CIRJE-F-287

On Detail-Free Mechanism Design and Rationality

Hitoshi Matsushima
The University of Tokyo

June 2004

On Detail-Free Mechanism Design and Rationality*

Hitoshi Matsushima+

June 28, 2004

Abstract

The study of mechanism design is sometimes criticized, because the designed mechanisms depend on the fine detail of the model specification, and agents' behavior relies on the strong common knowledge assumptions on their rationality and others. Hence, the study of 'detail-free' mechanism design with weak informational assumptions is the most important to make as the first step towards a practically useful theory. This paper will emphasize that even if we confine our attentions to detail-free mechanisms with week rationality, there still exist a plenty of scope for development of new ideas on how to design a mechanism to play the powerful role. We briefly explain my recent works on this line, and argue that the use of stochastic decision works much in large exchange economics, and agents' moral preferences can drastically improve implementability of social choice functions.

JEL Classification Numbers: C72, D41, D44, D78

* This paper was prepared for the 2004 JEA Nakahara Prize Lecture presented at the annual meeting of the Japanese Economic Association at Okayama University on 25-26 September 2004. The research for this paper was supported by Grant-In-Aid for Scientific Research (KAKENHI 15330036) from JSPS and MEXT of the Japanese Government.

+ Faculty of Economics, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan. Fax: +81-3-3818-7082. E-Mail: hitoshi@e.u-tokyo.ac.jp

1. Introduction

The study of mechanism design such as auction, implementation, and contract theory, which is one of my main research fields[1] besides repeated games, has been pervasive in application of game theory to economics. In order to achieve the desirable allocation, the central planner, or the auctioneer, designs a decentralized mechanism such as an auction scheme, a trading procedure, and a voting scheme in advance. The designed mechanism is assumed to be common knowledge among agents, or traders, who play equilibrium behavior in the game induced by this mechanism.

Many previous works in the mechanism design literature, however, depend on, not only this common knowledge assumption, but also the following unrealistic informational assumptions. In order to design a well-behaved mechanism, the central planner has to know the fine detail of the model specification such as agents' utility functions, the prior distribution, and the social choice function. In order for agents to calculate their equilibrium behavior, not only the model specification, but also agents' rationality, has to be common knowledge among agents. As many authors such as Hurwicz (1972), Wilson (1985, 1987), and Dasgupta and Maskin (2000) have pointed out as a criticism on mechanism design theory, these assumptions make the mechanisms difficult to put into practice.

Based on this observation, we can say that how to design mechanisms that are *detail-free*, i.e., do not depend on the detail of the model specification, and induce agents to play desirable behavior without requiring any restrictive knowledge assumption is the most important to investigate as the first step towards a practically useful theory. The purpose of the present paper is to show that even if we confine our attentions to the class of detail-free mechanisms, there might still be a plenty of scope for ingenious researchers to develop new ideas of mechanism design that can drastically improve economic welfare.

The present paper will explain two recent researches of mine, introducing two new ideas on how to design detail-free mechanisms with weak rationality assumptions. First, we will argue that in implementing any social choice function, it is quite useful from the practical viewpoint to take agents' moral preferences into account, in addition to their material interests. Whenever any social choice function is implementable in Nash equilibrium, then it must be contingent only on factors relevant to agents' preferences. This might sometimes be quite restrictive, especially in the case that only agents' material interests are relevant to their preferences. Hence, in order to implement many non-trivial social choice functions, we should take aspects of agents' moral preferences into account besides their material interests.

This point was firstly pointed out by my paper entitled "Universal Mechanisms and Moral Preferences in Implementation" (Matsushima (2003)). Matsushima (2003) showed that agents' moral preferences for *honesty* play a powerful role in, not only implementing a wide variety of social choice functions, but also making the mechanisms detail-free in a very strict sense. Section 4 will explain Matsushima (2003) in more detail.

---

[1] For instance, see Abreu and Matsushima (1992a, 1992b, 1994) and Matsushima (1988, 1990, 1991a, 1991b, 1993).

Second, we will argue that *stochastic* decision plays a powerful role in large double auction environments. According to the well-specified detail-free auction format, traders are randomly divided into distinct groups. The members of each group trade at the market-clearing price vector in *another* group. Hence, each agent's activity in the auction game never influences the price vector at which she trades. This will be the driving force of agents' having incentive to adopt price-taking behavior.

This idea of stochastic auction design was firstly explored by my paper entitled "Large Double Auction Design in Dominance" (Matsushima (2004)). Since traders with private values adopt price-taking behavior as dominant strategies, it follows from the law of large numbers that the auctioneer can achieve the perfectly competitive allocations in the limit as the number of traders grows. Of particular importance, in the interdependent value case, the auctioneer can achieve the rational expectations equilibrium allocation. Here, traders adopt price-taking behavior as twice iteratively undominated strategies, which implies that we do not need to assume any strong rationality assumption. Sections 5 and 6 will explain this paper in more detail.

The next section shows the basic model of implementation. Section 3 shows a brief survey on the background. This paper, however, does not intend to give a comprehensive survey on the general literature of implementation and auction theory.

2. Basic Model

This section introduces the basic model of implementation.[2] Let N denote the finite set of agents. Let A denote the finite set of pure alternatives. Let Δ denote the set of lotteries over pure alternatives.

A mechanism is defined as $G=(M,g)$, where $M=\prod_{i\in N} M_i$, M_i is the finite set of messages for agent $i\in N$, and

$$g: M \to \Delta.$$

The central planner designs a mechanism $G=(M,g)$ in advance, and then requires each agent $i\in N$ to announce her message $m_i \in M_i$. When agents announce message profile $m\in A$, the central planner chooses any pure alternative $a\in A$ with probability $g(m)(a)$.

Agent $i's$ utility function is defined as a function

$$u_i : A\times M_i \to R,$$

which satisfies the expected utility hypothesis. In Section 4, we allow each agent's announcement to have *intrinsic* value for her welfare. For instance, each agent may have *moral* preference in that she has positive physiological cost of announcing *dishonestly*. When agent $i's$ announcement has no such intrinsic values, we will simply write $u_i(a)$ instead of $u_i(a,m_i)$. Let $U_i(M_i)$ denote the set of possible utility functions for agent i. Let $U(M)=\prod_{i\in N} U_i(M_i)$ and $u=(u_i)_{i\in N}\in U(M)$. Two utility function profiles $u\in U(M)$ and $u'\in U(M)$ are said to be *preference-equivalent* if for every $i\in N$, there exist $\beta_i>0$ and $\gamma_i\in R$ such that $u_i'=\beta_i u_i+\gamma_i$.

A combination of a mechanism G and a utility function profile $u\in U(M)$, i.e., (G,u), defines a game. A message profile $m\in M$ is said to be *dominant* in (G,u) if for every $i\in N$ and every $m'\in M$,

$$u_i(g(m_i,m_{-i}'),m_i)\geq u_i(g(m'),m_i').$$

A message profile $m\in M$ is said to be a *Nash equilibrium* in (G,u) if for every $i\in N$ and every $m_i'\in M_i$,

$$u_i(g(m),m_i)\geq u_i(g(m_i',m_{-i}),m_i').$$

We introduce the Bayesian framework for our model as follows. Let $\Omega=\prod_{i\in N}\Omega_i$ denote the finite set of states, where Ω_i is the set of agent $i's$ private signals. Let $W_i(M_i,\Omega)$ denote the set of functions $w_i:\Omega\to U_i(M_i)$. When the state is ω, agent

[2] For surveys on implementation, see Moore (1992), Palfrey (1992), Osborne and Rubinstein (1994, Chapter 10), Mas-Colell, Whinston, and Green (1995, Chapter 23), and Maskin and Sjöström (2002).

$i's$ utility function is given by $w_i(\omega) \in U_i(M_i)$. Let $W(M,\Omega) = \prod_{i \in N} W_i(M_i,\Omega)$ and $w = (w_i)_{i \in N} \in W(M,\Omega)$. We will say that agent i has *private values* if $w_i(\omega)$ does not depend on the other agents' private signals $\omega_{-i} \in \Omega_{-i}$. In this case, we will simply write $w_i(\omega_i)$ instead of $w_i(\omega)$. We will say that agent i has *interdependent values* if $w_i(\omega)$ depends on $\omega_{-i} \in \Omega_{-i}$.

A *social choice function* is defined as a function $f : \Omega \to \Delta$. We will say that a social choice function f is *deterministic* if for every $\omega \in \Omega$, the lottery $f(\omega)$ is degenerate, i.e.,

$$f(\omega)(a) = 1 \text{ for some } a \in A.$$

In this case, we will simply write $f(\omega) = a$ instead of $f(\omega)(a) = 1$. Otherwise, we will say that it is *stochastic*. Let $F(\Omega)$ denote the set of social choice functions. For every positive real number $\varepsilon > 0$, two social choice functions $f \in F(\Omega)$ and $\tilde{f} \in F(\Omega)$ are said to be $\varepsilon-$*close* if for every $\omega \in \Omega$ and every $a \in A$,

$$\left| f(\omega)(a) - \tilde{f}(\omega)(a) \right| \leq \varepsilon.$$

The *direct* mechanism associated with the social choice function $f \in F(\Omega)$ is denoted by $G(f) = (M, g)$ where

$$M_i = \Omega_i \text{ and } g = f.$$

Let $\psi : \Omega \to [0,1]$ denote a *common prior*, where the state ω occurs with probability $\psi(\omega)$. Let Ψ denote the set of common priors with full supports. A combination of a mechanism G, state-contingent utility function profile $w \in W(M,\Omega)$, and a common prior $\psi \in \Psi$, i.e., (G, w, ψ), defines a *Bayesian* game. A *strategy* for agent $i \in N$ is defined as a function $s_i : \Omega_i \to M_i$. We denote by $S_i(M_i, \Omega_i)$ the set of strategies for agent i. Let $S(M,\Omega) = \prod_{i \in N} S_i(M_i,\Omega_i)$ and $s = (s_i)_{i \in N} \in S(M,\Omega)$.

We denote by

$$w_i(s) = \sum_{\omega \in \Omega} w_i(\omega)(g(s(\omega)), s_i(\omega_i), \omega)\psi(\omega)$$

the *ex-ante* expected utility for agent i when agents play strategy profile $s \in S(M,\Omega)$. A strategy profile $s \in S(M,\Omega)$ is said to be a *Bayesian Nash equilibrium* in (G, w, ψ) if for every $i \in N$ and every $s_i' \in S_i(M_i, \Omega_i)$,

$$w_i(s) \geq w_i(s_i', s_{-i}).$$

A strategy $s_i \in S_i(M_i, \Omega_i)$ for agent i is said to be *undominated* in (G, w, ψ) if there exists no $s_i' \in S_i(M_i, \Omega_i)$ such that for every $s_{-i} \in S_{-i}(M_{-i}, \Omega_{-i})$,

$$w_i(s) \leq w_i(s_i', s_{-i}),$$

and the strict inequality holds for some $s_{-i} \in S_{-i}(M_{-i}, \Omega_{-i})$.

Fix a Bayesian game (G, w, ψ) arbitrarily. We define an infinite sequence of subsets of strategy profiles, denoted by S^0, S^1, S^2, …, as follows. Let $S^0 = S$. Let S^1 denote the set of undominated strategy profiles. For every integer $r \geq 2$ and every $i \in N$, let S_i^r denote the set of strategies $s_i \in S_i^{r-1}$ for agent i such that there exists no strategy $s_i' \in S_i^{r-1}$ for agent i such that for every strategy profile for the other agents $s_{-i} \in S_{-i}^{r-1}$,

$$w_i(s) \leq w_i(s_i', s_{-i}),$$

and the strict inequality holds for some $s_{-i} \in S_{-i}^{r-1}$. Let $S^\infty = \lim_{r \to \infty} S^r$. A strategy profile $s \in S(M, \Omega)$ is said to be *iteratively undominated* in (G, w, ψ) if

$$s \in S^\infty.$$

The set of iteratively undominated strategy profiles S^∞ is said to be *twice dominance solvable* in (G, w, ψ) if

$$S^2 = S^\infty.$$

When S^∞ is twice dominance solvable, in order to calculate S_i^∞ for each agent $i \in N$, she needs to know only that the other agents never play undominated strategies.

The set of iteratively undominated strategy profiles S^∞ is said to be *interchangeable* in (G, w, ψ) if every iteratively undominated strategy profile $s \in S^\infty$ is a Bayesian Nash equilibrium in (G, w, ψ). If S^∞ is interchangeable, then agents can reach a set of Bayesian Nash equilibria only by iteratively eliminating dominated strategies.

3. Implementation: Definitions and Literatures

This section shows several definitions and the background of implementation.

3.1. Dominant Strategy

A social choice function f is said to be *strategy-proof* for $w \in W(M,\Omega)$ if for every $\omega \in \Omega$, truth-telling is a dominant strategy profile in $(G(f), w(\omega))$, i.e., for every $i \in N$ and every $\omega' \in \Omega$,

$$w_i(\omega)(f(\omega_i, \omega'_{-i}), \omega_i) \geq w_i(\omega)(f(\omega'), \omega'_i) .$$

Most of the previous works on strategy-proofness have studied the case where all agents have private values, their announcements have no intrinsic values, and different private signals correspond to different preferences over pure alternatives. The use of dominant strategies in the implementation literature might be quite appropriate, because by definition it does not depend on the model specification, and we need no common knowledge assumption on rationality.

The seminal works by Gibbard (1973) and Satterthwaite (1975), however, by investigating the general social choice environments with the rich preference domain and with three or more alternatives, showed the impossibility theorem that if a social choice function is strategy-proof and deterministic, then it must be *dictatorial* in the sense that there exists an agent $i \in N$ such that for every $\omega \in \Omega$,

$$w_i(\omega_i)(f(\omega)) \geq w_i(\omega_i)(a) .$$

A sketch on the proof of the Gibbard-Satterthwaite theorem, which is available in the text books such as Mas-Colell, Whinston, and Green (1995, Chapter 21), is as follows. First, we show that any strategy-proof social choice function, irrespective of whether it is deterministic or stochastic, must be efficient, and *monotonic* in the sense that for every $\omega \in \Omega$ and $\omega' \in \Omega$, if for every $i \in N$ and every $\alpha \in \Delta$,

$$[\, w_i(\omega_i)(f(\omega)) \geq w_i(\omega_i)(\alpha)\,] \Rightarrow [\, w_i(\omega'_i)(f(\omega)) \geq w_i(\omega'_i)(\alpha)\,],$$

then $f(\omega') = f(\omega)$ must hold. Second, we show that if a social choice function is deterministic, efficient, and monotonic, then it must be dictatorial. This second step of the proof is regarded as the social-choice-function version of the Arrow's impossibility theorem (See Mas-Colell, Whinston, and Green (1995, Chapter 21)).

Matsushima (1988), however, showed that every stochastic social choice function, the values of which always have the full support, is monotonic, and therefore, almost every stochastic social choice function is monotonic. Hence, we cannot directly apply the above sketch to show whether any *stochastic* social choice function to be strategy-proof.

Gibbard (1977) and Benoit (2003) showed that there exists no non-trivial stochastic social choice function that satisfies strategy-proofness or any related requirements. Hence, in general social choice environments, it is quite hard to find non-trivial social choice functions that are strategy-proof, even if we take stochastic decision into account. In fact, it is clear from the definition of strategy-proofness that if a deterministic social choice function is not strategy-proof, then there exists a small but positive real number $\varepsilon > 0$ such that no $\varepsilon-$close stochastic social choice function is strategy-proof. As Matsushima (1988) has showed, we must note that almost every $\varepsilon-$close stochastic social choice function is monotonic, but not strategy-proof.

In contrast to these negative results, the former part of Matsushima (2004) showed that stochastic decision plays a powerful role when we confine our attention to the economic environments with quasi-linearity. For every $\varepsilon \in [0,1]$, a mechanism G is said to *virtually* $\varepsilon-$*implement a lottery* $\alpha \in \Delta$ *in dominance with respect to* $u \in U(M)$ if there exists a dominant message profile in (G,u), and any dominant message profile $m \in M$ in (G,u) satisfies

$$|g(m)(a) - \alpha(a)| \leq \varepsilon \text{ for all } a \in A.$$

Here, we require all dominant strategy profiles to achieve the lottery α approximately. Matsushima (2004) designed a double auction mechanism that is detail-free, i.e., does not depend on the detail of the model specification. Matsushima then showed that this mechanism always virtually implements the perfectly competitive allocation in very general private value cases when the number of traders is sufficiently large. Section 5 will show a brief explanation on the former part of this paper.

3.2. Nash Equilibrium

For every $\varepsilon \in [0,1]$, a social choice function $f \in F(\Omega)$ is *virtually* $\varepsilon-$*implementable with respect to* $w \in W(M,\Omega)$ if there exists a mechanism G such that for every $\omega \in \Omega$, there exists a Nash equilibrium in $(G(f), w(\omega))$, and every Nash equilibrium $s \in S(M,\Omega)$ in $(G(f), w(\omega))$ satisfies

$$|g(s(\omega))(a) - f(\omega)(a)| \leq \varepsilon \text{ for all } a \in A.$$

Here, we require all Nash equilibria to achieve the lottery α approximately. When $\varepsilon = 0$, we will replace "virtually $\varepsilon-$implement" with "exactly implement".

Maskin (1977/1999) showed that if a deterministic social choice function is exactly implementable in Nash equilibrium, then it must be monotonic, and therefore, must be dictatorial. On the other hand, Matsushima (1988) and Abreu and Sen (1991) showed, by extending the analysis by Maskin (1977/1999) on social choice correspondences, that monotonicity is necessary and sufficient for exact implementability of stochastic social choice functions in Nash equilibrium. Hence, almost every stochastic social choice function is exactly implementable in Nash equilibrium. This implies that every social choice function, irrespective of whether it is stochastic or deterministic, is virtually

$\varepsilon-$implementable in Nash equilibrium, where we can choose ε as close to zero as possible. This possibility result is in contrast with the case of strategy-proofness in the general social choice environments.

In spite of these possibility results, the previous works have investigated only a very limited class of social choice functions in the following sense. For every $\varepsilon \in [0,1]$, a mechanism G is said to *virtually* $\varepsilon-$*implement a lottery* $\alpha \in \Delta(A)$ *in Nash equilibrium with respect to* $u \in U(M)$ if there exists a Nash equilibrium in (G,u), and any Nash equilibrium $m \in M$ in (G,u) satisfies

$$|g(m)(a)-\alpha(a)| \leq \varepsilon \quad \text{for all } a \in A.$$

When $\varepsilon = 0$, we will replace "virtually $\varepsilon-$ implement" with "exactly implement". Suppose that a social choice function $f \in F(\Omega)$ is exactly implementable in Nash equilibrium. Then, there exists a mechanism G that exactly implements $f(\omega)$ with respect to $w(\omega)$ in Nash equilibrium for all $\omega \in \Omega$. Fix $\omega \in \Omega$ and $\omega' \in \Omega$ arbitrarily, and suppose that $w_i(\omega)$ and $w_i(\omega')$ are preference-equivalent for all $i \in N$. Then, it is clear from the definition of Nash implementation that

$$f(\omega) = f(\omega').$$

This implies that any social choice function that is exactly implementable in Nash equilibrium must depend only on agents' preferences.

This point might be controversial, especially when we do not allow each agent's announcement to have intrinsic value, and therefore, agents' preferences describe only their material interests on consequences. However, as Rawls (1971), Dworkin (1981), Sen (1982, 1985), and others have pointed out in their respective works on theoretical foundation of social choice and welfare, reasonable social choice functions should depend on factors such as interpersonal comparison that include information other than agents' material interests. Hence, in order to implement such reasonable social choice functions, any relevant factors besides agents' material interests must be verifiable, and the mechanism must directly be contingent on these factors.[3]

Instead of investigating only material-interest-oriented agents, Matsushima (2003) focused on agents who have small moral preferences, where each agent suffers small disutility from announcing dishonestly. Matsushima then showed that there exists a single mechanism that always exactly implements any alternative in Nash equilibrium whenever agents regard this alternative as being socially optimal. Section 4 will show a brief explanation on this paper.

[3] For related arguments, see Maskin and Tirole (1999) and Tirole (1999).

3.3 Bayesian Framework

For every $\varepsilon \in [0,1]$, a mechanism G is said to *virtually* $\varepsilon -$ *implement a social choice function* $f \in F(\Omega)$ *in Bayesian Nash equilibrium (iterative dominance) with respect to* $(w,\psi) \in W(M,\Omega) \times \Psi$ if there exists a Bayesian Nash equilibrium (iteratively undominated strategy profile) in (G,w,ψ), and any Bayesian Nash equilibrium (iteratively undominated strategy profile, respectively) $s \in S(M,\Omega)$ in (G,w,ψ) satisfies

$$|g(s(\omega))(a) - f(\omega)(a)| \leq \varepsilon \text{ for all } \omega \in \Omega \text{ and all } a \in A.$$

When $\varepsilon = 0$, we will replace "virtually $\varepsilon -$ implement" with "exactly implement". Several works such as Jackson (1991), Matsushima (1993), Abreu and Matsushima (1992b), Duggan (1997), and Serrano and Vohra (2002) showed the possibility of social choice functions being implementable in Bayesian Nash equilibrium or iterative dominance. The mechanisms constructed in these works depend on the detail of the model specification such as the prior distribution, and therefore, may not be practically useful. Moreover, agents' equilibrium behaviors in these mechanisms depend crucially on strong common knowledge assumptions among agents on the prior distribution, their rationality, and others.

Based on this observation, several recent works such as Bergemann and Morris (2003a, 2003b) and Chung and Ely (2004) investigated ex post equilibrium instead of Bayesian Nash equilibrium or iteratively undominated strategy profile. A strategy profile $s \in S(M,\Omega)$ is said to be an *ex post equilibrium* in (G,w) if it is a *Bayesian Nash equilibrium* in (G,w,ψ) for all $\psi \in \Psi$. A mechanism G is said to exactly implement a social choice function $f \in F(\Omega)$ in ex post equilibrium with respect to $w \in W(M,\Omega)$ if there exists an ex post equilibrium in (G,w), and any ex post equilibrium $s \in S(M,\Omega)$ in (G,w) satisfies

$$g(s(\omega)) = f(\omega) \text{ for all } \omega \in \Omega.$$

Bergemann and Morris (2003b) investigated the possibility of a social choice function being exactly implementable in ex post equilibrium. However, implementation in ex post equilibrium does *not* require that for every common prior $\psi \in \Psi$, any Bayesian Nash equilibrium $s \in S(M,\Omega)$ in (G,w,ψ) induce the values of the social choice function $f \in F(\Omega)$, i.e., $g(s(\omega)) = f(\omega)$ for all $\omega \in \Omega$.

The recent work by Matsushima and Ohashi (2004) required the mechanism to have no unwanted Bayesian Nash equilibrium irrespective of how to specify the prior distribution. For every $\varepsilon \in [0,1]$, a mechanism G is said to *virtually* $\varepsilon -$ *implement a social choice function* $f \in F(\Omega)$ *in Bayesian Nash equilibrium with respect to* $w \in W(M,\Omega)$ if for every $\psi \in \Psi$, there exists a Bayesian Nash equilibrium in (G,w,ψ), and any Bayesian Nash equilibrium $s \in S(M,\Omega)$ in (G,w,ψ) satisfies

$$|g(s(\omega))(a) - f(\omega)(a)| \leq \varepsilon \text{ for all } \omega \in \Omega \text{ and all } a \in A.$$

Matsushima and Ohashi assumed that the size of the set of alternatives is sufficiently large relatively to the size of the set of states, and showed that for every $\varepsilon \in [0,1]$, any social choice function $f \in F(\Omega)$ is virtually $\varepsilon -$implementable in Bayesian Nash equilibrium with respect to $w \in W(M,\Omega)$ with small fines whenever truth-telling is an ex post equilibrium in $(G(f), w)$. Matsushima and Ohashi designed mechanisms that do not depend on the prior distribution.[4]

The latter part of Matsushima (2004) investigated very general double auction environments with interdependent values. Matsushima (2004) provided a new idea of detail-free auction design as an extension of auction design with private values studied in the former part of this paper. Matsushima then showed that the designed auction mechanism always virtually implements the rational expectations equilibrium allocations in iterative dominance. Here, the set of iteratively undominated strategy profiles is twice dominance solvable. Hence, in order to calculate this set, each agent needs to know only that the other agents never play dominated strategies. Moreover, the set of iteratively undominated strategy profiles is interchangeable. Hence, agents can reach a set of Bayesian Nash equilibria only by twice iteratively eliminating dominated strategies. Section 6 will show a brief explanation of the latter part of this paper.

[4] The result of Matsushima and Ohashi does not depend on the existence of common prior distribution, and could exclude any unwanted subjective correlated equilibrium. For related analysis, see Brandenburger and Dekel (1987), Battigalli and Siniscalchi (2003), and Bergemann and Morris (2003b).

4. Moral Preferences

This section shows a brief explanation on Matsushima (2003), which assumed that agents have small moral preferences, and then showed that there exists a detail-free mechanism that exactly implements any alternative in Nash equilibrium, whenever agents regard this alternative as being socially optimal.

Let $N = \{1,2,3\}$. Consider the follows mechanism $G^* = (M, g)$ such that

$$M_i = A^{2k} \text{ for all } i \in \{1,2,3\},$$

where k is an arbitrary positive integer that is sufficiently large. Hence, each agent $2k$ times announces her opinions on which alternative the central planner should choose. Let $m_i = (m_i^1,...,m_i^{2k})$. For every $l \in \{1,...,2k\}$, let $m^l = (m_i^l)_{i \in N}$ denote the $l-th$ message profile. The central planner picks up one integer l from the set $\{k+1,...,2k\}$ at random, i.e., with probability $\frac{1}{k}$, and then chooses a pure alternative $x(m^l) \in A$ where

$$x(m^l) = a \text{ if } m_i^l = a \text{ for two or three agents and } a \text{ is enforceable,}$$

and

$$x(m^l) = \tilde{a} \text{ if there exists no such } a \text{, where } \tilde{a} \text{ is regarded as “the status quo”.}$$

The central planner fines any agent a small positive monetary amount $\frac{\xi}{2} > 0$ if she is the last deviant from one of her first k messages. That is, the central planner fines agent $i \in N$ if and only if there exist $l \in \{1,...,k\}$ and $l' \in \{k+1,...,2k\}$ such that

$$m_i^l \neq m_i^{l'},$$

and

$$m^r = m^{r'} \text{ for all } r \in \{1,...,k\} \text{ and all } r' \in \{l'+1,...,2k\}.$$

The mechanism G^* specified above is detail-free in the following very strict sense. It never depends on the model specification such as agents' utility functions and the social choice function. In fact, it is simply described by the following statement that the central planner gives to agents.

> “Tell me $2k$ times on what I should do. I will pick up one opinion profile among the $2k$ announced profiles. If at least two of you recommend me to do the same thing and I can enforce it, then I will do it. Otherwise, I will do nothing.”

Here, all we need to assume is that whether any recommendation to be enforceable is verifiable. We do not even need to specify the set of enforceable alternatives in advance.

Fix any alternative $a^* \in A$ arbitrarily, which agents regard as being socially optimal. Matsushima (2003) confined attentions to utility functions for each agent $i \in N$ such that whenever the number of her opinions that do not recommend the socially optimal alternative a^* is more than or equals k, then she suffers psychological disutility that is at least equivalent to the amount ξ of monetary loss. We denote by $U_i(M_i, a^*, \xi)$ the set of such utility functions for agent i. Let $U(M, a^*, \xi) = \prod_{i \in N} U_i(M_i, a^*, \xi)$. Matsushima (2003) showed that for every $a^* \in A$ and every $u \in U(M)$, the detail-free mechanism G^* specified above exactly implements a^* in Nash equilibrium with respect to u whenever $u \in U(M, a^*, \xi)$.

By letting the number of opinions $2k$ sufficiently large, we can choose the monetary fine $\frac{\xi}{2}$ and the lower bound of disutility for dishonest reporting ξ as close to zero as possible. Hence, only by introducing moral preferences in the *minimal* way, we can exactly implement any alternative in Nash equilibrium by using the *single* detail-free mechanism, as long as agents regard it as being socially optimal. Matsushima (2003) could extend this result to the incomplete information environments.

5. Perfect Competition

This section shows a brief explanation of the former part of Matsushima (2004), which investigated double auction design in the general equilibrium model where many sellers and many buyers with quasi-linear preferences trade multiple commodities.

Let $N=\{1,...,4n\}$. There exist $2n$ sellers, i.e., sellers 1, …, and $2n$, and $2n$ buyers, i.e., buyers $2n+1$, …, and $4n$, where n is a sufficiently large positive integer. There exist k distinct commodities to be traded. Each seller can produce multiple units of each commodity up to l units. Each buyer will demand multiple unit of each demand up to l units. All agents have private values.

Matsushima (2004) introduced a new idea on how to design auction formats as follows. First, the auctioneer randomly divides all agents into two distinct groups, i.e., groups 1 and 2. The auctioneer randomly, i.e., with probability $\frac{1}{(2n!)^2}$, chooses a one-to-one function on the set of agents, i.e., $\phi : N \to N$, that is defined as a combination of a permutation on the set of sellers and a permutation on the set of buyers. As the members of group 1, the auctioneer picks up the n sellers $i \in \{1,...,2n\}$ whose values of the function ϕ belong to $\{1,...,n\}$, i.e., $\phi(i) \in \{1,...,n\}$. As the members of group 1, the auctioneer also picks up the n buyers $i \in \{2n+1,...,4n\}$ whose values of ϕ belong to $\{2n+1,...,3n\}$, i.e., $\phi(i) \in \{2n+1,...,3n\}$. The other agents belong to group 2.

The auctioneer asks each buyer to announce a demand function. At the same time, the auctioneer asks each seller to announce a supply function. The auctioneer calculates the (approximate) market-clearing price vector within each group. Importantly, agents in each group trade at the market-clearing price vector in *the other group*. The auctioneer determines the trading amounts for each agent according to the following rationing rule. In each group, and in each commodity market within this group, if there exists excess supply, then any buyer in this group can buy the same amount as what she intends to demand, but some sellers in this group whose values of ϕ are relatively high cannot sell the same amount as what she intends to supply. Similarly, if there exists excess demand, then any seller can sell the same amount as what she intends to supply, but some buyers whose values of ϕ are relatively high cannot buy the same amount as what she intends to demand.

Note that each agent's announcement *never* influences the price vector at which she trades. This, together with the specification of rationing rule, will be the driving force of incentivizing each seller (buyer) to announce her true competitive supplies (demands, respectively) as the dominant strategy. When the number of agents is sufficiently large, it is almost sure that the market-clearing price vector within each group approximates the market-clearing price vector in the whole markets combining both groups. This implies that the auctioneer can achieve the perfectly competitive allocation in the limit as the number of agents grows. Hence, the former part of Matsushima (2004) showed that for every $\varepsilon > 0$

and every sufficiently large n, the above auction format always virtually $\varepsilon-$implements the perfectly competitive allocation in dominance.

This result is in sharp contrast with the previous works on the mechanism design literature. Groves (1973) showed that there exists no non-trivial strategy-proof deterministic social choice function that is efficient and satisfies *budget-balancing* in that the sum of the monetary transfers always equals zero. In contrast, Matsushima's auction format satisfies budget-balancing. Barbarà and Jackson (1995) showed that any anonymous and deterministic social choice function that is strategy-proof is inefficient even in the limit as the number of agents grows. Matsushima format satisfies anonymity also, but allows stochastic decision. McAfee (1992) showed an idea of efficient double auction design where the budgetary deficit never occurs. McAfee's analysis, however, relies crucially on the assumption of single-unit demands and supplies. In contrast, Matsushima (2004) covers very wide class of double auction environments where we allow multiple commodities to be traded, each seller (buyer) to supply (demand) multiple units, and allow any mixture of complements and substitutes among distinct commodities for each traders.

There are researches on a strategic foundation of perfect competition by using naïve double auction models. See Wilson (1977), Rustichini, Satterthwaite, and Williams (1994), Fudenberg, Mobius, and Szeidl (2003), Jackson and Swinkels (2004), and others. These works investigated Bayesian Nash equilibria instead of dominant strategies, and therefore, inevitably depend on the strong common knowledge and rationality assumptions.

6. Rational Expectations Equilibrium

This section shows a brief explanation of the latter part of Matsushima (2004), where agents have *interdependent* values in the double auction environments. Let $N = \{1,...,2rn\}$, where r and n are sufficiently large positive integers. There exist rn sellers, i.e., sellers 1, …, and rn, and rn buyers, i.e., buyers $rn+1$, …, and $2rn$. There exist k distinct commodities to be traded. Each seller can produce multiple units of each commodity up to l units. Each buyer will demand multiple unit of each demand up to l units. Since agents have interdependent values, the auctioneer has to make each agent informed of the other agents' private signals in order to achieve *ex post* efficiency. The latter part of Matsushima (2004) could extend the idea of random grouping in the private value case to the interdependent value case in the following way.

First, the auctioneer randomly divides all agents into r distinct groups, i.e., groups 1, …, and r. The auctioneer randomly, i.e., with probability $\frac{1}{(rn!)^2}$, chooses a one-to-one function on the set of agents, i.e., $\phi : N \to N$, which is defined as a combination of a permutation on the set of sellers and a permutation on the set of buyers. As the members of each group $h \in \{1,...,r\}$, the auctioneer picks up the n sellers $i \in \{1,...,rn\}$ whose values of the function ϕ belong to $\{(h-1)n+1,...,hn\}$, i.e., $\phi(i) \in \{(h-1)n+1,...,hn\}$. As the members of each group $h \in \{1,...,r\}$, the auctioneer also picks up the n buyers $i \in \{rn+1,...,2rn\}$ whose values of ϕ belong to $\{(h-1)n+1,...,hn\}$, i.e., $\phi(i) \in \{(h-1)n+1,...,hn\}$. According to the following three stages, the auctioneer three times asks each buyer to announce a demand function, and three times asks each seller to announce a supply function.

At stage 1, each buyer (seller) announces a demand (supply, respectively) function as her first message. A the end of stage 1, agents in each group $h \in \{1,...,r\}$ observe the first messages announced in the *subsequent* group $h+1$, where we denote $r+1=1$.

At stage 2, each buyer (seller) announces a demand (supply, respectively) function as her second message. At the end of stage 2, agents in each group $h \in \{1,...,r\}$ observe the *first* messages announced in all groups except her group, i.e., except group h. By regarding the second messages as agents' supplies and demands, the auctioneer calculates the (approximate) market-clearing price vector p^h in each group $h \in \{1,...,r\}$.

At stage 3, each buyer (seller) announces a demand (supply, respectively) function as her third message. At the end of stage 3, agents in each group $h \in \{1,...,r\}$ almost certainly trade at the market-clearing price vector in the subsequent group, i.e., at p^{h+1}, which the auctioneer calculated at the end of step 2. Here, the auctioneer determines the trading amounts by regarding their third messages as agents' supplies and demands. With small but positive probability, the auctioneer randomly picks up an arbitrary price vector, and then agents trade at this price vector. Here, the auctioneer randomly chooses one of the three message profiles, and determines the trading amounts by regarding this profile as agents' supplies and demands. Whenever there exists either excess supply or excess demand in each commodity market within each group, then the auctioneer will use a modified version of the rationing rule specified in the private value case.

The above auction format is detail-free in that the auctioneer needs no information about the model specification such as agents' utility functions and the prior distribution. Note that each agent's three messages never influence the price vector at which she trades. This, together with the well-specified rationing rule, implies that without any knowledge on the other agents' private signals, each seller (buyer) has incentive to truthfully announce her competitive supplies (demands, respectively) as her undominated first message. With a minor informational condition, agents' first message announcement can fully reveal their private signals. Similarly, with full knowledge on the private signals of the members in the subsequent group, each seller (buyer) has incentive to truthfully announce her competitive supplies (demands, respectively) as her twice iteratively undominated second message.

We shall assume that by observing only $(r-1)n$ sellers' and $(r-1)n$ buyers' private signals in addition to her private signal, each seller (buyer) can calculate her true competitive supplies (demand, respectively) based on full knowledge about the state. This assumption might be unrestrictive whenever r is sufficiently large. Hence, the auctioneer can incentivize each seller (buyer) to truthfully announce her competitive supplies (demands, respectively) based on full knowledge about the state as her twice iteratively undominated third message.

When n is sufficiently large, it is almost certain from the law of large numbers that each seller's (buyer's) second message approximates her true competitive supplies (demands, respectively) based on full knowledge about the state. Hence, it is almost certain that the market-clearing price vector in each group approximates the market-clearing price vector in the whole markets combining all groups.

Based on the above observations, the latter part of Matsushima (2004) could show that for every $\varepsilon > 0$, every sufficiently large n, and every sufficiently large r, the above auction format always virtually $\varepsilon-$implements the ex post efficient allocation in iterative dominance. Here, the set of iteratively undominated strategy profiles is twice dominance solvable and interchangeable. Any iteratively undominated strategy profile mimics price-taking behavior.

This possibility result is closely related to the study of rational expectations equilibrium in competitive economies originated in the seminal work by Lucas (1972). The recent works by Reny and Perry (2003) investigated a strategic foundation of rational expectations equilibrium by using a naïve single-object and single-unit double auction, and showed that when agents' private signals are strictly affiliated and the number of agents is sufficiently large, there exists an approximately ex post efficient pure strategy Bayesian Nash equilibrium that mimics price-taking behavior.

The study of rational expectations equilibrium presumes that all agents' rational behavior is common knowledge among agents. In contrast, the latter part of Matsushima (2004) does not need to assume such restrictive assumptions. All we need is to assume that each agent knows that the other agents never play dominated strategies. Matsushima (2004) covers a very wide class of trading environments even with interdependent values. We do not require the private signals to be affiliated, we can allow multiple objects to be traded, and we can allow any mixture of complements and substitutes for every agent.

References

Abreu, D. and H. Matsushima (1992a): "Virtual Implementation in Iteratively Undominated Strategies: Complete Information," *Econometrica* 60, 993-1008. Also in *Recent Development in Game Theory, The International Library of Critical Writings in Economics 109*, edited by E. Maskin, Edward Elgar.

Abreu, D. and H. Matsushima (1992b): "Virtual Implementation in Iteratively Undominated Strategies: Incomplete Information," mimeo.

Abreu, D. and H. Matsushima (1994): "Exact Implementation," *Journal of Economic Theory* 64, 1-19.

Abreu, D. and A. Sen (1991): "Virtual Implementation in Nash equilibrium," *Econometrica* 59, 997-1021.

Barberà, S. and M. Jackson (1995): "Strategy-Proof Exchange," *Econometrica* 63, 51-87.

Battigalli, P. and M. Siniscalchi (2003): "Rationalization and Incomplete Information," *Advances in Theoretical Economics* 3 (1), Article 3.

Benoit, J.-P. (2002): "Strategic Manipulation in Voting Games When Lotteries and Ties are Permitted," *Journal of Economic Theory* 102, 421-436.

Bergemann, D. and S. Morris (2003a): "Robust Mechanism Design," mimeo, Yale University.

Bergemann, D. and S. Morris (2003b): "Robust Implementation: The Role of Large Type Spaces," mimeo, Yale University.

Brandenburger, A. and E. Dekel (1987): "Hierarchies of Beliefs and Common Knowledge," *Journal of Economic Theory* 59, 189-198.

Chung, K.-S. and J. Ely (2004): "Foundation of Dominant Strategy Mechanisms," mimeo.

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

Duggan, J. (1997): "Virtual Bayesian Implementation," *Econometrica* 65, 1175-1199.

Dworkin, R. (1981): "What is equality ? Part 1: Equality of Welfare, Part 2: Equality of Resources," *Philosophy and Public Affairs* 10, 185-246, 283-345.

Fudenberg, D, M. Mobius, and A. Szeidl (2003): "Existence of Equilibrium in Large Double Auctions," mimeo.

Gibbard, A. (1977): "Manipulation of Schemes that Mix Voting with Chance," *Econometrica* 45, 665-681.

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

Hurwicz, L. (1972): "On Informational Decentralized Systems," in *Decision and Organization*, edited by R. Radner and C. McGuire, North Holland.

Jackson, M. (1991): "Bayesian Implementation," *Econometrica* 59, 461-477.

Jackson, M. and J. Swinkels (2004): "Existence of Equilibrium in Single and Double Private Value Auctions," forthcoming in *Econometrica*.

Lucas, R. (1972): "Expectation and Neutrality of Money," *Journal of Economic Theory* 4, 103-124.

Mas-Colell, A., M. Whinston, and J. Green (1995): *Microeconomic Theory*, Oxford

University Press.
Maskin, E. (1999/1977): "Nash equilibrium and Welfare Optimality," *Review of Economic Studies* 66, 23-38.
Maskin, E. and T. Sjöström (2002): "Implementation Theory," in *Handbook of Social Choice and Welfare Volume 1*, ed. by K. Arrow, A. Sen, and K. Suzumura. Elsevier.
Maskin, E. and J. Tirole (1999): "Unforeseen Contingencies and Incomplete Contracts," *Review of Economic Studies* 66, 83-114.
Matsushima, H. (1988): "A New Approach of the Implementation Problem," *Journal of Economic Theory* 45, 128-144.
Matsushima, H. (1990): "Dominant Strategy Mechanisms with Mutually Payoff-Relevant Private Information and with Public Information," *Economics Letters* 34, 109-112.
Matsushima, H. (1991a): "Incentive Compatible Mechanisms with Full Transferability," *Journal of Economic Theory* 50, 198-203.
Matsushima, H. (1991b): "Coalitionally Dominant Strategy Mechanisms with Limited Public Information," *Economics Letters* 37, 371-375.
Matsushima, H. (1993): "Bayesian Monotonicity with Side Payments," *Journal of Economic Theory* 39, 107-121.
Matsushima, H. (2003): "Universal Mechanisms and Moral Preferences in Implementation," Discussion Paper CIRJE-F-254, Faculty of Economics, University of Tokyo.
Matsushima, H. (2004): "Large Double Auction Design in Dominance," Discussion Paper CIRJE-F-282, Faculty of Economics, University of Tokyo.
Matsushima, H. and Y. Ohashi (2004): "Belief-Free Implementation," in preparation.
McAfee, P. (1992): "A Dominant Strategy Double Auction," *Journal of Economic Theory* 56, 434-450.
Moore, J. (1992): "Implementation in Environments with Complete Information," in *Advances in Economic Theory: Sixth World Congress*, edited. by J.-J. Laffont, Cambridge University Press.
Osborne, M. and A. Rubinstein (1994): *A Course in Game Theory*, MIT Press.
Palfrey, T. (1992): "Implementation in Bayesian Equilibrium: the Multiple Equilibrium Problem in Mechanism Design," in *Advances in Economic Theory: Sixth World Congress*, ed. by J.-J. Laffont, Cambridge University Press.
Rawls, J. (1971): *A Theory of Justice*, Cambridge: Harvard: Harvard University Press.
Rustichini, A., M. Satterthwaite, and S. Williams (1994): "Convergence to Efficiency in a Simple Market with Incomplete Information," *Econometrica* 62, 1041-1063.
Reny, P. and M. Perry (2003): "Toward a Strategic Foundation for Rational Expectation Equilibrium," working paper, University of Chicago.
Sen, A. (1982): *Choice, Welfare and Measurement*, Oxford: Blackwell.
Sen, A. (1985): *Commodities and Capabilities*, Amsterdam: North-Holland.
Serrano, R. and R. Vohra (2002): "A Characterization of Virtual Bayesian Implementation," mimeo.
Tirole, J. (1999): "Incomplete Contracts: Where do we stand?," *Econometrica* 67, 741-781.
Wilson, R. (1977): "A Bidding Model of Perfect Competition," *Review of Economic Studies* 44, 511-518.

Wilson, R. (1985): "Incentive Efficiency in Double Auctions," *Econometrica* 53, 1101-1115.

Wilson, R. (1987): "Game-Theoretic Approaches to Trading Processes," in *Advances in Economic Theory: Fifth World Congress*, edited. by T. Bewley, Cambridge University Press.