smorodinsky solutions. Vartiainen (2006) demonstrates that there is no efficient and symmetric bargaining solution that is induced by any Nash implementable allocation rule.

However, these studies focus on a *given bargaining solution*. Thus, little is known about what types of bargaining solutions are induced by allocation rules implementable in a *given game-*

[1] Roemer (1988, 1996) first focuses on the allocation rules in the context of axiomatic bargaining theory. However, he does not consider the incentive problem in his framework.

theoretical solution concept. In this paper, we will attempt to demonstrate the class of bargaining solutions that are induced by implementable allocation rules. To accomplish this goal, we explore the relationships between implementability and the axioms of bargaining solutions.

Nash or subgame perfect implementation imposes unrealistic informational assumptions, e.g., common knowledge about each other's preferences. This signifies that mechanisms under Nash or subgame perfect implementation heavily depend on the informational assumptions, such that they are not robust. Thus, this paper concentrates on dominant strategy implementation. Dominant strategy implementability is a "robustness" property; this implies that in this case, it is not necessary to impose strong informational assumptions such as the common knowledge assumption. The robustness of mechanism has recently received attention in implementation theory (e.g., see Bergemann and Morris (2005) for details). Dominant strategy implementability is also a "practicality" property in that it is sufficient for each agent to understand her own types instead of knowing other agents' types when participating in a mechanism. Mizukami and Wakayama (2007) provide a necessary and sufficient condition for dominant strategy implementation in economic environments. Their results are applicable to our model.

The paper is organized as follows: Section 2 provides the model and defines several concepts and axioms. Section 3 presents the main results. Section 4 contains some concluding remarks. All proofs are relegated to the appendix.

2 Preliminaries

2.1 The Model

Consider n agents who are to divide a commodity that is perfectly divisible and freely disposable. Without loss of generality, we assume that the amount of the divisible commodity is equal to 1. If all the agents obey a recommendation by an impartial arbitrator, then they receive the outcome recommended by her; otherwise they get nothing.

Let $N := \{1, 2, \dots, n\}$ be the set of *agents*, where $2 \leq n < +\infty$.

Let $X := \left\{x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n_+ \colon \sum_{i \in N} x_i \leq 1\right\}$ be the set of *pure outcomes*, where agent $i \in N$ receives x_i.

Let $\triangle$ be the set of *lotteries* over X with *finite support*. Let $\ell(x)$ denote the probability that the lottery $\ell \in \triangle$ assigns to the pure outcome $x \in X$. Let $\triangle_D := \{\ell^x \in \triangle \colon x \in X\}$ be the set of *degenerate lotteries*, where ℓ^x denotes the degenerate lottery whose support consists of the single pure outcome x. For each agent $i \in N$, let ℓ_i denote her marginal distribution of the lottery $\ell \in \triangle$ over $[0, 1]$.

Let Θ_i be agent i's set of admissible *types*, each of which prescribes von Neumann–Morgenstern (henceforth abbreviated v.N–M) preferences over $\triangle$. We assume that each agent $i \in N$ is equipped with a utility function $u_i \colon X \times \Theta_i \to \mathbb{R}$. Given a lottery $\ell \in \triangle$, the expected utility of agent $i \in N$ conditional on the type $\theta_i \in \Theta_i$ is denoted by

$$U_i(\ell; \theta_i) = \sum_{x \in \mathrm{supp}(\ell)} \ell(x) u_i(x; \theta_i),$$

where $\mathrm{supp}(\ell)$ denotes the support of the lottery ℓ, i.e., $\mathrm{supp}(\ell) := \{x \in X \colon \ell(x) > 0\}$. Moreover, we assume that for each type $\theta_i \in \Theta_i$, the function $u_i(\cdot; \theta_i) \colon X \to \mathbb{R}$ is (i) continuous; (ii) strictly increasing; and (iii) strictly concave.

The *domain* is a set $\Theta := \Theta_1 \times \Theta_2 \times \cdots \times \Theta_n$. A *type profile* is a list $\theta := (\theta_1, \theta_2, \dots, \theta_n) \in \Theta$. For each lottery $\ell \in \triangle$ and each type profile $\theta \in \Theta$, let $U(\ell; \theta) := (U_1(\ell; \theta_1), U_2(\ell; \theta_2), \dots, U_n(\ell; \theta_n))$.

We often denote $N \setminus \{i\}$ by "$-i$." With this notation, (θ_i', θ_{-i}) is the type profile where agent i's type is θ_i' and the type of agent $j \in N \setminus \{i\}$ is θ_j.

For each type profile $\theta \in \Theta$, let $S(\theta) := \{U(\ell;\theta) \colon \ell \in \triangle\} \subset \mathbb{R}^n$ be a *utility possibility set*. Note that this set is strictly convex by the definition of Θ. For each type profile $\theta \in \Theta$, a *disagreement point* is the list $d(\theta) = (d_1(\theta), d_2(\theta), \ldots, d_n(\theta)) := U(\ell^{\mathbf{0}};\theta)$, where $\ell^{\mathbf{0}}$ denotes the degenerate lottery that selects $\mathbf{0} := (0, 0, \ldots, 0) \in X$. In particular, we write $d^{\mathbf{0}}(\theta)$ if $U(\ell^{\mathbf{0}};\theta) = \mathbf{0}$. Given a type profile $\theta \in \Theta$, let $\partial S(\theta) := \{s \in S(\theta) \colon$ there is no $s' \in S(\theta)$ such that $s' > s\}$.[2] Note that by definition, $d(\theta) \in S(\theta) \setminus \partial S(\theta)$ for each $\theta \in \Theta$. For each type profile $\theta \in \Theta$, a *bargaining problem* is the pair $(S(\theta), d(\theta))$. Let $\Sigma := \bigcup_{\theta \in \Theta} (S(\theta), d(\theta))$ be the class of bargaining problems.

2.2 Allocation Rules and Implementation

An "allocation rule," or briefly, a *rule* is a (possibly multi-valued) mapping $f \colon \Theta \twoheadrightarrow \triangle$ that assigns a nonempty subset of $\triangle$ to each type profile $\theta \in \Theta$.[3] If $f(\theta)$ is a singleton, then we slightly abuse notation and denote by $f(\theta)$ the single element. Following Roemer (1988), we impose two assumptions on rules:

***Essential single-valuedness*:** For each $\theta \in \Theta$, if $\ell, \ell' \in f(\theta)$, then $U(\ell;\theta) = U(\ell';\theta)$.

***Full correspondence*:** For each $\theta \in \Theta$, if $\ell \in f(\theta)$ and $U(\ell';\theta) = U(\ell;\theta)$, then $\ell' \in f(\theta)$.[4,5]

Essential single-valuedness implies that the lotteries that are chosen by a rule are welfare equivalent; in other words, all agents are indifferent to all the lotteries chosen by a rule. The full correspondence assumption states that a rule chooses a set of all lotteries that all agents are indifferent to.

We introduce additional notation and definitions about implementation. Let M_i denote a *strategy space* of agent $i \in N$. We call $m_i \in M_i$ a *strategy* of agent $i \in N$. A *mechanism* is a pair $\Gamma = (M, g)$, where $M := M_1 \times M_2 \times \cdots \times M_n$ and $g \colon M \to \triangle$ is an *outcome function*. A *strategy profile* is denoted by $m := (m_1, m_2, \ldots, m_n) \in M$.

A strategy $m_i^* \in M_i$ is a *dominant strategy* of mechanism (M, g) at $\theta_i \in \Theta_i$ if $U_i(g(m_i^*, m_{-i});\theta_i) \geq U_i(g(m_i', m_{-i});\theta_i)$ for each $m_i' \in M_i$ and each $m_{-i} \in M_{-i} := \prod_{j \neq i} M_j$. For each agent $i \in N$, let $DS_i^{\Gamma}(\theta_i) \subseteq M_i$ be the set of her dominant strategies of mechanism Γ at $\theta_i \in \Theta_i$.

A strategy profile $m^* = (m_1^*, m_2^*, \ldots, m_n^*) \in M$ is a *dominant strategy equilibrium* of mechanism (M, g) at $\theta \in \Theta$ if, for each agent $i \in N$, $U_i(g(m_i^*, m_{-i});\theta_i) \geq U_i(g(m_i', m_{-i});\theta_i)$ for each $m_i' \in M_i$ and each $m_{-i} \in M_{-i}$. Let $DSE^{\Gamma}(\theta) \subseteq M$ be the set of dominant strategy equilibria of mechanism Γ at $\theta \in \Theta$. Let $g(DSE^{\Gamma}(\theta)) := \left\{\ell \in \triangle \colon \ell = g(m) \text{ for some } m \in DSE^{\Gamma}(\theta)\right\}$ be the set of *dominant strategy equilibrium outcomes* of mechanism Γ at $\theta \in \Theta$. A mechanism $\Gamma = (M, g)$ *dominant strategy implements* a rule f if $g(DSE^{\Gamma}(\theta)) = f(\theta)$ for each $\theta \in \Theta$. Dominant strategy implementability states that there exists a mechanism whose equilibrium outcome coincides with the outcome of a given rule for every type profile.

***Dominant strategy implementability*:** There exists a mechanism $\Gamma = (M, g)$ such that $g(DSE^{\Gamma}(\theta)) = f(\theta)$ for each $\theta \in \Theta$.

---

[2] Vector inequalities are denoted as follows: given $x, x' \in \mathbb{R}^n$, $x \geqq x'$ means $x_i \geq x_i'$ for each $i \in N$; $x \geq x'$ means $x \geqq x'$ and $x \neq x'$; $x > x'$ means $x_i > x_i'$ for each $i \in N$.

[3] The arrow $\twoheadrightarrow$ is used for a multi-valued mapping.

[4] Many results obtained in this paper depend strongly on the full correspondence assumption. Indeed, if the full correspondence assumption is dropped, many rules appear (See Barberà et al. (1998) for details).

[5] A rule that satisfies the full correspondence assumption is often said to be "Pareto indifferent."

$$
\begin{tikzcd}
\Theta \arrow[r] \arrow[drr, "f"'] & \Sigma \arrow[rrr, "F"] & & & \mathbb{R}^n \\
& & \triangle \arrow[urr, "U"'] & &
\end{tikzcd}
$$

Figure 1: The commutative diagram.

2.3 Bargaining Solutions and Axioms

A *bargaining solution* is a single-valued mapping $F\colon \Sigma \to \mathbb{R}^n$, which assigns to each bargaining problem $(S,d) \in \Sigma$ a list of welfare levels in $S \subset \mathbb{R}^n$. We denote by $F_i(S,d)$ the welfare level assigned to agent $i \in N$.

Remark 1. Given a bargaining solution F and type profiles $\theta, \theta' \in \Theta$, if $(S(\theta), d(\theta)) = (S(\theta'), d(\theta'))$, then $F(S(\theta), d(\theta)) = F(S(\theta'), d(\theta'))$. Thus, we can write $F(S,d)$ for $F(S(\theta), d(\theta))$ for each $\theta \in \Theta$.

Since the notion of implementability is usually defined on the commodity (or lottery) space, we cannot directly apply the notion of implementability to bargaining solutions. Thus, in order to make a connection between rules and bargaining solutions, we introduce the following concept: a bargaining solution F is *induced* by a rule f if for each $\theta \in \Theta$ and each $\ell \in f(\theta)$,

$$F(S(\theta), d(\theta)) = U(\ell; \theta).$$

The relationship is illustrated in Figure 1.

Remark 2. For each bargaining solution, there exists a rule inducing it. However, there exists a rule that does not induce any bargaining solution. For example, some constant rules do not induce any bargaining solution as we prove in the appendix.

Our primary objective is to explore a relationship between dominant strategy implementability and the axioms of bargaining solutions. We now provide several axioms. The first axiom is a standard one for bargaining solutions, namely, *efficiency*; it requires that there should be no feasible point at which all the agents are better off.

***Efficiency*:** For each $(S,d) \in \Sigma$, $F(S,d) \in \partial S$.

The second axiom—*scale invariance*—reflects the fact that v.N–M utility functions are unique up to positive affine transformations. Scale invariance entails that the bargaining solution should yield the same list of welfare levels if two bargaining problems are generated by the type profiles that prescribe the same v.N–M *preference* profile. To give the formal definition of scale invariance, we need some notation. For each agent $i \in N$, $\tau_i\colon \mathbb{R} \to \mathbb{R}$ is an *affine transformation* of agent $i \in N$ if $\tau_i\colon t \mapsto at + b$ for each $t \in \mathbb{R}$, where $a \in \mathbb{R}_{++}$ and $b \in \mathbb{R}$. Let T_i be the set of all affine transformations of agent $i \in N$. Let $T := T_1 \times T_2 \times \cdots \times T_n$. Given $\tau \in T$ and $(S,d) \in \Sigma$, let $\tau(S) := \{s' \in \mathbb{R}^n \colon s' = (\tau_1(s_1), \tau_2(s_2), \ldots, \tau_n(s_n))$ for some $s \in S\}$ and $\tau(S,d) := (\tau(S), \tau(d))$.

***Scale invariance*:** For each $(S,d),(S',d') \in \Sigma$ and each $\tau \in T$, if $(S',d') = \tau(S,d)$, then $F(S',d') = \tau(F(S,d))$.

The last axiom—*strong monotonicity*—is an appealing one from a normative viewpoint; it entails that all of the agents should benefit from *any* expansion of opportunities.

***Strong monotonicity*:** For each $(S,d),(S',d) \in \Sigma$, if $S \subseteq S'$, then $F(S,d) \leqq F(S',d)$.

3 Main Results

In this section, we explore a relationship between dominant strategy implementability and the axioms of bargaining solutions. We first consider the two-agent case. Next, we discuss the case in which there are three or more agents. All proofs of the results can be found in the appendix.

3.1 The Two-agent Case

In this subsection, we consider the two-agent case. Theorem 1 given below provides a necessary and sufficient condition for a bargaining solution to be induced by a dominant strategy implementable rule satisfying an auxiliary condition, provided there are only two agents.

Theorem 1. *Suppose that $n = 2$. A rule f inducing a bargaining solution F is dominant strategy implementable and satisfies the non-disagreement condition if and only if F induced by f satisfies efficiency, scale invariance, and strong monotonicity.*

Theorem 1 argues that from a dominant strategy implementability viewpoint, these axioms described above—efficiency, scale invariance, and strong monotonicity—are desirable.

A few remarks should be made at this point. First, it should be noted that dominant strategy implementability alone does not imply efficiency. For example, a rule that always chooses the degenerate lottery $\ell^{\mathbf{0}}$ is dominant strategy implementable, but the bargaining solution induced by the rule violates efficiency. However, such a rule would not be desirable for agents. Therefore, we impose the following "plausible" condition on rules.

***The non-disagreement condition*:** There exists no $\theta \in \Theta$ such that $f(\theta) \neq \{\ell^{\mathbf{0}}\}$.

The non-disagreement condition is first introduced by this paper, and so it is a new concept. This condition requires that a rule should not choose only the disagreement lottery. This condition is "plausible" in the sense that as mentioned in the introduction, agents would adhere to a recommendation of the arbitrator in order to avoid disagreements. Note that if a bargaining solution satisfies efficiency, then a rule inducing the bargaining solution satisfies the non-disagreement condition. This fact holds not only for the two-agent case but also in the case of three or more agents.

Second, we show Theorem 1 by means of Roemer's (1996) result. Before stating his result, let us introduce the following solution—*dictatorial solution*—that favors one agent at the expense of others.

Dictatorial solutions (for $n = 2$), D^i: There exists an agent $i \in N$ such that for each bargaining problem $(S,d) \in \Sigma$, $D^i(S,d)$ is the maximal point s of S with $s_j = d_j$. That is, there exists $i \in N$ such that for each $(S,d) \in \Sigma$, $D^i_i(S,d) = a_i(S,d)$ and $D^i_j(S,d) = d_j$, where $a_i(S,d) := \max\{s_i \colon s \in S\}$.

Theorem 2 (Roemer, 1996; Theorem 2.7). *Suppose that $n = 2$. A bargaining solution F satisfies efficiency, scale invariance, and strong monotonicity if and only if it is dictatorial.*

For the original proof of Theorem 2, see Roemer (1996). We give a simple alternative proof in the appendix.

Third, it is meaningful to study what types of bargaining solutions are induced by dominant strategy implementable rules satisfying the non-disagreement condition. The answer to this question follows from Theorems 1 and 2.

Corollary 1. *Suppose that $n = 2$. A rule f inducing a bargaining solution F is dominant strategy implementable and satisfies the non-disagreement condition if and only if F induced by f is dictatorial.*

From Corollary 1, there are few bargaining solutions *induced* by dominant strategy implementable rules satisfying the non-disagreement condition. In contrast, as we prove in the appendix, there are many dominant strategy implementable rules satisfying the non-disagreement condition. Indeed, any constant rules, each of which always assigns only an efficient lottery, are dominant strategy implementable and satisfy the non-disagreement condition. However, almost all constant rules violate *welfarism*, which requires that if two type profiles give rise to the same bargaining problem, then rules should assign to each of the type profiles lotteries that are indistinguishable in terms of utility across the type profiles.[6] Welfarism is a necessary condition for a rule to induce a bargaining solution. We can say that welfarism is so strong in the light of implementability.

3.2 The Case of Three or More Agents

In this subsection, we discuss the case where there are three or more agents. We apply Roemer (1996) to obtain the results in the previous subsection. However, we cannot apply Roemer (1996) to the case of three or more agents because his characterization holds only for the two-agent case. Thus, the results of the two-agent case do not easily extend to the case where there are more than two agents. Indeed, in axiomatic bargaining theory as well as implementation theory, the results of the two-agent case often differ from those of the case of three or more agents.

According Theorem 3 stated below, even if there are more than two agents, efficiency, scale invariance, and strong monotonicity are still sufficient for a bargaining solution to be induced by a dominant strategy implementable rule.

Theorem 3. *If a bargaining solution F satisfies efficiency, scale invariance, and strong monotonicity, then a rule f inducing F is dominant strategy implementable.*

Theorem 4 states that efficiency is still necessary for a bargaining solution to be induced by a dominant strategy implementable rule satisfying the non-disagreement condition.

Theorem 4. *If a rule f satisfying the non-disagreement condition is dominant strategy implementable, then a bargaining solution F induced by f satisfies efficiency.*

From the above results, it can be inferred that Theorem 1 also holds for the case of more than two agents. However, Theorem 1 does not hold for the case where there are three or more agents. Thus, there is a clear gap between the case of two agents and that of more than two agents. Indeed, there exists a bargaining solution induced by a dominant strategy implementable rule satisfying the non-disagreement condition that violates strong monotonicity. The following example illustrates this bargaining solution.

[6] Welfarism is first introduced by Roemer (1988) in the context of axiomatic bargaining theory. See Footnote 14 for the formal description of welfarism.

Example 1. Let $N = \{1,2,3\}$. Let $\ell \in \triangle$ be such that $\ell((1,0,0)) = .5$ and $\ell(\mathbf{0}) = .5$. Let $\ell' \in \triangle$ be such that $\ell'((.3,0,0)) = 1$. Consider the following rule: for each $\theta \in \Theta$,

$$f^B(\theta) = \begin{cases} \ell^{(0,1,0)} & \text{if } U_1(\ell;\theta_1) \geq U_1(\ell';\theta_1) \\ \ell^{(0,0,1)} & \text{otherwise.} \end{cases}$$

Note that the rule f^B is dominant strategy implementable and satisfies the non-disagreement condition.[7] The bargaining solution F^B induced by the rule f^B is as follows. Let $(S,d) \in \Sigma$. Then, there exists $\lambda = (\lambda_1, \lambda_2, \lambda_3) \in \prod_{i \in N} \Lambda_i$ such that

$$S = \lambda(\text{co}\{(a_1(S,d), d_2, d_3), (d_1, a_2(S,d), d_3), (d_1, d_2, a_3(S,d))\}),$$

and $d = \lambda(d)$, where λ_i is a monotone transformation of agent $i \in N$.[8,9] Then,

$$F^B(S,d) = \begin{cases} (d_1, a_2(S,d), d_3) & \text{if } .5 \cdot a_1(S,d) + .5 \cdot d_1 \geq \lambda_1(.3 \cdot a_1(S,d) + .7 \cdot d_1) \\ (d_1, d_2, a_3(S,d)) & \text{otherwise.} \end{cases}$$

The bargaining solution F^B violates strong monotonicity (See Figure 2). This means that Theorem 1 does not hold for the case where there are three or more agents. ■

Efficiency plays a central role in the axiomatic bargaining theory, as is evident from the fact that there are many studies using efficiency. Moreover, scale invariance is indispensable in our setting. Therefore, a further direction on this study would be to investigate whether Theorem 1 holds if strong monotonicity is relaxed with another axiom in the case where there are three or more agents.

Remark 3. It would be evident from the proofs provided in the appendix that our results are not entirely dependant on the full correspondence assumption. The full correspondence assumption can be replaced by the following condition: for each $\theta \in \Theta$, if $\ell \in f(\theta)$ and $U(\ell^x;\theta) = U(\ell;\theta)$, then $\ell^x \in f(\theta)$. According to this condition, a rule should choose the degenerate lotteries that are indifferent to the one that the rule chooses for all the agents. This condition is different from the full correspondence assumption in that it is not necessary to choose *all* the lotteries that are indifferent to the one chosen by the rule for all the agents.

4 Conclusion

This paper explored a relationship between dominant strategy implementability and the axioms of bargaining solutions. It has been shown that efficiency, scale invariance, and strong monotonicity are necessary and sufficient for a bargaining solution to be induced by a dominant strategy implementable allocation rule satisfying the non-disagreement condition whenever there are two agents. From this, it follows that there are only dictatorial solutions that are induced by dominant strategy implementable allocation rules satisfying the non-disagreement condition in the two-agent case. Next, we demonstrated that these axioms are still sufficient for a bargaining solution to be induced by a dominant strategy implementable allocation rule even when there are three or more agents.

[7]Since f^B satisfies single-valuedness, ordinality, and Property λ, it is dominant strategy implementable. See Lemma 2 in the appendix for details.

[8]Given a set $S \subseteq \mathbb{R}^n$, co$S$ is the convex hull of S.

[9]Given a set $S \subseteq \mathbb{R}^n$, let $\lambda(S) := \{s' \in \mathbb{R}^n \colon s' = \lambda(s) = (\lambda_1(s_1), \lambda_2(s_2), \ldots, \lambda_n(s_n)) \text{ for some } s \in S\}$. See the appendix for the formal definition of monotone transformations.

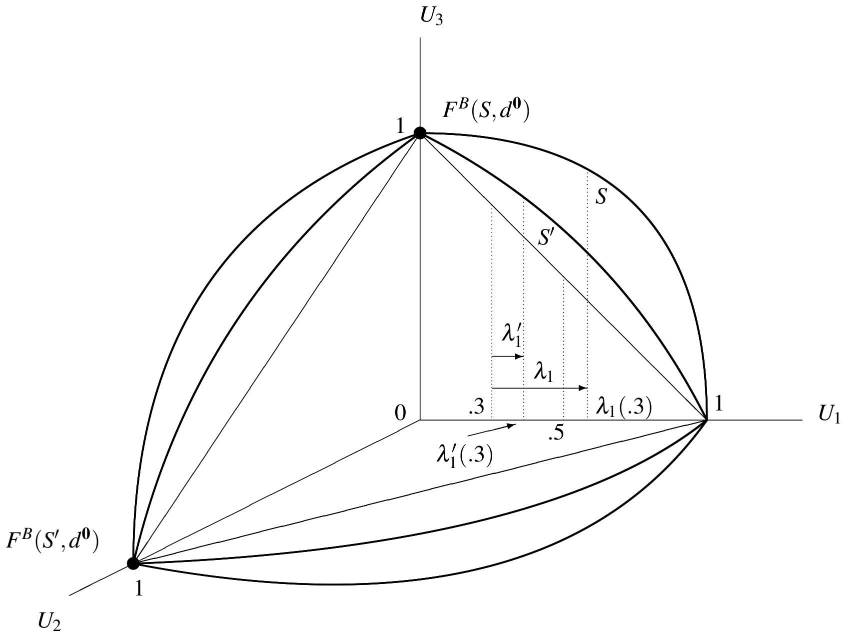

Figure 2: The bargaining solution F^B in Example 1. Let $(S,d^{\mathbf{0}}),(S',d^{\mathbf{0}}) \in \Sigma$ be such that $S' \subset S$, $a_i(S,d^{\mathbf{0}}) = 1$, and $a_i(S',d^{\mathbf{0}}) = 1$ for each $i \in N$. Then, there exist $\lambda, \lambda' \in \prod_{i \in N} \Lambda_i$ such that $S = \lambda(\text{co}\{(1,0,0),(0,1,0),(0,0,1)\})$ and $S' = \lambda'(\text{co}\{(1,0,0),(0,1,0),(0,0,1)\})$. Since $\lambda_1(.3) > .5$, $F^B(S,d^{\mathbf{0}}) = (0,0,1)$. Since $.5 > \lambda_1'(.3)$, $F^B(S',d^{\mathbf{0}}) = (0,1,0)$.

We discuss three remaining problems. First, this paper considered a *single-valued* bargaining solution and thereby derived an impossibility result. One way to avoid our impossibility result can be to consider a *multi-valued* bargaining solution. For instance, considering multi-valued bargaining solutions, Vartiainen (2006) derived possibility results for Nash implementation.

Second, although we provided a complete characterization of bargaining solutions that are induced by dominant strategy implementable allocation rules satisfying the non-disagreement condition in the two-agent case, we only provided a sufficient condition for a bargaining solution to be induced by a dominant strategy implementable allocation rule in the case of three or more agents. As we observed in Example 1, strong monotonicity is not necessary for a bargaining solution to be induced by a dominant strategy implementable allocation rule satisfying the non-disagreement condition. An interesting direction for future research would be to investigate the necessary conditions for a bargaining solution to be induced by a dominant strategy implementable allocation rule.

Finally, in this paper, we restricted our attention to the class of essentially single-valued allocation rules satisfying the full correspondence assumption. What are the types of bargaining solutions that are induced by dominant strategy implementable allocation rules that violate either essential single-valuedness or the full correspondence assumption? This issue is left to be addressed in future research.

A Appendix: Proofs

A.1 Preliminary Results

This section introduces two lemmas that are useful in the proofs of theorems. Given $i \in N$, $\theta_i \in \Theta_i$, and $x \in X$, let $A_i(x;\theta_i) := \{\ell \in \triangle \colon U_i(\ell;\theta_i) \geq U_i(\ell^x;\theta_i)\}$. That is, $A_i(x;\theta_i)$ is the set of lotteries that agent i with type θ_i likes at least as well as ℓ^x. A type $\theta_i' \in \Theta_i$ of agent $i \in N$ is *strictly more risk averse* than $\theta_i \in \Theta_i$ if $A_i(x;\theta_i) \supset A_i(x;\theta_i')$ for each $x \in X$ with $x_i \in (0,1)$. Given $i \in N$, let $e^i \in X$ be a pure outcome such that $e^i_i = 1$. Let $\triangle_P := \left\{\ell^x \in \triangle_D \colon \sum_{i \in N} x_i = 1\right\}$ be the set of *efficient lotteries*.

Lemma 1. *Let $\theta \in \Theta$, $i \in N$, and $\bar{\theta}_i \in \Theta_i$ be such that θ_i is strictly more risk averse than $\bar{\theta}_i$ and $f(\theta)$ is not a singleton. Let $\ell^x \in f(\theta) \cap \triangle_D$. Then, there exists $\tilde{\ell} \in f(\theta)$ such that $U_i(\tilde{\ell};\bar{\theta}_i) > U_i(\ell^x;\bar{\theta}_i)$.*

Proof: Let $\theta \in \Theta$, $i \in N$, and $\bar{\theta}_i \in \Theta_i$ be such that $f(\theta)$ is not a singleton and θ_i is strictly more risk averse than $\bar{\theta}_i$, i.e., $A_i(x';\bar{\theta}_i) \supset A_i(x';\theta_i)$ for each $x' \in X$ with $x'_i \in (0,1)$. Since $f(\theta)$ is not a singleton, by the full correspondence assumption and essential single-valuedness, $\ell^x \in f(\theta) \cap \left(\triangle_D \setminus \left(\triangle_P \cup \{\ell^{\mathbf{0}}\}\right)\right)$. Thus, $x_i \neq 1$. Note that $f(\theta) \cap \triangle_D$ is a singleton by the full correspondence assumption and essential single-valuedness.

We now prove the following claim:

Claim. $U_i(\hat{\ell};\bar{\theta}_i) \geq U_i(\ell^x;\bar{\theta}_i)$ for each $\hat{\ell} \in f(\theta)$.

Since θ_i is strictly more risk averse than $\bar{\theta}_i$ and $x_i \neq 1$, by definition,

$$\left\{\ell \in \triangle \colon U_i(\ell;\bar{\theta}_i) \geq U_i(\ell^x;\bar{\theta}_i)\right\} \supset \{\ell \in \triangle \colon U_i(\ell;\theta_i) \geq U_i(\ell^x;\theta_i)\} \text{ if } x_i \in (0,1); \text{ and}$$
$$\left\{\ell \in \triangle \colon U_i(\ell;\bar{\theta}_i) \geq U_i(\ell^x;\bar{\theta}_i)\right\} = \{\ell \in \triangle \colon U_i(\ell;\theta_i) \geq U_i(\ell^x;\theta_i)\} \text{ if } x_i = 0,$$

which imply that

$$\left\{\ell \in \triangle \colon U_i(\ell;\bar{\theta}_i) \geq U_i(\ell^x;\bar{\theta}_i)\right\} \supseteq \{\ell \in \triangle \colon U_i(\ell;\theta_i) \geq U_i(\ell^x;\theta_i)\}. \tag{1}$$

Note that by the full correspondence assumption and essential single-valuedness, we have $f(\theta) = \{\ell \in \triangle \colon U(\ell;\theta) = U(\ell^x;\theta)\}$. Thus, since $x_i \neq 1$,

$$\{\ell \in \triangle \colon U_i(\ell;\theta_i) \geq U_i(\ell^x;\theta_i)\} \supset \{\ell \in \triangle \colon U(\ell;\theta) = U(\ell^x;\theta)\} = f(\theta). \tag{2}$$

(1) and (2) together imply $\left\{\ell \in \triangle \colon U_i(\ell;\bar{\theta}_i) \geq U_i(\ell^x;\bar{\theta}_i)\right\} \supset f(\theta)$. This establishes that $U_i(\hat{\ell};\bar{\theta}_i) \geq U_i(\ell^x;\bar{\theta}_i)$ for each $\hat{\ell} \in f(\theta)$.

To complete the proof, suppose, by contradiction, that $U_i(\ell;\bar{\theta}_i) \leq U_i(\ell^x;\bar{\theta}_i)$ for each $\ell \in f(\theta)$. Then, by the claim,

$$U_i(\ell;\bar{\theta}_i) = U_i(\ell^x;\bar{\theta}_i) \tag{3}$$

for each $\ell \in f(\theta)$. Since θ_i is strictly more risk averse than $\bar{\theta}_i$, there exists an increasing and strictly concave function $k \colon \mathbb{R} \to \mathbb{R}$ such that $u_i(x';\theta_i) = k(u_i(x';\bar{\theta}_i))$ for each $x' \in X$ by Theorem 4 in Roth (1979). Let $\ell \in f(\theta) \setminus \{\ell^x\}$. Since k is strictly concave, by Jensen's inequality and (3), $U_i(\ell^x;\theta_i) = k\left(U_i(\ell^x;\bar{\theta}_i)\right) = k\left(U_i(\ell;\bar{\theta}_i)\right) = k\left(\sum_{\tilde{x} \in \mathrm{supp}(\ell)} \ell(\tilde{x}) u_i(\tilde{x};\bar{\theta}_i)\right) > \sum_{\tilde{x} \in \mathrm{supp}(\ell)} \ell(\tilde{x}) k\left(u_i(\tilde{x};\bar{\theta}_i)\right) = U_i(\ell;\theta_i)$, contradicting essential single-valuedness. $\square$

We now introduce a few properties that do not appear in the main text.

Ordinality: For each $\theta, \theta' \in \Theta$, whenever $f(\theta) \neq f(\theta')$, there exist $i \in N$ and $\ell, \ell' \in \triangle$ such that $U_i(\ell;\theta_i) \geq U_i(\ell';\theta_i)$ and $U_i(\ell';\theta_i') > U_i(\ell;\theta_i')$.

For each agent $i \in N$, $\lambda_i \colon \mathbb{R} \to \mathbb{R}$ is a *monotone transformation* of agent $i \in N$ if λ_i is a continuous and increasing function. Let Λ_i be the set of all monotone transformations of agent $i \in N$. Given a type profile $\theta \in \Theta$, let $\lambda_i(S(\theta)) := \{s' \in \mathbb{R}^n \colon s' = (\lambda_i(s_i), s_{-i}) \text{ for some } s \in S(\theta)\}$ and $\lambda_i(S(\theta), d(\theta)) := (\lambda_i(S(\theta)), \lambda_i(d(\theta)))$.

Remark 4. Since any distinct types of agent $i \in N$ induce the same ordinal ordering on X, we obtain the following fact: for each $\theta_i, \theta_i' \in \Theta_i$, there exists $\lambda_i \in \Lambda_i$ such that $u_i(x;\theta_i') = \lambda_i(u_i(x;\theta_i))$ for each $x \in X$.

Property $\boldsymbol{\lambda}$**:** For each $\theta \in \Theta$, each $i \in N$, each $\theta_i' \in \Theta_i$, and each $\lambda_i \in \Lambda_i$, if $(S(\theta_i', \theta_{-i}), d(\theta_i', \theta_{-i})) = \lambda_i(S(\theta), d(\theta))$, then $U_i(\ell';\theta_i') = \lambda_i(U_i(\ell;\theta_i))$ for each $\ell \in f(\theta)$ and each $\ell' \in f(\theta_i', \theta_{-i})$.

Lemma 2 below characterizes the set of dominant strategy implementable rules in terms of single-valuedness, ordinality, and Property λ. Lemma 2 also says that *none* of the multi-valued rules are dominant strategy implementable.

Lemma 2. *A rule f is dominant strategy implementable if and only if it satisfies single-valuedness, ordinality, and Property λ.*

Remark 5. If a rule f satisfies the full correspondence assumption and single-valuedness, then $f(\theta) \in \triangle_P \cup \{\ell^{\mathbf{0}}\}$ for each $\theta \in \Theta$.

Proof of Lemma 2: We first show the "only if" part. Let f be a dominant strategy implementable rule. Then, there exists a mechanism $\Gamma = (M, g)$ such that $g(DSE^{\Gamma}(\theta)) = f(\theta)$ for each $\theta \in \Theta$.

Step 1: f satisfies single-valuedness. Suppose, by contradiction, that f is not a single-valued function. Then $f(\theta)$ is not a singleton for some $\theta \in \Theta$. Since f satisfies the full correspondence assumption, $f(\theta)$ must contain a degenerate lottery, which we denote by ℓ^x.

Let $i \in N$ and $\bar{\theta}_i \in \Theta_i$ be such that θ_i is strictly more risk averse than $\bar{\theta}_i$. We denote by $\ell^{\bar{x}}$ a degenerate lottery contained in $f(\bar{\theta}_i, \theta_{-i})$.

Since f is dominant strategy implementable by Γ,

$$g(DSE^{\Gamma}(\theta)) = f(\theta), \tag{4}$$

$$g(DSE^{\Gamma}(\bar{\theta}_i, \theta_{-i})) = f(\bar{\theta}_i, \theta_{-i}). \tag{5}$$

If $m_i^* \in DS_i^{\Gamma}(\theta_i)$, then, by definition, $U_i(g(m_i^*, m_{-i});\theta_i) \geq U_i(g(m_i', m_{-i});\theta_i)$ for each $m_i' \in M_i$ and each $m_{-i} \in M_{-i}$. This implies that $U_i(g(m_i^*, m_{-i}^*);\theta_i) \geq U_i(g(\bar{m}_i, m_{-i}^*);\theta_i)$ for each $\bar{m}_i \in DS_i^{\Gamma}(\bar{\theta}_i)$ and each $m_{-i}^* \in \prod_{j \neq i} DS_j^{\Gamma}(\theta_j)$. Thus, $U_i(g(m^*);\theta_i) \geq U_i(g(\bar{m});\theta_i)$ for each $m^* \in DSE^{\Gamma}(\theta)$ and each $\bar{m} \in DSE^{\Gamma}(\bar{\theta}_i, \theta_{-i})$. Combined with (4) and (5), this implies that $U_i(\ell;\theta_i) \geq U_i(\bar{\ell};\theta_i)$ for each $\ell \in f(\theta)$ and each $\bar{\ell} \in f(\bar{\theta}_i, \theta_{-i})$. Hence, this implies that $U_i(\ell^x;\theta_i) \geq U_i(\ell^{\bar{x}};\theta_i)$, which implies that

$$x_i \geq \bar{x}_i \tag{6}$$

by the strict increasingness of $u_i(\cdot;\theta_i)$.

Similarly, if $\bar{m}_i^* \in DS_i^\Gamma(\bar{\theta}_i)$, then $U_i(g(\bar{m}_i^*, m_{-i}); \bar{\theta}_i) \geq U_i(g(m_i', m_{-i}); \bar{\theta}_i)$ for each $m_i' \in M_i$ and each $m_{-i} \in M_{-i}$. This implies that $U_i(g(\bar{m}_i^*, \bar{m}_{-i}^*); \bar{\theta}_i) \geq U_i(g(\hat{m}_i, \bar{m}_{-i}^*); \bar{\theta}_i)$ for each $\hat{m}_i \in DS_i^\Gamma(\theta_i)$ and each $\bar{m}_{-i}^* \in \prod_{j \neq i} DS_j^\Gamma(\theta_j)$. Thus, $U_i(g(\bar{m}^*); \bar{\theta}_i) \geq U_i(g(\hat{m}); \bar{\theta}_i)$ for each $\bar{m}^* \in DSE^\Gamma(\bar{\theta}_i, \theta_{-i})$ and each $\hat{m} \in DSE^\Gamma(\theta)$. Combined with (4) and (5), this implies that

$$U_i(\bar{\ell}; \bar{\theta}_i) \geq U_i(\ell; \bar{\theta}_i) \tag{7}$$

for each $\bar{\ell} \in f(\bar{\theta}_i, \theta_{-i})$ and each $\ell \in f(\theta)$.

Since $f(\theta)$ is not a singleton, by Lemma 1,

$$U_i(\tilde{\ell}; \bar{\theta}_i) > U_i(\ell^x; \bar{\theta}_i) \tag{8}$$

for some $\tilde{\ell} \in f(\theta)$. Since $\tilde{\ell} \in f(\theta)$, by (7),

$$U_i(\bar{\ell}; \bar{\theta}_i) \geq U_i(\tilde{\ell}; \bar{\theta}_i) \tag{9}$$

for each $\bar{\ell} \in f(\bar{\theta}_i, \theta_{-i})$. (8) and (9) together imply that $U_i(\bar{\ell}; \bar{\theta}_i) > U_i(\ell^x; \bar{\theta}_i)$ for each $\bar{\ell} \in f(\bar{\theta}_i, \theta_{-i})$. Since $\ell^{\bar{x}} \in f(\bar{\theta}_i, \theta_{-i})$, $U_i(\ell^{\bar{x}}; \bar{\theta}_i) > U_i(\ell^x; \bar{\theta}_i)$. Hence, the strict increasingness of $u_i(\cdot; \bar{\theta}_i)$ implies that $\bar{x}_i > x_i$, which contradicts (6).

Step 2: f satisfies ordinality. Suppose, by contradiction, that f violates ordinality. Then, there exist $\theta, \theta' \in \Theta$ such that $f(\theta) \neq f(\theta')$, but there are no $i \in N$ and $\ell, \ell' \in \triangle$ such that $U_i(\ell; \theta_i) \geq U_i(\ell'; \theta_i)$ and $U_i(\ell'; \theta_i') > U_i(\ell; \theta_i')$. This means that θ and θ' yield the same v.N–M preference profile. Therefore, $DSE^\Gamma(\theta) = DSE^\Gamma(\theta')$. By dominant strategy implementability, $f(\theta') = g(DSE^\Gamma(\theta'))$. Hence, $f(\theta) \neq f(\theta') = g(DSE^\Gamma(\theta')) = g(DSE^\Gamma(\theta))$, which is a contradiction to dominant strategy implementability.

Step 3: f satisfies quasi-constancy.[10] Suppose, by contradiction, that for some $\theta \in \Theta$, some $i \in N$, and some $\theta_i' \in \Theta_i$, there exist $\ell \in f(\theta)$ and $\ell' \in f(\theta_i', \theta_{-i})$ such that $\ell_i \neq \ell_i'$. By Step 1, $f(\theta)$ and $f(\theta_i', \theta_{-i})$ are singletons. Note that $f(\theta), f(\theta_i', \theta_{-i}) \in \triangle_P \cup \{\ell^{\mathbf{0}}\} \subset \triangle_D$ by Remark 5. Thus, we obtain $f_i(\theta) \neq f_i(\theta_i', \theta_{-i})$. Since $u_i(\cdot; \theta_i)$ and $u_i(\cdot; \theta_i')$ are strict increasing functions, this implies that

$$U_i(f(\theta); \theta_i) \neq U_i(f(\theta_i', \theta_{-i}); \theta_i), \tag{10}$$

$$U_i(f(\theta); \theta_i') \neq U_i(f(\theta_i', \theta_{-i}); \theta_i'). \tag{11}$$

So, since $u_i(\cdot; \theta_i)$ and $u_i(\cdot; \theta_i')$ are strictly increasing functions, by (10) and (11), $U_i(f(\theta); \theta_i) > U_i(f(\theta_i', \theta_{-i}); \theta_i)$ implies $U_i(f(\theta); \theta_i') > U_i(f(\theta_i', \theta_{-i}); \theta_i')$, or $U_i(f(\theta); \theta_i) < U_i(f(\theta_i', \theta_{-i}); \theta_i)$ implies $U_i(f(\theta); \theta_i') < U_i(f(\theta_i', \theta_{-i}); \theta_i')$. These imply that f is not strategy-proof.[11] Hence, by the revelation principle, it is a contradiction to dominant strategy implementability.

Step 4: f satisfies Property λ. Suppose that there exist $\theta \in \Theta$, $i \in N$, $\theta_i' \in \Theta_i$, and $\lambda_i \in \Lambda_i$ such that $(S(\theta_i', \theta_{-i}), d(\theta_i', \theta_{-i})) = \lambda_i(S(\theta), d(\theta))$, and $U_i(\ell'; \theta_i') \neq \lambda_i(U_i(\ell; \theta_i))$ for some $\ell \in f(\theta)$ and some $\ell' \in f(\theta_i', \theta_{-i})$. Since f is single-valued, by the full correspondence assumption, $\ell, \ell' \in \triangle_D$. Since $U_i(\ell'; \theta_i') \neq \lambda_i(U_i(\ell; \theta_i))$, we have $U_i(\ell'; \theta_i') \neq \lambda_i(U_i(\ell; \theta_i)) = U_i(\ell; \theta_i')$. This means that $\ell_i \neq \ell_i'$, which is a contradiction to quasi-constancy.

[10] A rule f satisfies *quasi-constancy* if for each $\theta \in \Theta$, each $i \in N$, each $\theta_i' \in \Theta_i$, each $\ell \in f(\theta)$, and each $\ell' \in f(\theta_i', \theta_{-i})$, we have $\ell_i = \ell_i'$.

[11] A single-valued rule f satisfies *strategy-proofness* if for each $\theta \in \Theta$, each $i \in N$, and $\theta_i' \in \Theta_i$, $U_i(f(\theta); \theta_i) \geq U_i(f(\theta_i', \theta_{-i}); \theta_i)$.

We next show the "if" part. Let f be a single-valued rule satisfying ordinality and Property λ. We would like to show that f is dominant strategy implementable. To show this, we first prove that Property λ implies quasi-constancy. Suppose, by contradiction, that for some $\theta \in \Theta$, some $i \in N$, and some $\theta_i' \in \Theta_i$, there exist $\ell \in f(\theta)$ and $\ell' \in f(\theta_i', \theta_{-i})$ such that $\ell_i \neq \ell_i'$. Since f is a single-valued function, by the full correspondence assumption, $f(\theta) \in \triangle_D$ for each $\theta \in \Theta$. By Remark 4, there exists $\lambda_i \in \Lambda_i$ such that for each $x \in X$, $u_i(x; \theta_i') = \lambda_i(u_i(x; \theta_i))$. So, $(S(\theta_i', \theta_{-i})) = \lambda_i((S(\theta), d(\theta)))$. Since $\ell_i \neq \ell_i'$ and $\ell, \ell' \in \triangle_D$, we obtain $U_i(\ell'; \theta_i') \neq U_i(\ell; \theta_i') = \lambda_i(U_i(\ell; \theta_i))$, which is a contradiction to Property λ. Hence, Property λ implies quasi-constancy. It is obvious that if a single-valued rule satisfies quasi-constancy, then it satisfies strategy-proofness. Hence, f is strategy-proof.

Let us prove that f is dominant strategy implementable. Let

$$\Theta_i^* := \left\{ \theta_i, \theta_i' \in \Theta_i : \begin{array}{l} \text{there exist } \ell, \ell' \in \triangle \text{ such that } U_i(\ell; \theta_i) > U_i(\ell'; \theta_i), \\ U_i(\ell'; \theta_i') > U_i(\ell; \theta_i'), U_i(\ell; \theta_i) > U_i(\ell^{\mathbf{0}}; \theta_i), \text{ and} \\ U_i(\ell'; \theta_i') > U_i(\ell^{\mathbf{0}}; \theta_i') \end{array} \right\} \subset \Theta_i.$$

It is easy to check that the domain Θ^* satisfies the following properties: (i) the existence of a common worst alternative; (ii) weak separability; and (iii) strict value distinction with a reference point.[12] Thus, we can use the mechanism $\Gamma^{MW} = (M^{MW}, g^{MW})$ constructed by Mizukami and Wakayama (2007) for dominant strategy implementation on the domains satisfying the three properties mentioned above. So, since f satisfies single-valuedness and strategy-proofness, then by Mizukami and Wakayama (2007),

$$f(\theta') = g^{MW}(DSE^{\Gamma^{MW}}(\theta')) \text{ for each } \theta' \in \Theta^*. \tag{12}$$

Let $\hat{\theta} \in \Theta \setminus \Theta^*$. Then, there exists $\theta' \in \Theta^*$ such that $\hat{\theta}$ and θ' yield the same v.N–M preference profile.[13] This implies that $DSE^{\Gamma^{MW}}(\hat{\theta}) = DSE^{\Gamma^{MW}}(\theta')$. By ordinality, $f(\hat{\theta}) = f(\theta')$. Thus, by (12), $f(\hat{\theta}) = f(\theta') = g^{MW}(DSE^{\Gamma^{MW}}(\theta')) = g^{MW}(DSE^{\Gamma^{MW}}(\hat{\theta}))$. Hence, we can conclude that f is dominant strategy implementable. □

A.2 Proof of Theorem 1

We first show the "only if" part. Let f be a dominant strategy implementable rule satisfying the non-disagreement condition.

---

[12] See Mizukami and Wakayama (2007) for the formal definitions of these properties.

[13] We now prove the following claim:

Claim. Let $\hat{\theta} \in \Theta \setminus \Theta^*$. Then, there exists $\theta' \in \Theta^*$ such that for each $i \in N$ and each $\ell, \ell' \in \triangle$,

(i) if $U_i(\ell; \hat{\theta}_i) > U_i(\ell'; \hat{\theta}_i)$, then $U_i(\ell; \theta_i') > U_i(\ell'; \theta_i')$;

(ii) if $U_i(\ell; \hat{\theta}_i) = U_i(\ell'; \hat{\theta}_i)$, then $U_i(\ell; \theta_i') = U_i(\ell'; \theta_i')$;

(iii) if $U_i(\ell; \hat{\theta}_i) < U_i(\ell'; \hat{\theta}_i)$, then $U_i(\ell; \theta_i') < U_i(\ell'; \theta_i')$.

To prove the claim, we shall derive a contradiction for the case (i). The proofs for other cases (ii) and (iii) are similar, so we omit them. Suppose that for each $\theta' \in \Theta^*$, there exist $i \in N$ and $\ell, \ell' \in \triangle$ such that $U_i(\ell; \hat{\theta}_i) > U_i(\ell'; \hat{\theta}_i)$ and $U_i(\ell; \theta_i') \leq U_i(\ell'; \theta_i')$. Assuming $U_i(\ell; \theta_i') < U_i(\ell'; \theta_i')$ immediately yields a contradiction to $\hat{\theta} \notin \Theta^*$. Thus, we assume $U_i(\ell; \theta_i') = U_i(\ell'; \theta_i')$. Let $\ell^i \in \triangle$ be such that $\ell^i(e^i) = \frac{1}{2}$ and $\ell^i(\mathbf{0}) = \frac{1}{2}$. By the axiom of independence, we have $U_i\left(\frac{1}{2}\ell + \frac{1}{2}\ell^i; \hat{\theta}_i\right) > U_i\left(\frac{1}{2}\ell' + \frac{1}{2}\ell^i; \hat{\theta}_i\right)$ and $U_i\left(\frac{1}{2}\ell + \frac{1}{2}\ell^i; \theta_i'\right) = U_i\left(\frac{1}{2}\ell' + \frac{1}{2}\ell^i; \theta_i'\right)$. Let $\hat{\ell}$ be such that $\hat{\ell}(e^i) = \frac{1}{2}\ell(e^i) + \frac{1}{2}\ell^i(e^i) - \varepsilon$, $\hat{\ell}(\mathbf{0}) = \frac{1}{2}\ell(\mathbf{0}) + \frac{1}{2}\ell^i(\mathbf{0}) + \varepsilon$, and $\hat{\ell}(x) = \frac{1}{2}\ell(x) + \frac{1}{2}\ell^i(x)$ for each $x \in X \setminus \{\mathbf{0}, e^i\}$. Let $\tilde{\ell}$ be such that $\tilde{\ell}(e^i) = \frac{1}{2}\ell'(e^i) + \frac{1}{2}\ell^i(e^i) + \varepsilon$, $\tilde{\ell}(\mathbf{0}) = \frac{1}{2}\ell'(\mathbf{0}) + \frac{1}{2}\ell^i(\mathbf{0}) - \varepsilon$, and $\tilde{\ell}(x) = \frac{1}{2}\ell'(x) + \frac{1}{2}\ell^i(x)$ for each $x \in X \setminus \{\mathbf{0}, e^i\}$. Then, by the continuity of U_i, there exists $\varepsilon > 0$ such that $U_i(\hat{\ell}; \hat{\theta}_i) > U_i(\tilde{\ell}; \hat{\theta}_i)$ and $U_i(\hat{\ell}; \theta_i') < U_i(\tilde{\ell}; \theta_i')$. Obviously, $\hat{\ell}, \tilde{\ell} \in \triangle$. However, this is a contradiction to $\hat{\theta} \notin \Theta^*$.

Step 1: $f(\theta) \in \triangle_P$ for each $\theta \in \Theta$. It immediately follows from Lemma 2, Remark 5, and the non-disagreement condition.

Step 2: For each $\theta, \theta' \in \Theta$, $f(\theta) = f(\theta')$. Suppose, by contradiction, that there exists $\theta, \theta' \in \Theta$ such that $f(\theta) \neq f(\theta')$. By Step 1, $f(\theta), f(\theta') \in \triangle_P$. Let $f(\theta) = \ell^x$ and $f(\theta') = \ell^{x'}$. Note that $x_2 \neq x_2'$. By Lemma 2, dominant strategy implementability implies quasi-constancy. Thus, by quasi-constancy, $f_1(\theta_1', \theta_2) = \ell_1^x$ and $f_1(\theta_1, \theta_2') = \ell_1^{x'}$. By Step 1, $f_2(\theta_1', \theta_2) = \ell_2^x$ and $f_2(\theta_1, \theta_2') = \ell_2^{x'}$. Then, since $u_2(\cdot; \theta_2)$ and $u_2(\cdot; \theta_2')$ are strictly increasing, if $x_2 < x_2'$, $U_2(f(\theta_1, \theta_2'); \theta_2) > U_2(f(\theta); \theta_2)$; otherwise $U_2(f(\theta_1', \theta_2); \theta_2') > U_2(f(\theta'); \theta_2')$. This is a contradiction to quasi-constancy. Thus, by Lemma 2, this is a contradiction to dominant strategy implementability.

Before starting Step 3, we introduce a notation and a class of rules. For each $\alpha \in [0,1]$, let $x(\alpha)$ be a pure outcome such that $x(\alpha) = (\alpha, 1-\alpha)$. A constant rule f^α which always assigns only an efficient lottery is defined as follows: for each $\theta \in \Theta$, $f^\alpha(\theta) = \ell^{x(\alpha)}$.

Step 3: For each $\alpha \in (0,1)$, f^α violates welfarism.[14] Let $\alpha \in (0,1)$. Obviously, f^α satisfies single-valuedness. There exist $\theta, \theta' \in \Theta$ and $(\lambda_1, \lambda_2) \in \Lambda_1 \times \Lambda_2$ such that (i) $\theta \neq \theta'$; (ii) $(S(\theta_1', \theta_2), d^{\mathbf{0}}(\theta_1', \theta_2)) = \lambda_1(S(\theta), d^{\mathbf{0}}(\theta))$; (iii) $(S(\theta'), d^{\mathbf{0}}(\theta')) = \lambda_2(S(\theta_1', \theta_2), d^{\mathbf{0}}(\theta_1', \theta_2))$; (iv) $(S(\theta'), d^{\mathbf{0}}(\theta')) = (S(\theta), d^{\mathbf{0}}(\theta))$; (v) $\lambda_1(d_1^{\mathbf{0}}(\theta)) = d_1^{\mathbf{0}}(\theta)$ and $\lambda_1(a_1(S(\theta), d^{\mathbf{0}}(\theta))) = a_1(S(\theta), d^{\mathbf{0}}(\theta))$; and (vi) $\lambda_2(d_2^{\mathbf{0}}(\theta_1', \theta_2)) = d_2^{\mathbf{0}}(\theta_1', \theta_2)$ and $\lambda_2(a_2(S(\theta_1', \theta_2), d^{\mathbf{0}}(\theta_1', \theta_2))) = a_2(S(\theta_1', \theta_2), d^{\mathbf{0}}(\theta_1', \theta_2))$. This is illustrated in Figure 3.

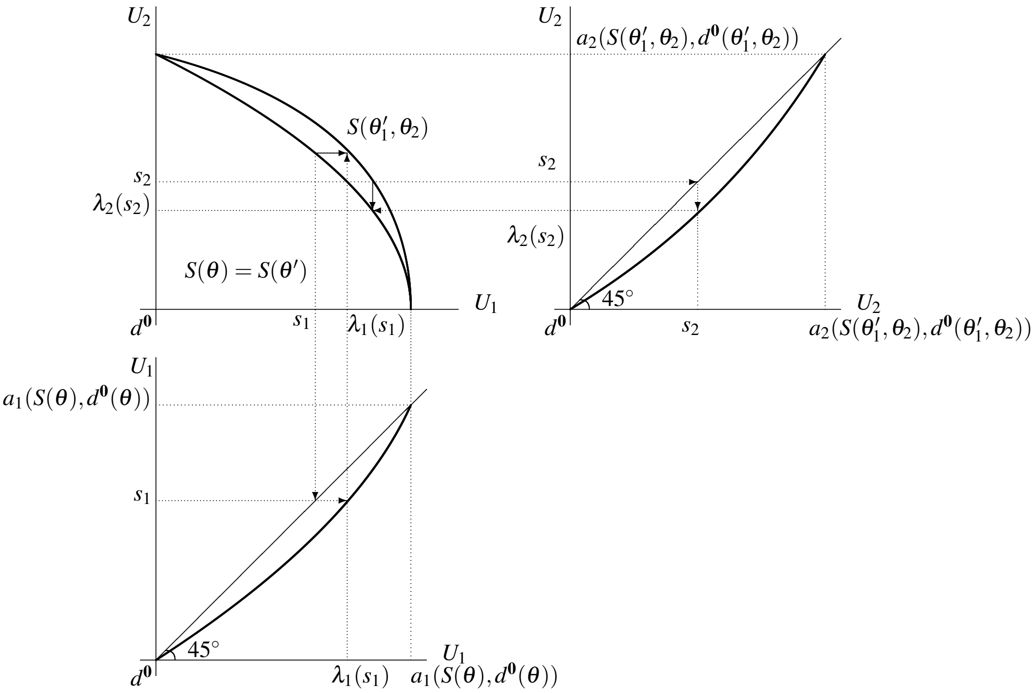

Figure 3: An example of transformation functions (λ_1, λ_2) in the proof of Theorem 1.

Since $\alpha \in (0,1)$, $a_1(S(\theta), d^{\mathbf{0}}(\theta)) > U_1(f^\alpha(\theta); \theta_1) > 0$ and $a_2(S(\theta), d^{\mathbf{0}}(\theta)) > U_2(f^\alpha(\theta); \theta_2) > 0$. Thus, $\lambda(U(f^\alpha(\theta); \theta)) \neq U(f^\alpha(\theta); \theta)$. Then, $(S(\theta), d^{\mathbf{0}}(\theta)) = (S(\theta'), d^{\mathbf{0}}(\theta'))$, but $U(f^\alpha(\theta); \theta) \neq \lambda(U(f^\alpha(\theta); \theta)) = U(f^\alpha(\theta); \theta') = U(f^\alpha(\theta'); \theta')$ (the last equality comes from the constancy of f^α). Hence, f^α violates welfarism.

[14] A rule f satisfies *welfarism* if for each θ, $\theta' \in \Theta$, if $(S(\theta), d(\theta)) = (S(\theta'), d(\theta'))$, then $U(\ell; \theta) = U(\ell'; \theta')$ for each $\ell \in f(\theta)$ and each $\ell' \in f(\theta')$.

Step 4: f^1 and f^0 satisfy welfarism. It suffices to show that the rule f^1 satisfies welfarism, since the proof for the rule f^0 can be dealt with in a parallel way. Let θ, θ' be such that $(S(\theta'), d(\theta')) = (S(\theta), d(\theta))$. Then, $U(f^1(\theta); \theta) = U(f^1(\theta'); \theta')$, since $(a_1(S(\theta), d(\theta)), d_2(\theta)) = (a_1(S(\theta'), d(\theta')), d_2(\theta'))$. Hence, f^1 satisfies welfarism.

Step 5: Concluding. By Step 3, for each $\alpha \in (0,1)$, f^α cannot induce any bargaining solutions. Thus, by Step 4, it is sufficient to consider f^1 and f^0. Without loss of generality, we assume $f = f^1$. Then, for each $(S,d) \in \Sigma$, $F(S,d) = (a_1(S,d), d_2)$, where F is the bargaining solution induced by the rule f^1. That is, F is dictatorial. It follows from Theorem 2.7 in Roemer (1996) that F satisfies efficiency, scale invariance, and strong monotonicity.

We next show the "if" part. Let F be a bargaining solution satisfying efficiency, scale invariance, and strong monotonicity. By Theorem 2.7 in Roemer (1996), the bargaining solution F is dictatorial. Then, there exists a rule f inducing the bargaining solution F and an agent $i \in N$ such that for each type profile $\theta \in \Theta$, $f(\theta) = \ell^{e^i}$. Hence, it is easily checked that the rule f is dominant strategy implementable and satisfies the non-disagreement condition. □

A.3 Alternative proof of Theorem 2

Since the "if" part is straightforward, we prove the "only if" part. Let F be a bargaining solution satisfying efficiency, scale invariance, and strong monotonicity.

Step 1: $F(S,d) \in (a_1(S,d), d_2) \cup (d_1, a_2(S,d))$ for each $(S,d) \in \Sigma$. Suppose, by contradiction, that $F(S,d) \notin (a_1(S,d), d_2) \cup (d_1, a_2(S,d))$ for some $(S,d) \in \Sigma$. Then, efficiency implies that

$$F(S,d) \in \partial S \setminus \{(a_1(S,d), d_2), (d_1, a_2(S,d))\}. \tag{13}$$

Consider $(S',d) \in \Sigma$ such that $S' \supset S$ and $a_i(S',d) = a_i(S,d)$ for each $i \in \{1,2\}$. Then, strong monotonicity implies $F(S',d) \geqq F(S,d)$.

Next, consider $(S'',d) \in \Sigma$ and $\tau'' \in T$ such that $S'' \supset S'$ and $(S'',d) = \tau''(S,d)$, where $\tau''(S) = \left\{s'' \in \mathbb{R}^2 \colon\ s'' = \tau''(s) = (\tau_1''(s_1), s_2) \text{ for some } s \in S\right\}$. Then, scale invariance implies $F(S'',d) = \tau''(F(S,d))$. Since $S'' \supset S'$, strong monotonicity implies $F(S'',d) \geqq F(S',d)$. Thus, we obtain $F(S'',d) = \tau''(F(S,d)) \geqq F(S',d) \geqq F(S,d)$, which implies

$$F_2(S',d) = F_2(S,d). \tag{14}$$

Finally, consider $(S''',d) \in \Sigma$ and $\tau''' \in T$ such that $S''' \supset S'$ and $(S''',d) = \tau'''(S,d)$, where $\tau'''(S) = \left\{s''' \in \mathbb{R}^2 \colon\ s''' = \tau'''(s) = (s_1, \tau_2'''(s_2)) \text{ for some } s \in S\right\}$. Then, by reasoning similar to that above, scale invariance and strong monotonicity together implies

$$F_1(S',d) = F_1(S,d). \tag{15}$$

By (14) and (15), we obtain $F(S',d) = F(S,d)$. Then, (13) and $S' \supset S$ together imply $F(S',d) \notin \partial S'$, contradicting efficiency.

Step 2: F is dictatorial. Consider $(S,d) \in \Sigma$. Then, by Step 1, we have either $F(S,d) = (a_1(S,d), d_2)$ or $F(S,d) = (d_1, a_2(S,d))$. Without loss of generality, we assume that $F(S,d) = (a_1(S,d), d_2)$.

Consider $(S',d') \in \Sigma$. We shall show $F(S',d') = (a_1(S',d'), d_2')$. Let $(\bar{S},d') \in \Sigma$ be such that (i) $\bar{S} \subset S'$; (ii) $a_1(\bar{S},d') = a_1(S',d')$; and (iii) $(\bar{S},d') = \tau(S,d)$ for some $\tau \in T$. Then, scale invariance implies $F(\bar{S},d') = \tau(F(S,d))$. Thus $F(\bar{S},d') = (a_1(\bar{S},d'), d_2')$. Therefore, by (i), strong monotonicity implies that $F(\bar{S},d') \leqq F(S',d')$. Combined with (ii), these imply $(a_1(\bar{S},d'), d_2') = (a_1(S',d'), d_2') \leqq F(S',d')$, which establishes $F(S',d') = (a_1(S',d'), d_2')$. □

A.4 Proof of Theorem 3

Let F be a bargaining solution satisfying efficiency, scale invariance, and strong monotonicity. Then, there exists a rule f inducing the bargaining solution F by Remark 2. By Lemma 2, it is sufficient to show that the rule f satisfies single-valuedness, ordinality, and Property λ.

Step 1: f satisfies single-valuedness. Since the utility possibility set is strictly convex, all efficient lotteries are degenerate. Since F satisfies efficiency, f inducing F satisfies the following: $f(\theta)$ chooses only an efficient lottery for each $\theta \in \Theta$. Hence, f is single-valued.

Step 2: f satisfies ordinality. Suppose that f violates ordinality. Then, there exist $\theta, \theta' \in \Theta$ such that $f(\theta) \neq f(\theta')$, but there exist no $i \in N$ and $\ell, \ell' \in \triangle$ such that $U_i(\ell; \theta_i) \geq U_i(\ell'; \theta_i)$ and $U_i(\ell'; \theta_i') > U_i(\ell; \theta_i')$. This implies that θ and θ' yield the same v.N–M preference profile. Then, there exists $\tau \in T$ such that $\tau(S(\theta), d(\theta)) = (S(\theta'), d(\theta'))$. Since f is induced by F and satisfies single-valuedness by Step 1, $U(f(\theta); \theta) = F(S(\theta), d(\theta))$ and $U(f(\theta'); \theta') = F(S(\theta'), d(\theta'))$. Hence,

$$\tau(F(S(\theta), d(\theta))) = \tau(U(f(\theta); \theta)) = U(f(\theta); \theta') \neq U(f(\theta'); \theta') = F(S(\theta'), d(\theta')),$$

which contradicts scale invariance.

Step 3: f satisfies Property λ. Let $\theta \in \Theta$, $i \in N$, $\theta_i' \in \Theta_i$, and $\lambda_i \in \Lambda_i$ be such that $(S(\theta_i', \theta_{-i}), d(\theta_i', \theta_{-i})) = \lambda_i(S(\theta), d(\theta))$. Let $(S, d) := (S(\theta), d(\theta))$ and $(S', d') := (S(\theta_i', \theta_{-i}), d(\theta_i', \theta_{-i}))$. By scale invariance, $F(S', d^{\mathbf{0}}) = F(S', d') - d'$ and $F(S, d^{\mathbf{0}}) = F(S, d) - d$.

Choose $\tau \in T$ such that

$$\tau(S') = \{s \in \mathbb{R}^n \colon\ s = \tau(s') = (\tau_i(s_i'), s_{-i}') \text{ for some } s' \in S'\} \supset S.$$

and $\tau(d^{\mathbf{0}}) = d^{\mathbf{0}}$. Then, scale invariance implies that

$$F(\tau(S', d^{\mathbf{0}})) = \tau(F(S', d^{\mathbf{0}})). \tag{16}$$

Since $\tau(S') \supset S$, strong monotonicity implies

$$F(\tau(S', d^{\mathbf{0}})) \geqq F(S, d^{\mathbf{0}}). \tag{17}$$

Next, consider $\tau' \in T$ such that

$$\tau'(S) = \{s' \in \mathbb{R}^n \colon\ s' = \tau'(s) = (\tau_i'(s_i), s_{-i}) \text{ for some } s \in S\} \supset \tau(S').$$

and $\tau'(d^{\mathbf{0}}) = d^{\mathbf{0}}$. Then, scale invariance implies

$$F(\tau'(S, d^{\mathbf{0}})) = \tau'(F(S, d^{\mathbf{0}})). \tag{18}$$

Since $\tau'(S) \supset \tau(S')$, by strong monotonicity

$$F(\tau'(S, d^{\mathbf{0}})) \geqq F(\tau(S', d^{\mathbf{0}})). \tag{19}$$

By (17) and (19), $F(\tau'(S, d^{\mathbf{0}})) \geqq F(\tau(S', d^{\mathbf{0}})) \geqq F(S, d^{\mathbf{0}})$. Thus, by (16) and (18), $\tau'(F(S, d^{\mathbf{0}})) \geqq \tau(F(S', d^{\mathbf{0}})) \geqq F(S, d^{\mathbf{0}})$, which implies $F_j(S', d^{\mathbf{0}}) = F_j(S, d^{\mathbf{0}})$ for each $j \in N \setminus \{i\}$. Since $(S', d') = \lambda_i(S, d)$, $d_j' = d_j$ for each $j \in N \setminus \{i\}$. Thus, by scale invariance, $F_j(S', d') = F_j(S', d^{\mathbf{0}}) + d_j' = F_j(S, d^{\mathbf{0}}) + d_j = F_j(S, d)$ for each $j \in N \setminus \{i\}$. By efficiency,

$$F(S', d') = \max_{s_i} \left\{ s \in S' \colon s_j = F_j(S, d) \text{ for any } j \in N \setminus \{i\} \right\} = \lambda_i(F(S, d)).$$

Since F is induced by f and satisfies single-valuedness by Step 1, $U(f(\theta);\theta) = F(S(\theta),d(\theta))$ and $U(f(\theta_i',\theta_{-i});(\theta_i',\theta_{-i})) = F(S(\theta_i',\theta_{-i}),d(\theta_i',\theta_{-i}))$. Thus, we can conclude that $U_i(f(\theta_i',\theta_{-i});\theta_i') = \lambda_i(U_i(f(\theta);\theta_i))$. Hence, f satisfies Property λ.

Step 4: Concluding. By Steps 1–3, it follows from Lemma 2 that f is dominant strategy implementable. $\square$

A.5 Proof of Theorem 4

Let $(S,d) \in \Sigma$. Let $\theta \in \Theta$ be such that $(S(\theta),d(\theta)) = (S,d)$. Since f is dominant strategy implementable, it follows from Lemma 2 that f is single-valued. By Remark 5, $f(\theta) \in \triangle_P \cup \{\ell^{\mathbf{0}}\}$. By the non-disagreement condition, $f(\theta) \in \triangle_P$, which implies $U(f(\theta);\theta) \in \partial S(\theta)$. Since F induced by f, we have $F(S,d) = F(S(\theta),d(\theta)) = U(f(\theta);\theta) \in \partial S(\theta) = \partial S$. Hence, F satisfies efficiency. $\square$

References

BARBERÀ, S., A. BOGOMOLNAIA, AND H. VAN DER STEL (1998): "Strategy-proof Probabilistic Rules for Expected Utility," *Mathematical Social Sciences*, **35**, 89–103.

HOWARD, J. V. (1992): "A Social Choice Rule and Its Implementation in Perfect Equilibrium," *Journal of Economic Theory*, **56**, 142–159.

KALAI, E. AND M. SMORODINSKY (1975): "Other Solutions to Nash's Bargaining Problem," *Econometrica*, **43**, 513–518.

KALAI, E. (1977): "Proportional Solutions to Bargaining Situations: Interpersonal Utility Comparisons," *Econometrica*, **45**, 1623–1630.

MIYAGAWA, E. (2002): "Subgame-Perfect Implementation of Bargaining Solutions," *Games and Economic Behavior*, **41**, 292–308.

MIZUKAMI, H. AND T. WAKAYAMA (2007): "Dominant Strategy Implementation in Economic Environments," forthcoming in *Games and Economic Behavior*.

MOULIN, H. (1984): "Implementing the Kalai-Smorodinsky Bargaining Solution," *Journal of Economic Theory*, **33**, 32–45.

NASH, J. F. (1950): "The Bargaining Problem," *Econometrica*, **18**, 155–162.

RAIFFA, H. (1953): "Arbitration Schemes for Generalized Two-person Games," in *Contributions to the Theory of Games II*, ed. by H. Kuhn and A. Tucker, 361–378. Princeton University Press.

ROEMER, J. E. (1988): "Axiomatic Bargaining Theory on Economic Environments," *Journal of Economic Theory*, **45**, 1–31.

ROEMER, J. E. (1996): *Theories of Distributive Justice*. Harvard University Press.

ROTH, A. E. (1979): *Axiomatic Models of Bargaining*, vol. 170 of Lecture Notes in Economics and Mathematical Systems. Springer-Verlag.

VARTIAINEN, H. (2006): "Nash Implementation and the Bargaining Problem," forthcoming in *Social Choice and Welfare*.