Bugs in the proofs of revelation principle

Wu, Haoyang

Wan-Dou-Miao Research Lab, Suite 1002, 790 WuYi Road,
Shanghai, China.

10 August 2010

Online at https://mpra.ub.uni-muenchen.de/31285/
MPRA Paper No. 31285, posted 05 Jun 2011 15:18 UTC
Bugs in the proofs of revelation principle

Haoyang Wu

Wan-Dou-Miao Research Lab, Suite 1002, 790 WuYi Road, Shanghai, 200051, China.

Abstract

In the field of mechanism design, the revelation principle has been known for decades. Myerson, Mas-Colell, Whinston and Green gave formal proofs of the revelation principle respectively. However, in this paper, I argue that there are bugs hidden in their proofs.

JEL codes: D7

Key words: Revelation principle; Mechanism design; Implementation theory.

The revelation principle is well-known in the economics literature. See Page 884, Line 24 [1]: “The implication of the revelation principle is ... to identify the set of implementable social choice functions, we need only identify those that are truthfully implementable.” But, in this paper I will argue that there are bugs in the proofs given by Mas-Colell, Whinston and Green [1] and Myerson [2] respectively. Coincidentally, the bugs are relevant to the same word “imply”. Related definitions and proofs are given in Appendices, which are cited from Section 8.E, 23.B and 23.D [1] and Ref. [2]. Two remarks are added in Appendix 1 and 3 respectively.

1 The bug in the proof by Mas-Colell, Whinston and Green

Here, the notation is referred to Ref. [1]. See the proof of Proposition 23.D.1: “... Condition (23.D.2) implies that for all i and all \( \theta_i \in \Theta_i \).” To derive formula (23.D.3), the term “\( \hat{s}_i \) (\( \forall \hat{s}_i \in S_i, i = 1, \cdots, I \))” in formula (23.D.2) is replaced by “\( s^*_i(\theta_i) \)” (\( \forall \theta_i \in \Theta_i, i = 1, \cdots, I \)). Since formula (23.D.2) holds for all 

* Corresponding author.

Email address: hywch@mail.xjtu.edu.cn, Tel: 86-18621753457 (Haoyang Wu).
\( s_i \in S_i \), it looks straightforward to do so at first sight. However, in what follows I will argue that this “straightforward” implication does not hold.

First, note that both formula (23.D.2) and (23.D.3) correspond to the same indirect mechanism \( \Gamma = (S_1, \cdots, S_I, g(\cdot)) \) given in Proposition 23.D.1. The reason is that these two formulas are based on the same outcome function \( g(\cdot) \), and after formula (23.D.3) the condition \( g(s^*(\theta)) = f(\theta) \) (for all \( \theta \in \Theta \)) also comes from the indirect mechanism \( \Gamma \).

Next, according to the definition of a pure strategy for player \( i \) in a Bayesian game, the legal input of the function \( s^*_i(\cdot) \) should be some realized type of agent \( i \) (see Remark 1).

Now consider the right part of formula (23.D.3), \( \mathbb{E}_{\theta_{-i}}[u_i(g(s^*_i(\hat{\theta}_i), s^*_{-i}(\theta_{-i})), \theta_i) | \theta_i] \). According to Proposition 8.E.1, the expectation is taken over realizations of the other players’ random types conditional on player \( i \)’s realized type \( \theta_i \). Given that agent \( i \)’s type has been realized as \( \theta_i \), none of \( \hat{\theta}_i (\forall \hat{\theta}_i \in \Theta_i, \hat{\theta}_i \neq \theta_i) \) can be such realized type. Therefore, in formula (23.D.3), the term “\( s^*_i(\hat{\theta}_i) \)” (\( \forall \hat{\theta}_i \in \Theta_i, \hat{\theta}_i \neq \theta_i \)) is actually illegal. Put differently, formula (23.D.3) is illegal. Hence, the aforementioned “straightforward” implication does not hold. That is the bug.

One may think the variable \( \hat{\theta}_i \) in formula (23.D.3) can be agent \( i \)’s announced type in a direct mechanism, and so can be different from the realized type \( \theta_i \). But as shown before, the mechanism corresponding to formula (23.D.3) is the indirect mechanism \( \Gamma \) given in Proposition 23.D.1, not a direct mechanism. It is illegal to let agent \( i \) directly announce a type in the Bayesian game induced by such indirect mechanism \( \Gamma \).

2 The bug in the proof by Myerson

Here, the notation is referred to Ref. [2]. See the proof of Theorem 2: “... Furthermore, the equilibrium inequalities (14) for \( \pi \) imply the incentive compatible inequalities (6) for \( \pi’ \)...”. Let us consider the right part of the incentive compatible inequalities (6) for \( \pi’ \). For all \( i, a_i \in A_i, b_i \in A_i \),

\[
Z_i(\pi', b_i|a_i) = \sum_{a \in A_1 \times \cdots \times A_n} \sum_{c \in C} P_i(\alpha|a_i)\pi'(c|\alpha_{-i}, b_i)U_i(c, \alpha)
= \sum_{a \in A_1 \times \cdots \times A_n} \sum_{s \in S_1 \times \cdots \times S_n} \sum_{c \in C} P_i(\alpha|a_i) \cdot \pi(c|s) \cdot [\prod_{j=1, j \neq i}^{n} \sigma_j(s_j|a_j) \times \sigma_i(s_i|b_i)] \cdot U_i(c, \alpha)
\]
As specified in the left term “$Z_i(\pi', b_i|a_i)$”, agent $i$’s type is realized as $a_i$. Therefore, according to Remark 2, the term “$\sigma_i(s_i|b_i)$” (for all $b_i \in A_i$, $b_i \neq a_i$) is actually illegal. Put differently, the incentive compatible inequalities (6) for $\pi'$ is illegal. That is the bug.

Appendix 1: Definitions and proof in Section 8.E [1]

According to page 255 [1], formally, in a Bayesian game, each player $i$ has a payoff function $u_i(s_i, s_{-i}, \theta_i)$, where $\theta_i \in \Theta_i$ is a random variable chosen by nature that is observed only by player $i$. The joint probability distribution of the $\theta_i$’s is given by $F(\theta_1, \cdots, \theta_I)$, which is assumed to be common knowledge among the players. Letting $\Theta = \Theta_1 \times \cdots \times \Theta_I$, a Bayesian game is summarized by $[I, \{S_i\}, \{u_i(.)\}, \Theta, F(.)]$. A pure strategy for player $i$ in a Bayesian game is a function $s_i(\theta_i)$, or decision rule, that gives the player’s strategy choice for each realization of his type $\theta_i$. Player $i$’s pure strategy set $S_i$ is therefore the set of all such functions. Player $i$’s expected payoff given a profile of pure strategies for the $I$ players $(s_1(\cdot), \cdots, s_I(\cdot))$ is then given by:

$$\tilde{u}_i(s_1(\cdot), \cdots, s_I(\cdot)) = E_{\theta}[u_i(s_1(\theta_1), \cdots, s_I(\theta_I), \theta_i)], \quad (8.E.1)$$

**********************************************************************************

**Remark 1:** Following page 148 [3], the timing of a static Bayesian game is as follows:

Step 1: Nature chooses a type vector $\theta = (\bar{\theta}_1, \cdots, \bar{\theta}_I)$, where $\bar{\theta}_i$ is the realized type of agent $i$;

Step 2: Nature reveals $\bar{\theta}_i$ to player $i$ but not to any other player;

Step 3: The players simultaneously output $(s_1(\bar{\theta}_1), \cdots, s_I(\bar{\theta}_I))$;

Step 4: Each player $i$ receives the payoff $u_i(s_1(\theta_1), \cdots, s_I(\theta_I), \theta_i)$.

For each player $i = 1, \cdots, I$, consider his strategy function $s_i(\cdot)$, then:

1) $s_i(\cdot)$ is chosen (or controlled) by player $i$, and is his private information;

2) In a static Bayesian game, player $i$’s type can be realized as any element of $\Theta_i$. The realized type of player $i$ is his private information;

3) The legal input parameter of $s_i(\cdot)$ must be some realized type $\bar{\theta}_i$ in $\Theta_i$, and the output of $s_i(\cdot)$ is $s_i(\bar{\theta}_i)$ which is observable to the outside agent (either principal or mediator).

4) Suppose player $i$’s type has been realized as $\bar{\theta}_i$ in Step 1, then in Step 3, it is illegal to let player $i$ output $s_i(\bar{\theta}_i)$ for any $\theta_i \in \Theta_i$, $\theta_i \neq \bar{\theta}_i$.

**********************************************************************************

**Definition 8.E.1:** A (pure strategy) Bayesian Nash equilibrium for the Bayesian...
game \([I, \{S_i\}, \{u_i(\cdot)\}, \Theta, F(\cdot)]\) is a profile of decision rules \((s_1(\cdot), \cdots, s_I(\cdot))\) that constitutes a Nash equilibrium of game \(\Gamma_N = [I, \{S_i\}, \{\tilde{u}_i(\cdot)\}]\). That is, for every \(i = 1, \cdots, I\),

\[\tilde{u}_i(s_i(\cdot), s_{-i}(\cdot)) \geq \tilde{u}_i(s'_i(\cdot), s_{-i}(\cdot))\]

for all \(s'_i(\cdot) \in \mathcal{S}_i\), where \(\tilde{u}_i(s_i(\cdot), s_{-i}(\cdot))\) is defined as in Eq\(8.\)E.1).

A very useful point to note is that in a (pure strategy) Bayesian Nash equilibrium each player must be playing a best response to the conditional distribution of his opponents’ strategies for each type that he might end up having. Proposition 8.E.1 provides a more formal statement of this point.

**Proposition 8.E.1**: A profile of decision rules \((s_1(\cdot), \cdots, s_I(\cdot))\) is a Bayesian Nash equilibrium in Bayesian game \([I, \{S_i\}, \{u_i(\cdot)\}, \Theta, F(\cdot)]\) if and only if, for all \(i\) and all \(\theta_i \in \Theta_i\) occurring with positive probability,

\[E_{\theta_{-i}}[u_i(s_i(\bar{\theta}_i), s_{-i}(\bar{\theta}_{-i}, \theta_i)|\theta_i)] \geq E_{\theta_{-i}}[u_i(s'_i, s_{-i}(\theta_{-i}, \bar{\theta}_i)|\theta_i)], \tag{8.E.2}\]

for all \(s'_i \in S_i\), where the expectation is taken over realizations of the other players’ random variables conditional on player \(i\)’s realization of his signal \(\bar{\theta}_i\).

**Proof**: For necessity, note that if Eq(8.E.2) did not hold for some player \(i\) for some \(\bar{\theta}_i \in \Theta_i\) that occurs with positive probability, then player \(i\) could do better by changing his strategy choice in the event he gets realization \(\bar{\theta}_i\), contradicting \((s_1(\cdot), \cdots, s_I(\cdot))\) being a Bayesian Nash equilibrium. In the other direction, if condition Eq(8.E.2) holds for all \(\bar{\theta}_i \in \Theta_i\) occurring with positive probability, then player \(i\) cannot improve on the payoff he receives by playing strategy \(s_i(\cdot)\). \(\square\)

**Appendix 2: Definitions and proof in Section 23.B and 23.D [1]**

(P858) Consider a setting with \(I\) agents, indexed by \(i = 1, \cdots, I\). These agents make a collective choice from some set \(X\) of possible alternatives. Prior to the choice, each agent \(i\) privately observes his type \(\theta_i\) that determines his preferences. The set of possible types for agent \(i\) is denoted as \(\Theta_i\). The vector of agents’ types \(\theta = (\theta_1, \cdots, \theta_I)\) is drawn from set \(\Theta = \Theta_1 \times \cdots \times \Theta_I\) according to probability density \(\phi(\cdot)\). Each agent \(i\)’s Bernoulli utility function when he is of type \(\theta_i\) is \(u_i(x, \theta_i)\).

**Definition 23.B.1**: A social choice function is a function \(f : \Theta_1 \times \cdots \times \Theta_I \to X\) that, for each possible profile of the agents’ types \((\theta_1, \cdots, \theta_I)\), assigns a collective choice \(f(\theta_1, \cdots, \theta_I) \in X\).
**Definition 23.B.3:** A mechanism $\Gamma = (S_1, \cdots, S_I, g(\cdot))$ is a collection of $I$ strategy sets $S_1, \cdots, S_I$ and an outcome function $g : S_1 \times \cdots \times S_I \to X$.

**Definition 23.B.4:** The mechanism $\Gamma = (S_1, \cdots, S_I, g(\cdot))$ implements social choice function $f(\cdot)$ if there is an equilibrium strategy profile $(s^*_1(\cdot), \cdots, s^*_I(\cdot))$ of the game induced by $\Gamma$ such that $g(s^*_1(\theta_1), \cdots, s^*_I(\theta_I)) = f(\theta_1, \cdots, \theta_I)$ for all $(\theta_1, \cdots, \theta_I) \in \Theta_1, \cdots, \Theta_I$.

**Definition 23.B.5:** A direct revelation mechanism is a mechanism in which $S_i = \Theta_i$ for all $i$ and $g(\theta) = f(\theta)$ for all $\theta \in \Theta_1 \times \cdots \times \Theta_I$.

**Definition 23.B.6:** The social choice function $f(\cdot)$ is truthfully implementable (or incentive compatible) if the direct revelation mechanism $\Gamma = (S_1, \cdots, S_I, f(\cdot))$ has an equilibrium $(s^*_1(\cdot), \cdots, s^*_I(\cdot))$ in which $s^*_i(\theta_i) = \theta_i$ for all $\theta_i \in \Theta_i$ and all $i = 1, \cdots, I$; that is, if truth telling by each agent $i$ constitutes an equilibrium of $\Gamma = (S_1, \cdots, S_I, f(\cdot))$.

**Definition 23.D.1:** The strategy profile $s^*(\cdot) = (s^*_1(\cdot), \cdots, s^*_I(\cdot))$ is a Bayesian Nash equilibrium of mechanism $\Gamma = (S_1, \cdots, S_I, g(\cdot))$ if, for all $i$ and all $\theta_i \in \Theta_i$,

$$E_{\theta_{-i}}[u_i(g(s^*_1(\theta_1), s^*_i(\theta_{-i})), \theta_i)|\theta_i] \geq E_{\theta_{-i}}[u_i(g(s_{i}^*(\theta_{-i})), \theta_i)|\theta_i]$$

for all $\hat{s}_i \in S_i$.

**Definition 23.D.2:** The mechanism $\Gamma = (S_1, \cdots, S_I, g(\cdot))$ implements the social choice function $f(\cdot)$ in Bayesian Nash equilibrium if there is a Bayesian Nash equilibrium of $\Gamma$, $s^*(\cdot) = (s^*_1(\cdot), \cdots, s^*_I(\cdot))$, such that $g(s^*(\theta)) = f(\theta)$ for all $\theta \in \Theta$.

**Definition 23.D.3:** The social choice function $f(\cdot)$ is truthfully implementable in Bayesian Nash equilibrium if $s^*_i(\theta_i) = \theta_i$ for all $\theta_i \in \Theta_i$ and $i = 1, \cdots, I$ is a Bayesian Nash equilibrium of the direct revelation mechanism $\Gamma = (\Theta_1, \cdots, \Theta_I, f(\cdot))$. That is, if for all $i = 1, \cdots, I$ and all $\theta_i \in \Theta_i$,

$$E_{\theta_{-i}}[u_i(f(\theta_i, \theta_{-i})), \theta_i)] \geq E_{\theta_{-i}}[u_i(f(\hat{\theta}_i, \theta_{-i})), \theta_i],$$

(23.D.1)

for all $\hat{\theta}_i \in \Theta_i$.

**Proposition 23.D.1 (The Revelation Principle for Bayesian Nash Equilibrium)** Suppose that there exists a mechanism $\Gamma = (S_1, \cdots, S_I, g(\cdot))$ that implements the social choice function $f(\cdot)$ in Bayesian Nash equilibrium. Then $f(\cdot)$ is truthfully implementable in Bayesian Nash equilibrium.

**Proof:** Since $\Gamma = (S_1, \cdots, S_I, g(\cdot))$ implements $f(\cdot)$ in Bayesian Nash equilibrium, then there exists a profile of strategies $s^*(\cdot) = (s^*_1(\cdot), \cdots, s^*_I(\cdot))$ such
that \( g(s^*(\theta)) = f(\theta) \) for all \( \theta \), and for all \( i \) and all \( \theta_i \in \Theta_i \),

\[
E_{\theta_{-i}}[u_i(g(s_i^*(\theta_i), s_{-i}^*(\theta_{-i})), \theta_i)|\theta_i] \geq E_{\theta_{-i}}[u_i(g(\hat{s}_i, s_{-i}^*(\theta_{-i})), \theta_i)|\theta_i], \tag{23.D.2}
\]

for all \( s_i \in S_i \). Condition (23.D.2) implies that for all \( i \) and all \( \theta_i \in \Theta_i \),

\[
E_{\theta_{-i}}[u_i(g(s_i^*(\theta_i), s_{-i}^*(\theta_{-i})), \theta_i)|\theta_i] \geq E_{\theta_{-i}}[u_i(g(s_i^*(\hat{\theta}_i), s_{-i}^*(\theta_{-i})), \theta_i)|\theta_i], \tag{23.D.3}
\]

for all \( \hat{\theta}_i \in \Theta_i \). Since \( g(s^*(\theta)) = f(\theta) \) for all \( \theta \), (23.D.3) means that, for all \( i \) and all \( \theta_i \in \Theta_i \),

\[
E_{\theta_{-i}}[u_i(f(\theta_i, \theta_{-i}), \theta_i)|\theta_i] \geq E_{\theta_{-i}}[u_i(f(\hat{\theta}_i, \theta_{-i}), \theta_i)|\theta_i], \tag{23.D.4}
\]

for all \( \hat{\theta}_i \in \Theta_i \). But, this is precisely condition (23.D.1), the condition for \( f(\cdot) \) to be truthfully implementable in Bayesian Nash equilibrium. \( \text{Q.E.D.} \)

**Appendix 3: Definitions and proof in Ref. [2]**

The arbitrator’s problem is described by a *Bayesian collective choice problem*, an object of the form:

\[
(C, A_1, A_2, \ldots, A_n, U_1, U_2, \ldots, U_n, P), \quad (1)
\]

The individual members of the group, or *players*, are numbered 1, 2, \( \cdots \), \( n \). \( C \) is the set of choices available to the group. For each player \( i \), \( A_i \) is the set of possible *types* for player \( i \). Each \( U_i : C \times A_1 \times \cdots \times A_n \rightarrow \mathbb{R} \) is a utility function such that each \( U_i(c, a_1, \cdots, a_n) \) is the payoff which player \( i \) would get if \( c \in C \) were chosen and if \( (a_1, \cdots, a_n) \) were the true vector of player types. \( P \) is a probability distribution on \( A_1 \times \cdots \times A_n \) such that \( P(a_1, \cdots, a_n) \) is the probability, as judged by the arbitrator, that \( (a_1, \cdots, a_n) \) is the true vector of types for the \( n \) players.

For some collection of *response sets* \( S_1, \cdots, S_n \), a *choice mechanism* is defined as a real-valued function \( \pi \) with a domain of the form \( C \times (S_1 \times \cdots \times S_n) \) such that:

\[
\sum_{c' \in C} \pi(c'|s_1, \cdots, s_n) = 1, \text{ and } \pi(c|s_1, \cdots, s_n) \geq 0 \text{ for all } c, \quad (2)
\]

for every \( (s_1, \cdots, s_n) \in S_1 \times \cdots \times S_n \).

Given a choice mechanism \( \pi \), for any player \( i \) and for any \( a_i \in A_i \) and \( b_i \in A_i \), let:

\[
Z_i(\pi, b_i|a_i) = \sum_{a \in A_1 \times \cdots \times A_n} \sum_{c \in C} P_i(\alpha|a_i)\pi(c|\alpha_{-i}, b_i)U_i(c, \alpha), \quad (5)
\]
where \((\alpha_i, b_i) = (\alpha_1, \cdots, \alpha_{i-1}, b_i, \alpha_{i+1}, \cdots, \alpha_n)\), \(P(\alpha|a_i) = 0\) if \(\alpha_i \neq a_i\). \(Z_i(\pi, b_i|a_i)\) is the conditionally-expected utility payoff for player \(i\), given that his type is \(a_i\), if he says that his type is \(b_i\) when \(\pi\) is the choice mechanism and when all other players are expected to tell the truth.

A choice mechanism \(\pi\) using the standard response sets is said to be **Bayesian incentive compatible** if

\[
Z_i(\pi, a_i|a_i) \geq Z_i(\pi, b_i|a_i), \quad \text{for all } i, a_i \in A_i, b_i \in A_i, \tag{6}
\]

If choice mechanism \(\pi\) is used and if everyone is honest, then player \(i\)'s conditionally-expected payoff when he knows \(a_i\) is:

\[
V_i(\pi|a_i) = Z_i(\pi, a_i|a_i), \tag{7}
\]

The allocation of conditionally-expected payoffs associated with mechanism \(\pi\) is the vector:

\[
V(\pi) = ((V_i(\pi|a_i))_{a_i \in A_i})_{i=1}^n. \tag{8}
\]

This is a vector of \(\sum_{i=1}^n |A_i|\) real numbers, indexed on the disjoint union of the \(A_i\) sets. If the arbitrator could use any choice mechanism and expect honest responses, then we would define the **feasible set** of expected allocation vectors to be:

\[F = \{V(\pi) : \pi \text{ is a choice mechanism}\}.\]

The set of **incentive-feasible** expected allocation vectors is defined to be:

\[F^* = \{V(\pi) : \pi \text{ is Bayesian incentive compatible}\}.\]

A **response plan** for player \(i\) is a function \(\sigma_i\) mapping each type \(a_i \in A_i\) onto a probability distribution over his response set \(S_i\). That is, if \(\sigma_i\) is player \(i\)'s response plan, then \(\sigma_i(s_i|a_i)\) is the probability that player \(i\) will tell the arbitrator \(s_i\) if his true type is \(a_i\).

********************************************************************************

**Remark 2**: Like Remark 1, I list the timing of a static Bayesian game as follows:

Step 1: Nature chooses a type vector \((\bar{a}_1, \cdots, \bar{a}_n)\), where \(\bar{a}_i\) is the **realized** type of agent \(i\);

Step 2: Nature reveals \(\bar{a}_i\) to player \(i\) but not to any other player;

Step 3: Player \(i\) tells his response \(s_i\) to the arbitrator according to the probability \(\sigma_i(s_i|\bar{a}_i)\). All players tell the arbitrator simultaneously.

Step 4: The arbitrator assigns choice \(c\) to all players according to the probability \(\pi(c|s_1, \cdots, s_n)\).

Step 5: Each player \(i\) receives the payoff \(U_i(c, \bar{a}_1, \cdots, \bar{a}_n)\).

For each player \(i = 1, \cdots, n\), consider his response plan \(\sigma_i(s_i|\cdot)\), then:

1) \(\sigma_i(s_i|\cdot)\) is chosen (or controlled) by player \(i\), and is his private information;
2) In a static Bayesian game, player $i$’s type can be realized as any element of $A_i$. The realized type of player $i$ is his private information;
3) The legal input parameter of $\sigma_i(s_i|\cdot)$ must be some realized type $\bar{a}_i$ in $A_i$, and the output of $\sigma_i(s_i|\cdot)$ is the probability that player $i$ will tell the arbitrator $s_i$ if his realized type is $\bar{a}_i$.
4) Suppose player $i$’s type has been realized as $\bar{a}_i$ in Step 1, then in Step 3, it is illegal to let player $i$ act using another response plan $\sigma_i(s_i|b_i)$ for any $b_i \in A_i$, $b_i \neq \bar{a}_i$.

If $(\sigma_1, \ldots, \sigma_n)$ lists the players’ response plans for the choice mechanism $\pi$, and if player $i$ knows that $a_i$ is his true type, then player $i$’s expected utility payoff is:

$$W_i(\pi, \sigma_1, \ldots, \sigma_n|a_i) = \sum_{\alpha \in A_i} \sum_{s \in S_i} P_i(\alpha|a_i) \sum_{\pi, \sigma_i, \bar{a}_i} W_i(\pi, \sigma_1, \ldots, \sigma_n|a_i).$$

The vector of conditionally-expected payoffs generated by $(\sigma_1, \ldots, \sigma_n)$ is:

$$\mathbf{W}(\pi, \sigma_1, \ldots, \sigma_n) = ((W_i(\pi, \sigma_1, \ldots, \sigma_n|a_i))_{a_i \in A_i}).$$

This is a vector with $\sum_{i=1}^n |A_i|$ components, indexed on the disjoint union of the $A_i$ sets, like the $\mathbf{V}(\pi)$. We say that $(\sigma_1, \ldots, \sigma_n)$ is a response-plan equilibrium for the choice mechanism $\pi$ if, for any player $i$ and type $a_i \in A_i$, for every possible alternative response plan $\sigma'_i$ for player $i$:

$$W_i(\pi, \sigma_1, \ldots, \sigma_n|a_i) \geq W_i(\pi, \sigma_1, \ldots, \sigma_{i-1}, \sigma'_i, \sigma_{i+1}, \ldots, \sigma_n|a_i).$$

The set of equilibrium-feasible expected allocation vectors is defined to be:

$$F^* = \{ \mathbf{W}(\pi, \sigma_1, \ldots, \sigma_n) : \pi \text{ is a choice mechanism, and } (\sigma_1, \ldots, \sigma_n) \text{ is a response-plan equilibrium for } \pi \}.$$  

**Theorem 2:** $F^* = F^*$.

**Proof:** If $(\sigma_1, \ldots, \sigma_n)$ is a response-plan equilibrium for a mechanism $\pi$ on $S_1, \ldots, S_n$, then we can define an equivalent choice mechanism $\pi'$ on $A_1, \ldots, A_n$ by:

$$\pi'(c|\alpha) = \sum_{s \in S_i} \pi(c|s) \cdot (\prod_{i=1}^n \sigma_i(s_i|\alpha_i)).$$

It is easy to check that $\mathbf{V}(\pi') = \mathbf{W}(\pi, \sigma_1, \ldots, \sigma_n)$, so that the allocations generated are the same. Furthermore, the equilibrium inequalities (14) for $\pi$ imply the incentive compatible inequalities (6) for $\pi'$. Thus $\mathbf{x} = \mathbf{W}(\pi, \sigma_1, \ldots, \sigma_n) \in F^*$ implies $\mathbf{x} = \mathbf{V}(\pi') \in F^*$. So $F^* \subseteq F^*$. I omit the rest of proof.

Q.E.D.
Acknowledgments

The author gratefully acknowledges helpful conversations with Dr. Hongtao Zhang. The author is also very grateful to Ms. Fang Chen, Hanyue Wu (Apple), Hanxing Wu (Lily) and Hanchen Wu (Cindy) for their great support.

References

