Virtual implementation in incomplete information environments with infinite alternatives and types

Guoqiang Tian

Department of Economics, Texas A & M University, College Station, TX 77843, USA
Submitted 1 March 1996; accepted 1 September 1996

Abstract

This paper considers virtual implementation in both dominant strategy equilibrium and Bayesian strategy equilibrium for incomplete information environments with general payoff functions, any number of agents, and arbitrary sets of alternatives and types. We characterize virtual implementability in private-value models by giving necessary and sufficient conditions. We show that a social choice function is virtually implementable in Bayesian (resp. dominant strategy) equilibrium if and only if it is Bayesian (resp. dominant strategy) incentive compatible. We also show that incentive compatibility is the necessary and sufficient condition for a class of common-value economic environments where there are two or more alternatives that have no valuation externalities. In addition, our mechanism for transferable utility models satisfies budget balancing constraints.

Keywords: Incentive compatibility; Mechanism design; Virtual implementation; Bayesian equilibrium; Dominant strategy equilibrium

1. Introduction

In many contexts, collective decision making has to preserve some kind of decentralization when direct control is impossible or inappropriate. Incentive compatibility then becomes a basic requirement to consider in the design of a social organization in general and an economic organization in particular. Since agents have private information, and may find that it is advantageous to distort the information they reveal in a way which cannot be sanctioned, they may use such
information strategically to advance their own interests. This implies that the basic principle of mechanism design with incomplete information must require the organization to provide individuals with appropriate incentives so that individuals' interests are consistent with the goals of the organization. An organization (social choice) rule which has this consistency property is called incentive compatible. This abstract formulation embraces a variety of specific applications, including problems of optimal taxation, choice of public project, contract theory, auction theory, principle-agent models, etc. The fundamental problem becomes the characterization of what various institutions can achieve using incentive compatible mechanisms. Mechanism design theory (implementation theory) studies precisely this problem. The answers to this problem, however, depend on equilibrium solution concepts which describe individuals' behavior and information requirements. Many solution concepts, including dominant strategy equilibrium, Nash equilibrium, Bayesian equilibrium, and maxmin equilibrium, have been used in the literature.¹

In the study of incentive compatibility of a social choice goal, the revelation principle has played a prominent role. Roughly, it says that, when solution concepts are given by a dominant strategy equilibrium, Bayesian equilibrium, or maxmin equilibrium, a social goal attainable by an abstract incentive mechanism can also be attained by a revelation (direct) mechanism in which every individual has an incentive to report his private information truthfully as long as all other individuals are also truthfully reporting their private information, and, further, the outcomes induced by the truthful reports coincide with the outcomes induced by non-cooperative equilibria of the original mechanism. Thus, the revelation principle provides a simple method of representing the constraints imposed by private information and strategic behavior. Until recently, this approach has been widely used in the literature to consider the incentive compatibility of a mechanism.

While the incentive compatibility requirement is central, it may not be sufficient for a mechanism to give all desirable outcomes when the mechanism has multiple equilibria. The goal is to design a mechanism which has a desirable equilibrium outcome, but which may also have an undesirable equilibrium outcome. The severity of this multiple equilibria problem has been illustrated by Demski and Sappington (1984), Postlewaite and Schmeidler (1986), Repullo (1986) and others. This raises the question of how to solve this kind of multiple equilibria problem. That is, given a social choice goal, does there exist a

mechanism that generates only outcomes which are consistent with the social choice goal under some solution concept of self-interested behavior? If so, we call it implementable.

In attempts to solve the multiple equilibria problem, three approaches have been used in the implementation literature. One approach is to 'exactly' implement a social choice rule. One wants to design an incentive compatible mechanism such that the set of equilibrium outcomes of the mechanism coincides with the set of socially desirable alternatives for all environments under consideration. A seminal characterization on exact Nash implementation was given by Maskin (1977). The complete proof and further generalizations were provided by Dasgupta, et al. (1979), Repullo (1987), Sajio (1988), Moore and Repullo (1990) and others. These characterization results show what is possible for implementation of a general social choice function (correspondence), but not what is realistic. For 'better' mechanism design, see Hurwicz (1979), Schmeidler (1980), Walker (1981), Postlewaite and Wettstein (1989), Tian (1989, 1990, 1993, 1994, 1996a) and many others. Characterization results for Bayesian implementation were given by Postlewaite and Schmeidler (1986), Palfrey and Srivastava (1989b), Mookherjee and Reichelstein (1990), Jackson (1991), Tian (1997), among many others. However, exact implementation is very restrictive. A social choice function is exactly Nash (Bayesian) implementable only if it is Maskin (resp. Bayesian) monotonic. For the incomplete information case, the problem is even worse since many desirable social correspondences are not monotonic. While admitting correspondences and introducing domain restrictions on environments may allow one to escape this difficulty, as Palfrey and Srivastava (1987) demonstrated, many commonly used social choice correspondences, such as Pareto efficient allocations, core allocations, boundary Walrasian and Lindahl allocations, are not monotonic and thus not implementable for incomplete information environments.

The second approach for solving the multiple equilibria problem is to use refinements of Nash (Bayesian) equilibrium (cf. Moore and Repullo (1988), Palfrey and Srivastava (1989a, 1991a), Abreu and Sen (1990), among many others). This approach has had as its primary objective to enlarge the class of implementable social choice functions (correspondences). The third approach is virtual implementation, which achieves the same goal of enlarging the class of implementable social choice rules. Although a social choice goal is not exactly implementable, it may be virtually implementable in the sense that there may exist an implementable social choice function which can be arbitrarily close to the original social choice goal. This paper will use the third approach to attempt to solve the multiple equilibria problem.

What is the class of virtually implementable social choice correspondences? For complete information environments, Matsushima (1988) and Abreu and Sen (1991) showed that in societies with at least three agents, all social choice correspondences for random allocations (mapping preferences to lotteries) are virtually implementable. For incomplete information environments, the class of
virtually implementable social choice functions is much more restrictive. Abreu and Matsushima (1990) showed that, when the set of types is finite, a social choice function for random allocations is virtually implementable in iteratively undominated strategies if and only if it is Bayesian incentive compatible and satisfies a technical measurability condition. Matsushima (1993) considered virtual implementation in Bayesian equilibrium when there are at least three agents who have quasi-linear utility functions, and where the sets of alternatives and types are finite. He showed that a social choice function for random allocations is virtually Bayesian implementable if it is strictly Bayesian incentive compatible. Recently, Duggan (1993) considered virtual Bayesian implementation by relaxing the finiteness of the set of types. He showed that, when utility functions of agents are quasi-linear and have private values (no valuation externalities), and the set of alternatives is finite, a social choice function is virtually Bayesian implementable if and only if it is Bayesian incentive compatible.

In this paper we complement and extend earlier work on virtual implementation in incomplete information environments to include infinite alternatives and types. We show that the multiple equilibrium problem can be solved in the large and important class of environments with private values. We consider virtual implementation in both dominant strategy equilibrium and Bayesian strategy equilibrium for incomplete information environments with general utility functions, any number of agents, and arbitrary sets of alternatives and types. We characterize virtual implementability in private-value models by giving necessary and sufficient conditions. We show that, under some regularity conditions on utility functions, a social choice function for random or non-random allocations is virtually implementable in Bayesian (dominant strategy) equilibrium if and only if it is Bayesian (dominant strategy) incentive compatible. We obtain similar results for models with side payments. We also show that incentive compatibility is the necessary and sufficient condition for some common-value models which contain both valuation externality and valuation externality-free alternatives. The mechanisms we construct improve upon those of Duggan (1993) in a number of respects. Our results consider infinite alternatives. In addition, our mechanisms with transfers are feasible in the sense that they have budget balancing for all message profiles. Furthermore, non-equilibrium outcomes of the mechanisms are $\varepsilon$-close to an equilibrium outcome when the social choice function is continuous and non-equilibrium strategies are $\varepsilon$-close to the equilibrium strategy. This property is highly desirable if one is to seriously consider accepting the mechanisms — in particular, if such a mechanism which is terminated prior to the attainment of an equilibrium

---

2 That is, each individual's utility function depends only on the outcome and on his own private information.

3 The private valuation assumption has been widely used in the mechanism design literature. For instance, Palfrey and Srivastava (1989a) considered the multiple equilibria problem by eliminating weakly dominated strategies only for private-value models.
is actually implemented, one would like the allocation at non-equilibrium to be close to the equilibrium outcome as long as the message is close to the equilibrium message.

Thus, our results significantly enlarge the class of implementable social choice rules by using the virtual implementation approach. The reason why virtual implementation is in an incomplete information setting leads to a much richer set than exact implementation is essentially because of the existence of a direct revelation mechanism, defined on a finite subset of alternatives, in which truth-telling is a strict dominant strategy equilibrium for private valuation economic environments. As a result, any convex combination (probability mixture) of any social choice function and the outcome function of this direct mechanism is implementable by a mechanism in which any strategy with truth-telling is a Bayesian (resp. dominant strategy) equilibrium, and any other non-truth-telling strategy is not a Bayesian (resp. dominant strategy) equilibrium. Since any neighborhood of an arbitrary social choice rule contains such a probability mixture, the social choice rule is virtually implementable. That is, any neighborhood of an arbitrary social choice rule contains a strictly Bayesian incentive compatible social choice rule which is Bayesian monotonic.

It is interesting to point out that this approach not only shows the virtual implementability of a general social choice function, but also gives a simple way to construct a mechanism which implements a social choice function that is arbitrarily close to the original social choice function. To reach the goal, one needs only to construct the canonical direct revelation mechanism which is defined only on a finite subset of alternatives (even two alternatives are enough). Then, the convex combination (probability mixture) of the original social choice function and the outcome function of this canonical revelation mechanism with an appropriate message space can be arbitrarily closed to the original social choice rule and thus virtually implements the original social choice rule.

This paper is organized as follows. Section 2 states some notation and definitions concerning the analysis framework. In Section 3 we provide the theorems which give necessary and sufficient conditions for virtual implementation in Bayesian and dominant strategy equilibrium for private-value models. We do so by providing a canonical direct revelation mechanism on finite approximation of alternatives, for which truthful revelation of types is a strict dominant strategy equilibrium. Section 4 considers virtual implementation for environments with side payments. Section 5 considers the same problem for common-value models. Concluding remarks are given in Section 6.

2. The analysis framework

Consider a society with \( n \) individuals. The set of individuals is denoted by \( N = \{1, 2, \ldots, n\} \). Let \( T_i \) be the set of possible types (preferences, parameters,
beliefs, signals, etc.) of agent $i$, which summarizes the preferences and information of the agent. Assume that $T_i$ is a metric space and is endowed with a Borel $\sigma$-field $\mathcal{T}_i$. A type profile is $t = (t_1, \ldots, t_n)$ and $T = \prod_{i \in N} T_i$ denotes the set of all such profiles. Denote by $T_{-i} = \prod_{j \neq i} T_j$ the set of possible profiles of the types of all agents other than $i$, and $T_{-(i,j)} = \prod_{k \neq i,j} T_k$ the set of possible profiles of the types of all agents other than $i$ and $j$. We summarize $i$'s beliefs about the other agents by means of a collection of conditional distribution functions (probability measures) on $T_{-i}$. For each $i$, let $G_i(t_{-i} | t_i)$ denote agent $i$'s regular conditional probability measure that other agents receive the profile of types $t_{-i}$ when agent $i$ receives the type $t_i$. Note that we do not make the so-called consistent beliefs assumption (Harsanyi's (1967-68) terminology). Instead, we allow the probability distribution to vary with individuals’ types, which captures the notion that the types of an agent may correspond to different beliefs about other agents.

Let $A$ denote an arbitrary set of social states (feasible alternatives, or allocations), which may be a finite, infinite, or continuum element set. Assume $A$ is a metric space. Let $\mathbb{P}(A)$ denote the space of all (Borel) probability measures on $A$. An element of $\mathbb{P}(A)$ is also called a random allocation. $\mathbb{P}(A)$ can be metrized into a metric space (cf. Parthasarathy (1967) or Billingsley (1968)) endowed with the weak* topology. An allocation rule is a function $x : T \to \mathbb{P}(A)$.

Given a profile of types $t$, and an alternative $a \in A$, the (ex-post) utility function of agent $i$ is given by $U_i(a, t_i)$, which is assumed to be bounded. We normalize the utility function so that for some allocation $a^0$, $U_i(a^0, t_i) = 0$ for all $t_i \in T_i$. When agent $i$ must make choices among random allocations, he will linearly extend his utility function over pure allocations to random allocations. In other words, we assume that agent $i$ has a von Neumann–Morgenstern expected utility function over random allocations. We use $\tilde{U}_i : \mathbb{P}(A) \to \mathbb{R}$ to denote agent $i$’s von Neumann–Morgenstern expected utility function over random allocations $x \in \mathbb{P}(A)$, which is defined by the integral of $U_i$ with respect to the probability measure $x$,

$$\tilde{U}_i(x, t_i) = \int_A U_i(a, t_i) \, dx.$$ 

For any two random allocations (probability measures) $x^1, x^2 \in \mathbb{P}(a)$, their convex combination with coefficient $0 \leq \lambda \leq 1$, $(1 - \lambda)x^1 \oplus \lambda x^2$, is defined by

$$((1 - \lambda)x^1 \oplus \lambda x^2)(B) := (1 - \lambda)x^1(B) + \lambda x^2(B)$$

for every Borel set $B$ in the Borel $\sigma$-field of $A$. It is clear that $(1 - \lambda)x^1 \oplus \lambda x^2$ is also a probability measure. The convex combination of $L$ random allocations, $x^1, \ldots, x^L$, with coefficients $0 \leq \lambda^l \leq 1$ such that $\sum_{l=1}^{L} \lambda^l = 1$, is defined similarly.

Given a random allocation rule, $x : T \to \mathbb{P}(A)$, the interim (conditional expected) utility of $x$ to type of agent $i$ is

$$W_i(x, t_i) = \int_{T_{-i}} \tilde{U}_i(x(t), t_i) \, dG_i(t_{-i} | t_i).$$
The tuple $e = (N, A, \mathcal{P}, T, \{G_i\}, \{U_i\})$ is called an environment. We assume, as is standard, that the structure of $e$ is common knowledge to all agents and that each agent $i$ knows his own type. Denote by $E$ the set of all such environments.

The designer is assumed to know $T$ and $G_i$, but does not know the true types of agents. To achieve a social goal, the designer needs to construct an informationally decentralized mechanism $\langle M, g \rangle$, where $M = \prod_{i \in N} M_i$, whose elements are denoted by $m = (m_1, \ldots, m_n)$. The set $M_i$ contains the possible messages agent $i$ can use, and $M$ is called a message space. $g : M \rightarrow \mathcal{P}(A)$ is an outcome function. For each $m \in M$, $g(m)$ yields an outcome in $\mathcal{P}(A)$. Given a mechanism $\langle M, g \rangle$, each agent $i$ chooses messages $m_i$ as a function of his types. We call a mapping $\sigma_i : T_i \rightarrow M_i$ a strategy for agent $i$, and $\Sigma_i$ its set of strategies. Note that, as usual, we do not allow mixed or behavior strategies (so we do consider the virtual implementation in the presence of preplay communication in this paper) for simplicity. The notations $\sigma_{-i}$ and $\Sigma_{-i}$ are similarly defined. Given a strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$, the interim utility to $i$ when of type $t_i$ is given by

$$\Pi_{\langle M, g \rangle}^i(\sigma; t_i) = \int_{T_{-i}} \hat{U}_i(g(\sigma(t)), t_i) \, dG_i(t_{-i} | t_i).$$

and the ex-ante utility to $i$ is given by

$$\Pi_{\langle M, g \rangle}^i(\sigma) = \int_{T_i} \Pi_{\langle M, g \rangle}^i(\sigma; t_i) \, dG(t_i).$$

where $G(t_i)$ is the marginal probability measure of agent $i$.

When there is no risk of confusion, the subscripts for mechanisms may be dropped.

**Definition 1.** A strategy profile $\sigma$ is a dominant strategy equilibrium of a mechanism $\langle M, g \rangle$ defined on $E$ if, for all $t_i \in T_i$,

$$\Pi_{\langle M, g \rangle}^i(\sigma_i, \hat{\sigma}_{-i}; t_i) \geq \Pi_{\langle M, g \rangle}^i(\hat{\sigma}; t_i)$$

for all $i \in N$ and all $\hat{\sigma} \in \Sigma$. When $\sigma$ is a dominant strategy equilibrium, $g(\sigma)$ is called a dominant strategy equilibrium outcome. Denote by $D_{\langle M, g \rangle}(e)$ the set of all such outcomes for environment $e$. The strategy profile $\sigma$ is called a strict dominant strategy equilibrium if the above inequality holds strictly for all $i \in N$ and all $\hat{\sigma}_i$ different from $\sigma_i$ on $\Sigma_i$.

**Definition 2.** A strategy profile $\sigma$ is a Bayesian equilibrium of a mechanism $\langle M, g \rangle$ defined on $E$ if, for all $t_i \in T_i$,

$$\Pi_{\langle M, g \rangle}^i(\sigma; t_i) \geq \Pi_{\langle M, g \rangle}^i(\hat{\sigma}_i, \sigma_{-i}; t_i).$$

for all $\hat{\sigma}_i \in \Sigma_i$. When $\sigma$ is a Bayesian equilibrium, $g(\sigma)$ is called a Bayesian equilibrium outcome. Denote by $B_{\langle M, g \rangle}(e)$ the set of all such allocations for
environment \( e \). The strategy profile \( \sigma \) is called a **strict Bayesian equilibrium** if the above inequality holds strictly for all \( i \in N \) and all \( \tilde{\sigma}_i \) different from \( \sigma_i \) on \( \Sigma_i \).

**Remark 1.** Note that, as long as a strict dominant strategy equilibrium exists, it is unique. This may not be true for a strict Bayesian equilibrium.

A mechanism \( \langle M, g \rangle \) is called a **direct** or **revelation mechanism** if \( M = T \). That is, each agent announces a, possibly false, type \( t_i \), which determines an outcome \( g(t_1, \ldots, t_n) \in \mathcal{P}(A) \). Let \( \tau \) denote the truthful strategy profile; that is, \( \tau_i(t_i) = t_i \) for all \( i \in N \) and all \( t_i \in T_i \). Then truth telling is simply the identity function. We call \( \sigma_i \) a **deception** by \( i \), the interpretation being that when \( i \) is of type \( t_i \), he acts as if he were of type \( \sigma_i(t_i) \). For any equilibrium \( \sigma \) of a given mechanism \( \langle M, g' \rangle \), we can define a direct mechanism \( \langle T, g \rangle \) by \( g = g' \circ \sigma \) (i.e., \( g(t) = g'(\sigma(t)) \) for all \( t \in T \)). Then, by the well-known revelation principle, we know that the truthful strategy \( \tau \) is a Bayesian equilibrium for this direct mechanism.

**Definition 3.** A social choice function \( f: T \to \mathcal{P}(A) \) is **incentive compatible in dominant strategy** if \( \tau \) is a dominant strategy equilibrium of the game induced by \( \langle T, f \rangle \). A social choice function \( f \) is **strictly incentive compatible in dominant strategy** if \( \tau \) is a strict dominant strategy equilibrium of the game induced by \( \langle T, f \rangle \).

**Definition 4.** A social choice function \( f: T \to \mathcal{P}(A) \) is **Bayesian incentive compatible** if \( \tau \) is a Bayesian equilibrium of the game induced by \( \langle T, f \rangle \). A social choice function \( f \) is **strictly Bayesian incentive compatible** if \( \tau \) is a strict Bayesian equilibrium of the game induced by \( \langle T, f \rangle \).

The concept of Bayesian incentive compatibility means that every agent will report his type truthfully when all other agents are employing their truthful strategies and thus every truthful strategy profile is a Bayesian equilibrium of the direct mechanism \( \langle T, f \rangle \). Notice that Bayesian incentive compatibility does not say what is the best response of an agent when other agents are not using truthful strategies. So it may also contain some undesirable equilibrium outcomes when a mechanism has multiple equilibria. The implementation problem involves designing mechanisms to ensure that all equilibria result in desirable outcomes which are captured by the social choice function.

**Definition 5.** A social choice function \( f: T \to \mathcal{P}(A) \) is **exactly implementable** (or simply called **implementable**) **in Bayesian** (resp. **dominant strategy**) equilibrium on \( T \) if there is a mechanism \( \langle M, g \rangle \) such that, for every Bayesian (resp. dominant strategy) equilibrium of the mechanism \( \langle M, g \rangle \), \( g(\sigma(t)) = f(t) \) for all \( t \in T \).
Rather than exactly implementing a social choice function, this paper considers virtual implementation.

**Definition 6.** A social choice function \( f: T \rightarrow \mathcal{P}(A) \) is virtually implementable in Bayesian (resp. dominant strategy) equilibrium on \( T \) if, for any \( \epsilon > 0 \), there is a social choice function \( g: T \rightarrow \mathcal{P}(A) \) such that

(i) \( g \) is implementable in Bayesian (resp. dominant strategy) equilibrium on \( T \),

(ii) \( d_p(f(t), g(t)) < \epsilon \) for all \( t \in T \).

Here \( d_p \) is the Prohorov metric over \( \mathcal{P}(A) \).

**Remark 2.** The notions of exact implementability and virtual implementability given above are slightly different from those of Duggan (1993). He considers the implementation problem by using the ex-ante utility maximization approach instead of using the interim utility maximization approach, and thus requires that exact implementability and virtual implementability hold not for all types, but only for a subset \( S_i \) of \( T_i \) with \( G_i(S_i) = 1 \). The interim utility maximization approach as used in this paper is more standard in the implementation literature.

### 3. Main results

In this section, we will characterize the virtual implementability in dominant strategy equilibrium and Bayesian equilibrium by giving necessary and sufficient conditions for general environments about types, preferences, and allocation sets which we specified in the above section. We will show that a social choice rule is virtually implementable in Bayesian equilibrium (resp. dominant strategy equilibrium) if and only if it is Bayesian (resp. dominant strategy) incentive compatible. To show this, we first give a lemma which shows the existence of a strictly dominant strategy incentive compatible mechanism for a finite, \( K \)-element subset (lattices) of \( A \), denoted by \( A^K \), which satisfies some regularity conditions. This mechanism puts higher probability on those outcomes in this pre-specified finite set with higher utilities. Elements of the set \( A^K \), together with \( a^0 \), are denoted \( \{a^0, a^1, \ldots, a^K\} \). Let

- \( u^k_i(t_i) = U_i(a^k, t_i) \) \( k = 0, 1, \ldots, K \).
- \( u_i(t_i) = (u^0_i(t_i), u^1_i(t_i), \ldots, u^K_i(t_i)) \).

---

\(^4\) For any \( \delta > 0 \) and any Borel subset \( B \), define \( B^\delta = \{a \in A : d(a, B) < \delta\} \) (visualize the open set \( B^\delta \) as \( B \) wearing a halo of thickness \( \delta \)), where \( d \) is the uniform metric. Define the Prohorov metric (distance) between two Borel probability measures \( \mu^1 \) and \( \mu^2 \) as

\[ d_p(\mu^1, \mu^2) = \inf\{\delta > 0 : \mu^1(B^\delta) + \delta \geq \mu^2(B) \text{ for every } B\} \]

For the proof of \( d_p \) being a metric, see Pollard (1984) or Billingsley (1968). Note that \( d_p(\mu^1, \mu^2) \leq 1 \) for any two Borel probability measures \( \mu^1 \) and \( \mu^2 \).
Notice that, by the definition of $a^0$, $u_i^k(t_i) = 0$ for all $t_i \in T_i$. We make the following regularity assumption.

**Assumption 1.** There exists a $K$-element subset $A^K \subset A$ with $K > 1$ such that

1. $\sum_{k=1}^{K} u_i^k(t_i) \neq 0$ for all $t_i \in T_i$,
2. for each $t_i \in T_i$, $u_i^k(t_i) \neq 0$ for at least two elements in $A^K$,
3. $u_i(t_i) \neq u_i(\hat{t}_i)$ for all $t_i \neq \hat{t}_i$.

**Remark 3.** Conditions (1)-(3) in Assumption 1 are very general. If we can normalize the utility function by choosing $a_0$ as a minimum of the utility function, then Conditions (1) and (2) will be satisfied for non-constant utility functions. Further, when utility functions are monotonic in $t_i$, Condition (3) will be satisfied. Note that Condition (3) can be dropped when the set of types consists only of the set of utility functions of agents and each agent is required to report his utility functions. In addition, if one uses the ex-ante utility approach and the set of types consists only of the set of utility functions of agents, Condition 1 is unnecessary. Note that these three conditions are automatically satisfied in the finite alternatives setting of Duggan (1993).

Denote by $P(A^K)$ the set of all random allocations (probability measures) distributed only over $A^K$. Precisely, a random allocation $\mu \in P(A)$ belongs to $P(A^K)$ if and only if $\mu(A^K) = 1$. Then $\mu(A \setminus A^K) = 0$ and thus $P(A^K)$ is a subset of $P(A)$. Note that $P(A^K) = \Delta^{K-1} := \{ x \in \mathbb{R}^K : \sum_{k=1}^{K} x_k = 1 \}$ is the $K - 1$ dimensional unit simplex.

**Lemma 1.** There exists a direct mechanism $\langle T, \hat{g} \rangle$ with $\hat{g} : T \rightarrow P(A^K)$, such that the truthful strategy profile, $\tau$, is a strict dominant strategy equilibrium of the game induced by $\langle T, \hat{g} \rangle$.

**Proof:** Define the direct mechanism $\langle T, \hat{g} \rangle$ as follows. For each type $\tilde{t}$ in $T$ reported by agents, the probability $1/(2K) + \sum_{i=1}^{n} u_i^k(\tilde{t}_i)/2nK \| u_i(\tilde{t}_i) \|$ is assigned to allocation $a^k$ for $k \geq 1$, and the probability $1/2 - \sum_{i=1}^{n} \sum_{k=1}^{K} u_i^k(\tilde{t}_i)/2nK \| u_i(\tilde{t}_i) \|$ is placed on the allocation $a^0$, where $\| \cdot \|$ is the Euclidian norm. That is,

$$\hat{g}^0(\tilde{t}) = 1/2 - \sum_{i=1}^{n} \sum_{k=1}^{K} \frac{u_i^k(\tilde{t}_i)}{2nK \| u_i(\tilde{t}_i) \|}$$

$$\hat{g}^k(\tilde{t}) = 1/2K + \sum_{i=1}^{n} \frac{u_i^k(\tilde{t}_i)}{2nK \| u_i(\tilde{t}_i) \|}$$

for $k \geq 1$. 

Then,
\[ \hat{U}_i(\hat{g}(\hat{t}), t_i) = \int_{A^K} U_i(a, t_i) \, d\hat{g} \]
\[ = u_i(t_i) \cdot \hat{g}(\hat{t}) \]
\[ = \sum_{k=1}^{K} u_i^k(t_i) \left( \frac{1}{2K} + \sum_{j=1}^{n} \frac{u_j^k(\hat{t}_j)}{2nK ||u_j^k(\hat{t}_j)||} \right) \] (by noting \( u_i^0 = 0 \))
\[ = \sum_{k=1}^{K} \frac{u_i^k(t_i)u_i^k(\hat{t}_i)}{2nK ||u_i^k(\hat{t}_i)||} + \sum_{k=1}^{K} u_i^k(t_i) \left( \frac{1}{2K} + \sum_{j \neq i}^{n} \frac{u_j^k(\hat{t}_j)}{2nK ||u_j^k(\hat{t}_j)||} \right) \]
\[ = \|u_i(t_i)\| \frac{u_i(t_i) \cdot u_i(\hat{t}_i)}{2nK ||u_i^k(\hat{t}_i)||} + C(\hat{t}_i) \] (4)

where \( C(\hat{t}_i) \) is some term which does not depend on \( \hat{t}_i \).

One can verify that the (normalized) solution of
\[ \max_{p_i \neq 0} \frac{u_i(t_i)}{\|u_i(t_i)\|} \cdot \frac{p_i}{\|p_i\|} \]
is unique and given by \( p_i = u_i(t_i) \). Then \( \hat{U}_i(\hat{g}(t_i, \hat{t}_i), t_i) > \hat{U}_i(\hat{g}(\hat{t}), t_i) \) for all \( \hat{t}_i \neq t_i \). Thus \( \int_{A^K} \hat{U}_i(\hat{g}(t), t_i) \, dG(t_i|t_i) > 0 \) for all \( A_{-i} \) with \( G(A_{-i}|t_i) > 0 \), and particularly \( \Pi_{(T, \hat{g})}(t_i, \sigma_{-i}; t_i) > \Pi_{(T, \hat{g})}(\sigma; t_i) \) for all \( \sigma_i \neq \tau_i \) and all \( t_i \in T_i \).

Therefore, the truthful strategy profile \( \tau \) is a strict dominant strategy equilibrium of the game induced by \( (T, \hat{g}) \) which is defined on \( A^K \). The proof is completed.

By using this lemma and the mechanism \( \hat{g} \) constructed above, we have the following results.

**Theorem 1.** A social choice function \( f: T \to \mathbb{P}(A) \) is virtually implementable in dominant strategy equilibrium if and only if it is dominant strategy incentive compatible.

**Proof:** For any \( \varepsilon > 0 \), let \( \lambda = \min(\varepsilon, 1) \). Define a direct mechanism \( (T, g) \) by letting
\[ g(t) = (1 - \lambda)f(t) \oplus \lambda \hat{g}(t) \]
for each \( t \in T \). We first show that every truthful strategy \( \tau \) is a strict dominant strategy equilibrium of the induced game and therefore it is unique (cf. Remark 1). Indeed, for any \( i \in N \), any \( \hat{\sigma}_i \in \Sigma_i \) such that \( \hat{\sigma}_i \neq \tau_i \), and for any \( \hat{\sigma}_{-i} \in \Sigma_{-i} \), we have
\[ \Pi_{(T, g)}(\tau_i, \hat{\sigma}_{-i}; t_i) = (1 - \lambda)\Pi_{(T, f)}(\tau_i, \hat{\sigma}_{-i}; t_i) + \lambda \Pi_{(T, \hat{g})}(\tau_i, \hat{\sigma}_{-i}; t_i) \]
\[ > (1 - \lambda)\Pi_{(T, f)}(\tau_i; t_i) + \lambda \Pi_{(T, \hat{g})}(\hat{\sigma}; t_i) \]
\[ = \Pi_{(T, g)}(\hat{\sigma}; t_i). \]
Here the strict inequality follows from the dominant strategy incentive compatibility of \( \langle T, f \rangle \) and the strict dominant strategy incentive compatibility of \( \langle T, \hat{g} \rangle \). Thus \( g \) has a unique dominant strategy equilibrium and therefore it is implementable. It is clear that \( d_p(g(t), f(t)) \leq \epsilon \) for all \( t \in T \). Thus, \( f \) is virtually implementable in dominant strategy.

We now prove necessity. Suppose, by way of contradiction, that \( f \) is not incentive compatible. Then there is an \( i \) and \( \hat{\sigma} \neq \tau \) such that

\[
\Pi_{\langle T, f \rangle}(\hat{\sigma}; t_i) - \Pi_{\langle T, f \rangle}(\tau_i, \hat{\sigma}_{-i}; t_i) = 3\eta
\]

for some \( t_i \in T_i \) and \( \eta > 0 \). Since \( U_i(\cdot) \) is bounded, there is a \( D > 0 \) such that \( U_i(a, t_i) \leq D \) for all \( a \in A \) and all \( t_i \in T_i \). Let \( \epsilon = \eta/D \). Since \( \langle T, f \rangle \) is virtually implementable in dominant strategy equilibrium, there exists an implementable social choice function \( g : T \to \mathcal{P}(A) \) such that \( d_p(g(t), f(t)) \leq \epsilon \) for all \( t \in T \). Consider the direct mechanism \( \langle T, g \rangle \). Then

\[
\left| \Pi_{\langle T, g \rangle}^i(\hat{\sigma}; t_i) - \Pi_{\langle T, f \rangle}^i(\hat{\sigma}; t_i) \right| = \left| \int_{T_{-i}} \int_A U_i(a, t_i)(d_i(t) - d_g(t)) \, dG_i(t_{-i} | t_i) \right| \\
\leq D \int_{T_{-i}} (x(t) - g(t)) \, dG_i(t_{-i} | t_i) \\
\leq \epsilon D \leq \eta.
\]

Similarly, we have \( \left| \Pi_{\langle T, f \rangle}^i(\tau_i, \hat{\sigma}_{-i}; t_i) - \Pi_{\langle T, g \rangle}^i(\tau_i, \hat{\sigma}_{-i}; t_i) \right| \leq \eta \). But then

\[
\Pi_{\langle T, g \rangle}^i(\hat{\sigma}; t_i) - \Pi_{\langle T, g \rangle}^i(\tau_i, \hat{\sigma}_{-i}; t_i) = \left[ \Pi_{\langle T, g \rangle}^i(\hat{\sigma}; t_i) - \Pi_{\langle T, f \rangle}^i(\hat{\sigma}; t_i) \right] \\
+ \left[ \Pi_{\langle T, f \rangle}^i(\hat{\sigma}; t_i) - \Pi_{\langle T, f \rangle}^i(\tau_i, \hat{\sigma}_{-i}; t_i) \right] \\
+ \left[ \Pi_{\langle T, f \rangle}^i(\tau_i, \hat{\sigma}_{-i}; t_i) - \Pi_{\langle T, g \rangle}^i(\tau_i, \hat{\sigma}_{-i}; t_i) \right] \\
\geq -\eta + 3\eta - \eta \\
= \eta > 0,
\]

which contradicts the fact that \( g \) has truth-telling as a dominant strategy equilibrium. Therefore, \( f \) must be dominant incentive compatible. The proof is completed.

Since our finite canonical strict dominant strategy mechanism \( \langle T, \hat{g} \rangle \) is also a strict dominant strategy mechanism for complete information, we have the following corollary.
Corollary 1. A social choice function \( f: E \to \mathcal{P}(A) \) is virtually implementable in dominant strategy equilibrium on the set of complete information environments, \( E \), if and only if it is dominant strategy incentive compatible on \( E \).

Proof: The proof is the same as for Theorem 1.

Now we turn to the main result of the paper, which gives the necessary and sufficient condition for virtual implementability in Bayesian equilibrium.

Theorem 2. A social choice function \( f: T \to \mathcal{P}(A) \) is virtually implementable in Bayesian equilibrium if and only if it is Bayesian incentive compatible.

Proof: We first prove sufficiency. When \( n = 1 \), the proof is given by Theorem 1. So we need to consider only the case of \( n > 1 \). For any \( \varepsilon > 0 \), let \( \lambda = \min(\varepsilon/3, 1/3) \). Define a mechanism \( \langle M, g \rangle \) as follows. Let \( M_i = T_i \times T_i \times (0, 1/2) \times (0, 1/2) \). For each message \( m = (m_1, m_2, m_3, m_4) \) with \( m_3 = \beta_3 \) and \( m_1 = \rho_i \), define the outcome function by

\[
g(m) = \begin{cases} 
(1 - 3\lambda) \lambda \sigma(t) \odot \lambda \sigma(t) \odot \lambda \sigma(t) \odot \lambda \sigma(t) & \text{if } |S(m)| \leq 1 \\
(1 - 3\lambda) \lambda \sigma(t) \odot \lambda \sigma(t) \odot \lambda \sigma(t) \odot \lambda \sigma(t) & \text{otherwise}
\end{cases}
\]

where \( S(m) = \{ j \in N : m_j \neq m_j \} \), \( \beta = \min(\beta_1, \ldots, \beta_n) \), \( \rho = \max(\rho_1, \ldots, \rho_n) \), \( x \in \Delta^k \) is the random allocation which places equal probability on \( (\alpha^0, \alpha^1, \ldots, \alpha^k) \), and \( e^0 \in \beta^k \) is the degenerate random allocation which yields \( \alpha^0 \) with probability 1 and \( (\alpha^1, \ldots, \alpha^k) \) with probability 0 (i.e., \( e^0 \) is an element in \( \Delta^k \) such that its first component is one and the other components are zero). Notice that the mechanism defined above is well defined in the sense that, for all message profiles, alternatives are assigned nonnegative probabilities which sum to one (by noting that \( 1 + (1 - 3\lambda) \lambda \sigma(t) \odot \lambda \sigma(t) \odot \lambda \sigma(t) \odot \lambda \sigma(t) > 0 \), \( e^0 \) is a probability measure).

Note that the above mechanism has a desired property that for all \( \sigma = (\sigma^1, \sigma^2, \sigma^3, \sigma^4) \in \Sigma \), \( d_i(g(\sigma(t)), f(\sigma(t))) \leq \varepsilon \) for all \( t \in T \). We now show that only those strategies with truth-telling are Bayesian equilibria and thus this mechanism implements the social choice function \( g': T \to \mathcal{P}(A) \) defined by \( g'(t) = (1 - 3\lambda) \lambda \sigma(t) \odot 2 \lambda \sigma(t) \odot \lambda \sigma(t) \) for each \( t \in T \).

We consider three cases for a strategy profile \( \sigma = (\sigma^1, \sigma^2, \sigma^3, \sigma^4) \in \Sigma: (1) \sigma^1 = \sigma^2 = \sigma^3 = \sigma^4 \in \Sigma \), \( (2) \sigma^1 = \sigma^3 \neq \sigma^2 \), and \( (3) \sigma^1 \neq \sigma^2 \).

Case 1. \( \sigma^1 = \sigma^2 = \tau \). In this case, \( |S(\tau(t))| = 0 \) for all \( t \in T \). We prove that any \( \sigma \in \Sigma \) with \( \sigma^1 = \sigma^2 = \tau \) (i.e., \( \sigma^1(t) = \sigma^2(t) = t \) for all \( t \in T \)) is a Bayesian
equilibrium. Indeed, for any \( i \in N \) and any \( \hat{d}_i \in \Sigma \), such that \( \hat{d}_i \neq \sigma_i \), we still have 
\[ |S(\hat{d}_i(t_i), \sigma_-(t_-))| \leq 1 \]
for all \( t \in T \) and thus
\[
\Pi_{(M, \xi)}^i(\sigma; t_i) = (1 - 3\lambda) \Pi_{(T, f)}^i(\tau; t_i) + \lambda \Pi_{(T, g)}^i(\sigma; t_i) \\
+ \Pi_{(T, d)}^i(\tau; t_i) + \lambda \Pi_{(T, \epsilon)}^i(\sigma; t_i) \\
> (1 - 3\lambda) \Pi_{(T, f)}^i(\hat{d}_i^1, \tau_{-i}; t_i) + \lambda \Pi_{(T, g)}^i(\hat{d}_i^1, \tau_{-i}; t_i) \\
+ \Pi_{(T, d)}^i(\tau_{-i}; t_i) + \lambda \Pi_{(T, \epsilon)}^i(\sigma; t_i) \\
= \Pi_{(M, \xi)}^i(\hat{d}_i, \sigma_{-i}; t_i),
\]
where the inequality follows from the Bayesian incentive compatibility of \( \langle T, f \rangle \) and the strict dominant strategy incentive compatibility of \( \langle T, g \rangle \). Thus any strategy profile \( \sigma \) with \( \sigma^1 = \sigma^2 = \tau \) is a Bayesian equilibrium, and \( d_p(g(\sigma(t)), f(t)) \leq \epsilon \).

Case 2. \( \sigma^1 = \sigma^2 \neq \tau \). In this case, \( |S(\sigma(t))| = 0 \) for all \( t \in T \). We prove that any \( \sigma \in \Sigma \) with \( \sigma^1 = \sigma^2 \neq \tau \) is not a Bayesian equilibrium. Since \( \sigma^1 = \sigma^2 \neq \tau \), there is \( i \in N \) and \( t_i \in T_i \) such that \( \sigma_i^1(t_i) = \sigma_i^2(t_i) \neq t_i \). Let \( \hat{d}_i = (\sigma_i^1, \tau_i, \sigma_i^3, \sigma_i^4) \). Then \( |S(\hat{d}_i(t_i))| \leq 1 \) for all \( i \in T \) and thus
\[
\Pi_{(M, \xi)}^i(\hat{d}_i, \sigma_{-i}; t_i) = (1 - 3\lambda) \Pi_{(T, f)}^i(\sigma^1; t_i) + \lambda \Pi_{(T, g)}^i(\sigma^1; t_i) \\
+ \Pi_{(T, d)}^i(\tau^2, t_i) + \lambda \Pi_{(T, \epsilon)}^i(\sigma^1; t_i) \\
> (1 - 3\lambda) \Pi_{(T, f)}^i(\sigma^1; t_i) + \lambda \Pi_{(T, g)}^i(\sigma^1; t_i) \\
+ \Pi_{(T, d)}^i(\sigma^2; t_i) + \lambda \Pi_{(T, \epsilon)}^i(\sigma^1; t_i) \\
= \Pi_{(M, \xi)}^i(\sigma; t_i),
\]
where the inequality follows from the strict dominant strategy incentive compatibility of \( \langle T, g \rangle \).

Case 3. \( \sigma^1 \neq \sigma^2 \). We prove that any \( \sigma \in \Sigma \) with \( \sigma^1 \neq \sigma^2 \) is not a Bayesian equilibrium. Since \( \sigma^1 \neq \sigma^2 \), there is \( i \in N \) and \( t_i \in T_i \) such that \( \sigma_i^1(t_i) \neq \sigma_i^2(t_i) \).

Define
\[
A_{-i}^w = \{ t_{-i} \in T_{-i} : \sigma_i^1(t_{-i}) \neq \sigma_i^2(t_{-i}) \}
\]
\[
B_{-i}^w = \{ t_{-i} \in T_{-i} : \sigma_i^1(t_{-i}) = \sigma_i^2(t_{-i}) \}.
\]
Note that these sets form a partition of \( T_{-i} \). Then, \( G_{-i}(A_{-i}^w | t_i) + G_{-i}(B_{-i}^w | t_i) = 1 \), and thus we must have either \( G_{-i}(A_{-i}^w | t_i) > 0 \) or \( G_{-i}(B_{-i}^w | t_i) = 0 \) (i.e., \( G_{-i}(B_{-i}^w | t_i) = 1 \)).

Case 3a. \( G_{-i}(A_{-i}^w | t_i) > 0 \). In this case, we have \( |S(\sigma(t_i), \sigma_-(t_-))| > 1 \) for all \( t_{-i} \in A_{-i}^w \), and \( |S(\sigma(t_i), \sigma_-(t_-))| = 1 \) for all \( t_{-i} \in B_{-i}^w \). Let \( \mathcal{H}_i = (\sigma_i^1, \sigma_i^2, \hat{d}_i^1, \hat{d}_i^2) \) with
\[
(\hat{d}_i^3(t_i), \hat{d}_i^4(t_i)) = \begin{cases} 
(\beta_3^i, \beta_4^i) & \text{if } \sum_{t_i=1}^K u_i^1(t_i) > 0 \\
(\beta_3^i, \beta_4^i) & \text{if } \sum_{t_i=1}^K u_i^1(t_i) < 0.
\end{cases}
\]
Suppose agent \( i \) chooses \( \hat{\beta}_i \) < \( \beta \) and \( \hat{\rho}_i \) > \( \rho \). Then \( \hat{\beta}_i = \hat{\beta} < \beta \) and thus \( \hat{\rho}_i = \hat{\rho} > \rho \). By noting \( |S(\hat{\sigma}(t_i), \sigma_{-i}(t_{-i}))| > 1 \) for all \( t_{-i} \in B_{-i}^\sigma \), \( |S(\hat{\sigma}(t_i), \sigma_{-i}(t_{-i}))| = 1 \) and \( g(\hat{\sigma}(t_i), \sigma_{-i}(t_{-i})) = g(\sigma(t)) \) for all \( t_{-i} \in B_{-i}^\sigma \), we have

\[
\Pi^i_{(M,K)}(\hat{\sigma}_i, \sigma_{-i}; t_i) - \Pi^i_{(M,K)}(\sigma; t_i)
= \int_{A^\sigma_i} \left[ U_i(g(\hat{\sigma}_i, \sigma_{-i}; t_i)) - U_i(g(\sigma; t_i)) \right] \, dG_i(t_{-i} | t_i)
= \int_{A^\sigma_i} \left[ \frac{\lambda(1 - \beta - \hat{\rho})}{K + 1} - \frac{\lambda(1 - \beta - \rho)}{K + 1} \right] u_i(t_i) \cdot (\bar{x} - e_0) \, dG_i(t_{-i} | t_i)
= \int_{A^\sigma_i} \frac{\lambda}{K + 1} \left[ (\beta - \hat{\beta}) + (\rho - \hat{\rho}) \right] u_i(t_i) \cdot (\bar{x} - e_0) \, dG_i(t_{-i} | t_i)
= \int_{A^\sigma_i} \frac{\lambda}{(K + 1)^2} \left[ (\beta - \hat{\beta}) + (\rho - \hat{\rho}) \right] \sum_{k=1}^K u^k_i(t_i) \, dG_i(t_{-i} | t_i)
= \begin{cases} 
\frac{\lambda}{(K + 1)^2} (\beta - \hat{\beta}) \sum_{k=1}^K u^k_i(t_i) G_i(A^\sigma_{-i} | t_i) & \text{if } \sum_{k=1}^K u^k_i(t_i) > 0 \\
\frac{\lambda}{(K + 1)^2} (\rho - \hat{\rho}) \sum_{k=1}^K u^k_i(t_i) G_i(A^\sigma_{-i} | t_i) & \text{if } \sum_{k=1}^K u^k_i(t_i) < 0
\end{cases}
> 0.
\]

So \( \sigma \in \Sigma \) is not a Bayesian equilibrium.

**Case 3b.** \( G_{-i}(B_{-i}^\sigma | t_i) = 1 \). In this case, we have \( |S(\sigma(t_i), \sigma_{-i}(t_{-i}))| = 1 \) for all \( t_{-i} \in B_{-i}^\sigma \). This can be divided into two subsubcases: (i) \( \sigma^1_i(t_{-i}) = \sigma^2_i(t_{-i}) = t_{-i} \) for all \( t_{-i} \in B_{-i}^\sigma \) and (ii) \( \sigma^1_i(t_{-i}) = \sigma^2_i(t_{-i}) = t_{-i} \) for some \( t_{-i} \in B_{-i}^\sigma \).

**Case 3b (i).** (i) \( \sigma^1_i(t_{-i}) = \sigma^2_i(t_{-i}) = t_{-i} \) for all \( t_{-i} \in B_{-i}^\sigma \). Since \( G_{-i}(B_{-i}^\sigma | t_i) = 1 \) and \( \sigma^1_i(t_{-i}) = \sigma^2_i(t_{-i}) = t_{-i} \) for all \( t_{-i} \in B_{-i}^\sigma \), we may suppose that \( \sigma^1_i = \sigma^2_i = \tau_{-i} \). When \( \sigma^1_i(t_i) \neq t_i \), let \( \hat{\sigma}_i = (\sigma^1_i, \tau_i, \sigma_i^2, \sigma_i^3) \). Then we have

\[
\Pi^i_{(M,K)}(\hat{\sigma}_i, \sigma_{-i}; t_i) = (1 - 3\lambda) \Pi^i_{(T,\xi)}(\sigma^1_i; t_i) + \lambda \Pi^i_{(T,\xi)}(\sigma^1; t_i)
+ \lambda \Pi^i_{(T,\xi)}(\sigma^2_i; t_i) + \lambda \Pi^i_{(T,\xi)}
> (1 - 3\lambda) \Pi^i_{(T,\xi)}(\sigma^1_i; t_i) + \lambda \Pi^i_{(T,\xi)}(\sigma^1; t_i)
+ \lambda \Pi^i_{(T,\xi)}(\sigma^2; t_i) + \lambda \Pi^i_{(T,\xi)}
= \Pi^i_{(M,K)}(\sigma; t_i).
\]
where the inequality follows from the strict dominant strategy incentive compatibility of \( \langle T, \hat{g} \rangle \).

When \( \sigma_i^1(t_i) = t_i, \sigma_i^2(t_i) \neq t_i \), let \( \hat{\sigma}_i = (\tau_i, \sigma_i^2, \sigma_i^3, \sigma_i^4) \). Then we have
\[
\Pi_{(M, \mathcal{X})}^{i}(\hat{\sigma}_i, \sigma_{-i}; t_i) = (1 - 3\lambda) \Pi_{(T, f)}^{i}(\tau; t_i) + \lambda \Pi_{(T, \hat{g})}^{i}(\tau)
\]
\[
+ \lambda \Pi_{(T, \hat{g})}^{i}(\sigma^2; t_i) + \lambda \Pi_{(T, \hat{g})}^{i}(\sigma^2)
\]
\[
> (1 - 3\lambda) \Pi_{(T, f)}^{i}(\sigma^1, \tau_{-i}; t_i)
\]
\[
+ \lambda \Pi_{(T, \hat{g})}^{i}(\sigma^1, \tau_{-i}; t_i)
\]
\[
+ \lambda \Pi_{(T, \hat{g})}^{i}(\sigma^2; t_i) + \lambda \Pi_{(T, \hat{g})}^{i}(\sigma^2)
\]
\[
= \Pi_{(M, \mathcal{X})}^{i}(\sigma^2; t_i),
\]

where the inequality follows from the Bayesian incentive compatibility of \( \langle T, f \rangle \) and the strict dominant strategy incentive compatibility of \( \langle T, \hat{g} \rangle \).

Case 3b (ii). \( \sigma^1_{-i}(t_{-i}) = \sigma^2_{-i}(t_{-i}) \neq t_{-i} \), for some \( t_{-i} \in B_{-i}^\alpha \). Then there is
\[
\hat{\sigma} = \{(\hat{\tau}_i), \hat{s}_i, t_i, t_{-i}, \tau_{-i} \}
\]
\[
\hat{\sigma} = \{(\hat{\tau}_i), \hat{s}_i, t_i, t_{-i}, \tau_{-i} \}
\]
\[
\hat{\sigma} = \{(\hat{\tau}_i), \hat{s}_i, t_i, t_{-i}, \tau_{-i} \}
\]
\[
\hat{\sigma} = \{(\hat{\tau}_i), \hat{s}_i, t_i, t_{-i}, \tau_{-i} \}
\]

Let
\[
B_{(i, j)}^\alpha = \{t_{-i, j} \in \prod_{l \neq i, j} T_l : (t_{-i, j}, t_j) \in B_{-i, j}^\alpha \}.
\]

Then, by \( G_{-i}(B_{-i}^\alpha | t_i) = 1 \), we must have \( G_{-i}(B_{-i}^\alpha | t_i) > 0 \).

Define
\[
A_{-i}^\alpha = \{i_t : \sigma_i^1(t_i) \neq \sigma_i^2(t_i) \}
\]
\[
B_{-i}^\alpha = \{i_t : \sigma_i^1(t_i) = \sigma_i^2(t_i) \}
\]

which form a partition of \( T \). So we must have either \( G_i(A_{-i}^\alpha | t_i) > 0 \) or \( G_i(A_{-i}^\alpha | t_i) = 0 \) (i.e., \( G_i(B_{-i}^\alpha | t_i) = 1 \)).

When \( G_i(A_{-i}^\alpha | t_i) > 0 \), we define \( \hat{\sigma}_j = (\sigma_j^1, \tau_j, \hat{\sigma}_j, \hat{\sigma}_j^4) \). Then we have
\[
| S(\hat{\sigma}_j, \sigma_{-j}(\hat{\tilde{t}}_j), \sigma_{-j}(\hat{\tilde{t}}_{-j})) | < 1 \text{ for all } \hat{\tilde{t}}_j \in A_{-j}^\alpha \times B_{-j}^\alpha \text{, and } | S(\hat{\sigma}_j, \sigma_{-j}(\hat{\tilde{t}}_{-j})) | = 1 \text{ otherwise. Then, by the strict dominant strategy incentive compatibility of } \langle T, \hat{g} \rangle,
\]

we have
\[
\Pi_{(M, \mathcal{X})}^{i}(\hat{\sigma}_i, \sigma_{-i}; t_i) - \Pi_{(M, \mathcal{X})}^{i}(\sigma^2; t_i)
\]
\[
= \left[ \Pi_{(M, \mathcal{X})}^{i}(\tau_j, \sigma^2; t_j) - \Pi_{(M, \mathcal{X})}^{i}(\sigma^2; t_j) \right]
\]
\[
+ \int_{A_{-i}^\alpha \times B_{-i}^\alpha} \frac{A(1 - \hat{\beta} - \hat{\rho})}{K + 1} u_i(t_i) \cdot (\tilde{x} - e^u) dG_i(t_{-i}, t_i)
\]
\[
> 0
\]
when $\hat{\beta}$ and $\hat{\rho}$ are sufficiently close to $1/2$. So $\sigma \in \Sigma$ is not a Bayesian equilibrium.

When $G_i(B_i \mid t_j) = 1$, let $\hat{\sigma}_i = (\sigma_j^1, \tau_j^2, \sigma_j^3, \sigma_j^4)$. Note that $|S(\hat{\sigma}(t_i), \sigma_j(\hat{\sigma}_{-j}))| \leq 1$ for all $\hat{\sigma}_{-i} \in T$. Then we have

$$
\Pi_{(M,S)}(\hat{\sigma}_i, \sigma_{-j}; t_i) = (1 - 3\lambda) \Pi_{(T,T)}(\sigma^1; t_j) + \lambda \Pi_{(T,T)}(\sigma^2; t_j)
$$

+ $\lambda \Pi_{(T,T)}(\sigma^3; t_j) + \lambda \Pi_{(T,T)}(\sigma^4; t_j)

$$
> (1 - 3\lambda) \Pi_{(T,T)}(\sigma^1; t_j) + \lambda \Pi_{(T,T)}(\sigma^2; t_j)
$$

+ $\lambda \Pi_{(T,T)}(\sigma^3; t_j) + \lambda \Pi_{(T,T)}(\sigma^4; t_j)

$$
= \Pi_{(M,S)}(\sigma; t_j).
$$

where the inequality follows from the strict dominant strategy incentive compatibility of $\langle T, \hat{\sigma} \rangle$. So $\sigma \in \Sigma$ is not a Bayesian equilibrium.

Thus we have proved only those strategies with $\sigma^1 = \sigma^2 = \tau$ are Bayesian equilibria and thus, for every Bayesian equilibrium $\sigma$, $g(\sigma(t)) = g'(t)$ for all $t \in T$. It is clear that $d_{\rho}(g(t), f(t)) \leq \epsilon$ for all $t \in T$. Hence, $f$ is virtually implementable in Bayesian equilibrium.

We now prove necessity. The proof is essentially the same as for Theorem 1. Suppose, by way of contradiction, that there is an $i$ and $\hat{\sigma}_i \neq \tau_i$ such that

$$
\Pi_{(T,T)}(\hat{\sigma}_i, \tau_{-i}; t_i) - \Pi_{(T,T)}(\tau; t_i) = 3\eta
$$

with $\eta > 0$. Since $U(\cdot)$ is bounded, there is a $D > 0$ such that $U_i(a, t_i) \leq D$ for all $a \in A$ and all $t_i \in T$. Let $\epsilon = \eta/D$. Since $\langle T, f \rangle$ is virtually implementable in Bayesian equilibrium, there exists a social choice function $g : T \rightarrow \mathbb{P}(A)$ such that $d_{\rho}(g(t), f(t)) \leq \epsilon$ for all $t \in T$. Consider the direct mechanism $\langle T, g \rangle$. Then

$$
|\Pi_{(T,T)}(\hat{\sigma}_i, \tau_{-i}; t_i) - \Pi_{(T,T)}(\hat{\sigma}_i, \tau_{-i}; t_i)|
$$

$$
= \int_T \int_A U_i(a, t_i)(dx(t) - dg(t)) \ dG_i(t_{-i} \mid t_i)
$$

$$
\leq D \int_T (x(t) - g(t)) \ dG_i(t_{-i} \mid t_i)
$$

$$
\leq \epsilon D \leq \eta.
$$

Similarly, we have $|\Pi_{(T,T)}(\tau; t_i) - \Pi_{(T,T)}(\tau; t_i)| \leq \eta$. But then

$$
\Pi_{(T,T)}(\hat{\sigma}_i, \tau_{-i}; t_i) - \Pi_{(T,T)}(\tau; t_i)
$$

$$
= [\Pi_{(T,T)}(\hat{\sigma}_i, \tau_{-i}; t_i) - \Pi_{(T,T)}(\hat{\sigma}_i, \tau_{-i}; t_i)]
$$

+ $[\Pi_{(T,T)}(\hat{\sigma}_i, \tau_{-i}; t_i) - \Pi_{(T,T)}(\tau; t_i)]

+ $[\Pi_{(T,T)}(\tau; t_i) - \Pi_{(T,T)}(\tau; t_i)]

$$
\geq - \eta + 3\eta - \eta
$$

$$
= \eta > 0.
$$
which contradicts the fact that \( g \) is Bayesian incentive compatible. Therefore, \( f \) must be Bayesian incentive compatible. The proof is completed.

**Remark 4.** Notice that the message space of the mechanism constructed above consists of two categories: agents' type messages and non-type messages. In the mechanism, agents are required to report their private information (types) and at the same time to send some auxiliary 'non-type' messages. Thus, the above mechanism is in the class of so-called augmented revelation mechanisms which was studied by Mookherjee and Reichelstein (1990). We have therefore shown that only truthful reporting (i.e., only for those \( \sigma \) with \( \sigma^1 = \sigma^2 = \tau \)) is a Bayesian equilibrium and any other non-truth-telling strategy is not a Bayesian equilibrium. This approach of constructing an augmented revelation mechanism is standard in studying Bayesian implementation in the literature.

**Remark 5.** Readers may find that all results on virtual implementation consider implementation of a social choice function for random allocations. The reason is mainly because it simplifies the problem by linearizing utility functions. When agents make choices among random allocations, they extend linearly their utility functions over pure allocations to random allocations so that the von Neumann–Morgenstern expected utility functions are linear in random allocations. In our approach, the strict dominant strategy mechanism must be defined over non-degenerate random allocations. An interesting question is whether or not one can consider virtual implementation of a social choice function for pure allocations. The answer is immediate. Even though the social choice function is pure, we can consider a pure allocation as a random allocation which yields that pure allocation with probability 1 and virtually implements it by a non-degenerate random allocation mechanism. Thus, as a corollary of the above theorems, our results cover environments with pure allocations as well.

**Corollary 2.** A social choice function \( f: T \to A \) is virtually implementable in Bayesian equilibrium (resp. dominant strategy equilibrium) if and only if it is incentive compatible in Bayesian equilibrium (resp. dominant strategy equilibrium).

### 4. Virtual implementation with side payments

The results given above are very general and cover many interesting economic models and results in the literature as special cases. Since we allow arbitrary sets of alternatives and types, the results for finite types and/or alternatives, which were given by Abreu and Matsushima (1990), Matsushima (1993) and Duggan (1993), can be obtained as corollaries of ours. Our results can also be used to characterize the virtual implementation of efficient allocations by efficient trading mechanisms which were studied by Myerson (1981), Palfrey and Srivastava.
(1991b), among many others. In this section, we consider virtual implementation of a social choice function with side payments (so that utility functions are quasi-linear) which also covers many interesting economic models. A particularly interesting case is the principle-agent model which in general has quasi-linear utility functions. For example, for a standard adverse selection model involving a principle and $n$ agents, in which the principle seeks to implement some project(s) $y$, agent $i$ obtains a benefit $V_i(y, t_i)$ (or alternatively, bears a cost $-V_i(y, t_i)$) of the project and pays (receives) a monetary transfer tax $r_i$ (payment $-r_i$) for consuming the project when it is implemented. Thus agents' payoff functions are quasi-linear. It is well known that Bayesian (dominant) incentive compatibility mechanisms for such types of utility functions exist. For instance, the Groves mechanism is dominant strategy incentive compatible for complete information environments and yields the efficient level of the project (cf. Groves (1973), Groves and Loeb (1975), Green and Laffont (1979)). However, the Groves mechanism in general does not satisfy the budget balancing constraint and is not ex-post (Pareto) efficient (cf. Walker (1978, 1980), Hurwicz and Walker (1990), Tian (1996b)).

By using the Bayesian approach, d'Aspremont and Gérard-Varet (1979) proposed to obtain balanced and outcome efficient mechanisms by using interim (expected) Bayesian incentive compatibility, and thus their mechanism yields interim (Pareto) efficient allocations. Although these mechanisms are incentive compatible, they may not be implementable since they may contain some undesirable equilibrium outcomes. More generally, one needs to design alternative mechanisms which (virtually) implement a social choice goal. Duggan (1993) considered virtual implementation for models with side payments. However, he considered virtual implementation only for finite types of alternatives.

We now consider virtual implementation of a social choice function with side payments. The mechanisms to be constructed allow budget balancing. To do this, let $Y$ be a set of feasible projects which is a metric space. We assume that agent $i$'s utility function is given by a quasi-$\epsilon$-near utility function:

$$U_i(y, r_i, t_i) = V_i(y, t_i) + r_i,$$

where $y \in Y$. $V_i(y, t_i)$ is the value function with $V_i(y^0, t_i) = 0$, and $r_i \in R$ is the transfer of agent $i$. Let $\mathbb{P}(Y)$ be the set of random allocations and $Y^K$ be a finite $K$-element subset of $Y$. Elements of the set $Y^K$, together with $y^0$, are denoted by $\{y^0, y^1, y^k, \ldots, y^K\}$. Let

$$v_{i}(t_i) = V_i(y^k, t_i) \quad k = 0, 1, \ldots, K.$$ 
$$v_{i}(t_i) = (v_{i}^0(t_i), v_{i}^1(t_i), \ldots, v_{i}^{K}(t_i)).$$

Assumption 2. There is a $K$-element subset $Y^K \subset Y$ with $K > 1$ such that

1. for each $t_i \in T_i$, $v_{i}(t_i) \neq 0$ for at least two elements in $Y^K$,
2. $v_{i}(t_i) \neq v_{i}(\hat{t}_i)$ for all $t_i \neq \hat{t}_i$. Denote by $\mathbb{P}(Y^K)$ the set of all random allocations which satisfy Assumption 2 and are distributed only over $Y^K$. 

Lemma 2. There exists a direct mechanism \( \langle T, \hat{g}, \hat{r} \rangle \) with \( \langle \hat{g}, \hat{r} \rangle : T \to \mathcal{P}(\mathbb{A}^K) \times \mathbb{R}^n \), such that the truthful strategy profile, \( \tau \), is a strict dominant strategy equilibrium of the game induced by \( \langle T, \hat{g}, \hat{r} \rangle \).

Proof: Define the direct mechanism \( \langle T, \hat{g}, \hat{r} \rangle \) as follows. For each reported type \( \hat{t} \in T \), the probability assigned to allocation \( a^k \) \((k = 0, 1, \ldots, n)\) is

\[
\hat{g}^0(\hat{t}) = \frac{1}{2} - \frac{1}{2nK} \sum_{i=1}^{n} \sum_{k=1}^{K} p(v^k_i(\hat{t}))
\]

\[
\hat{g}^k(\hat{t}) = \frac{1}{2} K + \frac{1}{2nK} \sum_{i=1}^{n} p(v^k_i(\hat{t}))
\]

for \( k \geq 1 \). Here

\[
p(v^k_i(\hat{t})) = \begin{cases} 
\frac{v^k_i(\hat{t})}{1 + v^k_i(\hat{t})} & \text{if } v^k_i(\hat{t}) > 0 \\
\frac{v^k_i(\hat{t})}{1 - v^k_i(\hat{t})} & \text{if } v^k_i(\hat{t}) \leq 0
\end{cases}
\]  

(7)

The transfer of agent \( i \) is defined by

\[
\hat{t}_i(\hat{t}) = -\sum_{k=1}^{K} \left[ \gamma(p(v^k_i(\hat{t}))) - \lambda(p(v^{k+1}_i(\hat{t}_{i+1}))) \right]
\]  

(8)

where \( n + 1 \) is regarded as 1 and \( \gamma(p) \) is defined by

\[
\gamma(p) = \begin{cases} 
-\ln(1 - p) - p & \text{if } p > 0 \\
-\ln(1 + p) + p & \text{if } p \leq 0
\end{cases}
\]  

(9)

Thus, the mechanism is feasible which allows a balanced budget for all strategy profiles, and

\[
\hat{O}(\hat{g}(\hat{t}), t_i) = \int_{\mathbb{A}^K} \mathcal{V}((a, t_i)) \, d\hat{g} + \hat{r}_i(\hat{t})
\]

\[
= v_i(t_i) \cdot \hat{g}(\hat{t}) + \hat{r}_i(\hat{t})
\]

\[
= \sum_{k=1}^{K} v_i^k(t_i) \left( \frac{1}{2K} + \frac{1}{2nK} \sum_{i=1}^{n} p(v^k_i(\hat{t})) \right) + \hat{r}_i(\hat{t})
\]

\[
= \frac{1}{2nK} \sum_{k=1}^{K} \left( v_i^k(t_i, p(v^k_i(\hat{t}))) - \lambda(p(v^k_i(\hat{t}))) \right)
\]

\[
+ \sum_{k=1}^{K} v_i^k(t_i) \left( \frac{1}{2K} + \frac{1}{2nK} \sum_{j \neq i}^{n} p(v^k_i(\hat{t})) + \gamma(p(v^k_{i+1}(\hat{t}_{i+1}))) \right)
\]
where \( C'(\hat{t}_{-i}) \) is some term which does not depend on \( \hat{t}_i \).

By the strict convexity of \( \gamma(\cdot) \), one can verify that the unique solution of

\[
\max_{p_i \in \mathbb{R}^K} \sum_{k=1}^{K} \left[ v_i^k(t_i) p_i^k - \gamma(p_i^k) \right]
\]

is that \( p_i = v_i(t_i) \). Thus, by monotonicity of \( p(v_i^k) \), \( \hat{U}_i(\hat{\sigma}(t_i, \hat{t}_{-i}), t_i) > \hat{U}_i(\hat{\sigma}(t_i, \hat{t}_{-i}), t_i) \) for all \( \hat{t}_i \neq t_i \). Notice that, as before, when the set of types consists only of the set of utility functions of agents and each agent is asked to report his utility functions, this inequality still holds even without the assumption \( v_i(t_i) \neq v_i(\hat{t}_i) \) for all \( t_i \neq \hat{t}_i \).

Thus we have \( \sum_{t_i \in T_i} \hat{U}_i(x(t), t_i) dG_i(t_{-i} | t_i) > 0 \) for all \( A_{-i} \) with \( G_i(A_{-i} | t_i) > 0 \), and particularly

\[
\hat{U}_i(\hat{\sigma}(t_i, \hat{t}_{-i}), t_i) > \hat{U}_i(\hat{\sigma}(\hat{t}_i), t_i)
\]

for all \( \hat{t}_i \neq t_i \), and thus \( \prod_{(T, \hat{\sigma})} \hat{U}_i(\hat{\sigma}; t_i) > \prod_{(T, \hat{\sigma})} \hat{U}_i(\hat{\sigma}; t_i) \) for all \( \sigma_i \neq \tau_i \). Therefore, the truthful strategy profile \( \tau \) is a strict dominant strategy equilibrium of the game induced by \( \langle T, \hat{\sigma}, \hat{\tau} \rangle \) which is defined on \( Y^K \). The proof is completed.

By applying this lemma, we have

**Theorem 3.** A social choice function \( (f, \lambda) : T \rightarrow \mathbb{R}^{Y \times R^n} \) is virtually implementable in Bayesian equilibrium (resp. dominant strategy equilibrium) if and only if it is Bayesian (resp. dominant strategy) incentive compatible.

**Proof:** The proof for the dominant strategy incentive compatibility is the same as for Theorem 1. We prove the theorem for the Bayesian incentive compatibility only for \( n \geq 2 \). The proof for necessity is also the same as for Theorem 2. We now prove sufficiency. For any \( \varepsilon > 0 \), let \( \lambda = \min \{ \varepsilon / 2, 1 / 2 \} \). Define a mechanism \( \langle M^*, \hat{g}^*, r^* \rangle \) as follows. Let \( M^*_i = T_i \times T_i \times (0, 1) \). For each message \( m = (m^1, m^2, m^3) \) with \( m^3 := \beta_i \), define the outcome function by

\[
g^*(m) = (1 - 2\lambda)f(m^1) \oplus \lambda \hat{g}(m^1) \oplus \lambda \hat{g}(m^2)
\]

and

\[
r^*_i(m) = \begin{cases} (1 - 2\lambda)r_i(m^1) + \lambda \hat{r}_i(m^1) + \lambda \hat{r}_i(m^2) & \text{if } S(m) = \{i\} \text{ or } \emptyset \\ (1 - 2\lambda)r_i(m^1) + \lambda \hat{r}_i(m^1) + \lambda \hat{r}_i(m^2) + \lambda (\beta_i - \beta_{i+}) \\ + \lambda (\beta_{i+ } - \beta_{i++)} \delta_{i+1}(m) & \text{otherwise} \end{cases}
\]
where \( S(m) = \{ j \in N : m^j_1 \neq m^j_2 \} \) and
\[
\delta_{i+1}(m) = \begin{cases} 
1 & \text{if } S(m) = \{ i + 1 \}, \\
0 & \text{otherwise}
\end{cases}
\] (13)

Here \( n + 1 \) and \( n + 2 \) are regarded as 1 and 2, respectively. One can see that \( r_\sigma^* \) is balanced for all \( \sigma \in \Sigma \). Also \( d_\rho(g^* (\sigma(t)), f(\sigma^1(t))) \leq \epsilon \) and \( \| r^*(\sigma(t)) - r(\sigma^1(t)) \| \leq \epsilon \) for all \( \sigma \in \Sigma \) and all \( t \in T \). We now show that only those strategies with truth-telling are Bayesian equilibria and thus this mechanism implements the social choice function \((\bar{g}, \bar{r}): T \to P(A) \times R^n\), which are defined by \( \bar{g}(t) = (1 - 2\lambda)f(t) \oplus 2\lambda\tilde{g}(t) \) and \( \bar{r}_m(t) = (1 - 2\lambda)\tau_1(t) + 2\lambda\tilde{r}_m(t) \) for all \( t \in T \).

Again, we can prove this by considering three cases for a strategy profile \( \sigma = (\sigma^1, \sigma^2, \sigma^3) \in \Sigma \): (1) \( \sigma^1 = \sigma^2 = \tau \), (2) \( \sigma^1 = \sigma^2 \neq \tau \), and (3) \( \sigma^1 \neq \sigma^2 \).

The proofs for Cases 1 and 2 are the same as those in Theorem 2. We need only to prove Case 3 in which any \( \sigma \in \Sigma \) with \( \sigma^1 \neq \sigma^2 \) is not a Bayesian equilibrium.

For each \( i \in N \), define
\[
A^i_\sigma = \{ t_i \in T_i : \sigma^1(t_i) \neq \sigma^2(t_i) \}.
\]

Case 3 can be divided into two subcases: (a) There is \( j \in N \) such that \( G_j(A^i_\sigma) > 0 \), and (b) \( G_j(A^i_\sigma) = 0 \) for all \( j \in N \), and there is \( i \in N \) and \( t_i \in T_i \) such that \( \sigma^1(t_i) \neq \sigma^2(t_i) \).

Case 3a. There is \( j \in N \) such that \( G_j(A^i_\sigma) > 0 \). Take \( i \neq j \) and let \( \hat{\delta}_i = (\sigma^1_i, \sigma^2_i, \hat{\sigma}_i^3) \) with \( \hat{\sigma}_i^3 = \beta_1 < \beta_2 \). Note that we have \( S(\hat{\delta}_i(t_i), \sigma^1(t_\cdot)) = S(\sigma(t)) \neq \{ i \} \) for all \( t \in A^i_\sigma \times T_\cdot \). Thus we have
\[
\Pi^i_{(M+\cdots;i)}(\hat{\delta}_i, \sigma^1; t_i) - \Pi^i_{(M+\cdots;i)}(\sigma; t_j)
\]
\[
\geq \int_{A^i_\sigma \times T_{-\{i,j\}}} \lambda(\beta_i - \hat{\beta}_i) dG(t_\cdot | t_i)
\]
\[
= \lambda(\beta_i - \hat{\beta}_i)G(A^i_\sigma \times T_{-\{i,j\}} | t_i) > 0.
\]

So \( \sigma \in \Sigma \) is not a Bayesian equilibrium.

Case 3b. \( G_j(A^i_\sigma) = 0 \) for all \( j \in N \) and there is \( i \in N \) and \( t_i \in T_i \) such that \( \sigma^1(t_i) \neq \sigma^2(t_i) \). This can be divided into two subsubcases: (i) \( \sigma^1(t_\cdot) = \sigma^2(t_\cdot) = t_\cdot \) for all \( t_\cdot \in T_{-\cdot} \), and (ii) \( \sigma^1(t_\cdot) = \sigma^2(t_\cdot) \neq t_\cdot \) for some \( t_\cdot \in T_{-\cdot} \).

Case 3b(i). \( \sigma^1(t_\cdot) = \sigma^2(t_\cdot) = t_\cdot \) for all \( t_\cdot \in T_{-\cdot} \). The proof is the same as for Case 3b(i) in Theorem 2.

\(^5\) It can be verified by considering three cases: \( |S(m)| = 0 \), \( |S(m)| = 1 \), and \( |S(m)| \geq 2 \).
Case 3b(ii). \( \sigma_1(\tau_{-i}) = \sigma_2(\tau_{-i}) \neq t_{-i} \) for some \( t_{-i} \in T_{-i} \). Then there is \( j \in N \) with \( j \neq i \), and \( t_j \in T_j \) such that \( \sigma_j^1(t_j) = \sigma_j^2(t_j) \neq t_j \). Let \( \theta_j = (\sigma_j^1, \sigma_j^2, \sigma_j^3) \). Then we have

\[
\Pi^i_{\langle M^*, \bar g^*, r^* \rangle}(\theta_j, \sigma_{-j}; t_j) = (1 - 2 \lambda) \Pi^i_{\langle T, f, r \rangle}(\sigma^1; t_j) + \lambda \Pi^i_{\langle T, \bar g, \bar r \rangle}(\sigma^1; t_j)
\]

\[
> (1 - 2 \lambda) \Pi^i_{\langle T, f, r \rangle}(\sigma^1; t_j) + \lambda \Pi^i_{\langle T, \bar g, \bar r \rangle}(\sigma^2; t_j)
\]

\[
= \Pi^i_{\langle M^*, \bar g^*, r^* \rangle}(\sigma^1; t_j),
\]

where the inequality follows from the strict dominant strategy incentive compatibility of \( \langle T, \bar g \rangle \). So \( \sigma \in \Sigma \) is not a Bayesian equilibrium. Thus we have proved that only those strategies with truth-telling are Bayesian equilibria and therefore, for every Bayesian equilibrium \( \sigma, g^*(\sigma(t)) = \bar g(t) \) and \( r_j^*(\sigma(t)) = \bar r_j(t) \) for all \( t \in T \). It is clear that \( d_j(\bar g(t), f(t)) \leq \varepsilon \) and \( \| \bar r(t) - r(t) \| \leq \varepsilon \). Thus \((f, r)\) is virtually Bayesian implementable.

5. Common values

The most significant assumption in the above discussion is private values. Although a large majority of applications using the Bayesian approach to economic problems and applications of the revelation principle of mechanism design have used this assumption, it is still restrictive. For instance, to invoke the private valuation assumption in most trading mechanism design papers (such as Palfrey and Srivastava (1991b) and references therein) rules out markets whose valuation structure generates externalities between traders. When these externalities are present, one’s willingness to trade a commodity depends not just on one’s own idiosyncratic preferences but also on the preferences of other traders. For such valuation externalities, Myerson (1981) first gave an efficient trading mechanism, and Gresik (1991) characterized the set of Bayesian incentive compatible efficient trading mechanisms. However, efficient allocations are in general not implementable because the efficient allocations are in general not Bayesian monotonic, and thus mechanisms yielding efficient allocations may also contain some undesirable equilibrium allocations. Then one may want to ask: what happens if one considers virtual implementation instead of considering exact implementation for economic environments with common values? Can we obtain similar results for common valuation economic environments?

Our results in general do not apply with nearly the same force in a setting with common values. In this case, incentive compatibility may not be a sufficient condition for virtual implementation in situations of incomplete information. When
Matsushima (1993) considered virtual implementation in Bayesian equilibrium and Fang (1993) considered virtual implementation in dominant strategy equilibrium, they both assumed the strict incentive compatibility condition with other additional conditions (such as finite types, quasi-linear utility functions, etc.). Even so, in some special cases we can get the same results, which we now discuss. For example, we can characterize the virtual implementation for efficient trader mechanisms in markets containing both types of commodities: commodities without valuation externalities and commodities with valuation externalities.

The framework itself is easily modified to incorporate common values. To do this, we write the utility function of agent \( i \) as \( U_i(a, t) \) instead of \( U_i(a, t_i) \). The definition of incentive compatibility is the same as before.

The following theorem shows that, even for common valuation economic environments, as long as there are no valuation externalities for at least two alternatives, we can get the same results. The proof of the theorem is the same as for Theorems 1 and 2.

**Theorem 4.** For common valuation economic environments, a necessary condition for a social choice function \( f : T \rightarrow P(A) \) to be virtually implementable in Bayesian equilibrium (resp. dominant strategy equilibrium) is that it is Bayesian incentive compatible (resp. dominant strategy equilibrium). It becomes sufficient if there are at least two alternatives which do not have valuation externalities (i.e., if there is a \( K \)-element set \( A^K \subset A \) with \( K \) greater than 1 such that \( U_i \) are independent of \( T_{-i} \) on \( A^K \) for all \( i \in N \)).

### 6. Conclusion

In this paper we have considered virtual implementation in both dominant strategy equilibrium and Bayesian strategy equilibrium for incomplete information environments with infinite alternatives and types. Our results stand in sharp contrast to previous results in the literature. We allow for general utility functions, any number of agents, and arbitrary sets of alternatives and types. In addition, our mechanisms with transfers satisfy the budget balance requirement. We characterize virtual implementation of a social choice function by giving necessary and sufficient conditions. We show that, under some technical conditions, a social function is virtually implementable in Bayesian (resp. dominant strategy) equilibrium if and only if it is Bayesian (resp. dominant strategy) incentive compatible. We obtain the same results for models with side payments.

These results not only show the virtual implementability of a general social choice function, but also give a simple way to construct an implementable social choice function which is arbitrarily close to the original social choice function. To solve the multiple equilibria problem of the original social choice function, one can use an implementable social choice function which is simply a convex
combination (probability mixture) of the original social choice function and the outcome function \(\hat{f}\) obtained in the paper. This canonical revelation outcome function is defined only on a finite set of alternatives and it has the property that truth-telling is a strict dominant strategy equilibrium of the game reduced by this outcome function. Thus, solving the multiple equilibria problem has been simply reduced to the problem of constructing this canonical revelation mechanism and then forming the probability mixture.

We also consider the virtual implementability for some types of common-value economic environments, and show that incentive compatibility is the necessary and sufficient condition for a class of economic environments where there are at least two alternatives that have no valuation externalities. The full characterization on virtual implementation for general common-values models remains to be done.

Acknowledgements

I wish to thank two anonymous referees for very helpful comments and suggestions. Financial support from the Private Enterprise Research Center and Texas A & M University’s Program to Enhance Scholarly and Creative Activities is gratefully acknowledged.

References


Duggan, J., 1993, Virtual implementation in Bayesian equilibrium with infinite types, Parts I and II, mimeo.

Fang, J., 1993, Approximate implementation in dominant strategy equilibrium, mimeo.


