# Bugs in the proofs of revelation principle

Wu, Haoyang

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

19 August 2010

# A note on revelation principle from an energy perspective

Haoyang Wu *

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

**Abstract**

The revelation principle has been known in the economics society for decades. In this paper, I will investigate it from a physical perspective, i.e., considering the energy consumed by agents and the designer in participating a mechanism.

JEL codes: D7

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

## 1 Introduction

The revelation principle is a fundamental theorem in economics theory. According to the wide-spread textbook given by Mas-Colell, Whinston and Green (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.*"

So far, the revelation principle has been applied to many disciplines such as auction, contract, the theory of incentives and so on. If we move eyes from economics to physics, it is well-known that the world is a physical world, doing any action requires energy. In this paper, I will investigate the revelation principle from a physical perspective, i.e., studying how much energy is required for agents and the designer in participating a mechanism. Section 2 and 3 are the main parts of this paper. Section 4 draws conclusions. Related definitions

* Corresponding author.

*Email address:* hywch@mail.xjtu.edu.cn, Tel: 86-18621753457 (Haoyang Wu).

and proofs are given in Appendix, which are cited from Section 23.B and 23.D [1].

## 2 Energy matrices

Let us consider a setting with $I$ agents, indexed by $i = 1, \cdots, I$ (page 858 [1]). 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 $u_i(x, \theta_i)$. A mechanism $\Gamma = (S_1, \cdots, S_I, g(\cdot))$ is a collection of $I$ sets $S_1, \cdots, S_I$, each $S_i$ containing agent $i$'s possible actions (or plans of action), and an outcome function $g : S \to X$, where $S = S_1 \times \cdots \times S_I$ (page 883, Line 7 [1]).

At first sight, it looks trivial to discriminate the exact format of agent $i$'s strategy. Because the two formats of strategies, actions and plans of action, just correspond to the same results in the traditional theory of mechanism design. However, from a physical perspective, an action should be viewed different from a plan of action.

For any agent $i$, if his strategy $s_i(\cdot)$ is of an action format, I denote by $E_a$ the energy required for agent $i$ to choose it (i.e., performing the action). Otherwise agent $i$'s strategy $s_i(\cdot)$ is of a message format (i.e., a plan of action), and I denote by $E_m$ the energy required for agent $i$ to choose it (i.e., selecting the message). Generally speaking, an action is laborious, to carry out it requires more energy; whereas a plan of action is an oral message, to select it requires less energy. This is consistent to the common sense in the real world. Therefore, it is natural to assume $E_a > E_m$. Note the private type of agent $i$ can also be represented as a message, because agent $i$ can announce it to the designer. In addition, I define by $E_{send}$ and $E_g$ the energy consumed in sending out a message and performing the outcome function $g(\cdot)$ respectively.

Now let us consider 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, and the corresponding direct revelation mechanism $\Gamma_{direct} = (\Theta_1, \cdots, \Theta_I, g(s^*(\cdot)))$. Let us consider two different cases:

Case 1: $\Gamma$ is oral, in which each agent $i$'s strategy is of a message format (i.e., a plan of action).

1) Participating $\Gamma$: Given any $\theta \in \Theta$, each agent $i$ selects the strategy $s_i^*(\theta_i)$ and send it to the designer. Hence, the energy consumed by $I$ agents is $I\cdot(E_m+E_{send})$. The designer receives $I$ messages and perform the outcome function $g(\cdot)$. Hence, the energy consumed by the designer is $E_g$.
2) Participating $\Gamma_{direct}$: Given any $\theta \in \Theta$, each agent $i$ announces a type as a message to the designer. Hence, the energy consumed by $I$ agents is $I \cdot E_{send}$. The designer receives $I$ messages and perform the outcome function $g(s^*(\cdot))$. Hence, the energy consumed by the designer is $I \cdot E_m + E_g$.

Case 2: $\Gamma$ is laborious, in which each agent $i$'s strategy is of an action format.
1) Participating $\Gamma$: Given any $\theta \in \Theta$, each agent $i$ performs his action $s_i^*(\theta_i)$. Hence, the energy consumed by $I$ agents is $I \cdot E_a$. The designer perform the outcome function $g(\cdot)$. Hence, the energy consumed by the designer is $E_g$.
2) Participating $\Gamma_{direct}$: Given any $\theta \in \Theta$, each agent $i$ announces a type as a message to the designer. Hence, the energy consumed by $I$ agents is $I \cdot E_{send}$. The designer receives $I$ messages and perform the outcome function $g(s^*(\cdot))$. Hence, the energy consumed by the designer is $I \cdot E_a + E_g$.

*Table 1: An energy matrix of $I$ agents and the designer. The first entry denotes the energy consumed by $I$ agents, and the second stands for the energy consumed by the designer.*

| Strategy format \ Mechanism | $\Gamma$ | $\Gamma_{direct}$ |
|---|---|---|
| Oral (a message) | $[I \cdot (E_m + E_{send}), E_g]$ | $[I \cdot E_{send}, I \cdot E_m + E_g]$ |
| Laborious (an action) | $[I \cdot E_a, E_g]$ | $[I \cdot E_{send}, I \cdot E_a + E_g]$ |

Usually, $E_m$, $E_g$ and $E_{send}$ are small. Suppose they can be neglected, then Table 1 is reduced to Table 2:

*Table 2: A simplified energy matrix of $I$ agents and the designer.*

| Strategy format \ Mechanism | $\Gamma$ | $\Gamma_{direct}$ |
|---|---|---|
| Oral (a message) | $[0, 0]$ | $[0, 0]$ |
| Laborious (an action) | $[I \cdot E_a, 0]$ | $[0, I \cdot E_a]$ |

In terms of computer science, when agents' strategies are actions instead of plans of action, the complexity of the energy consumed by the designer in $\Gamma_{direct}$ is $\mathcal{O}(I)$, which cannot be neglected. Therefore, in order to make the direct revelation mechanism $\Gamma_{direct}$ work, an energy condition should be added: *The designer possesses enough energy, at least the sum of energy that all agents would consume when they participate the original indirect mechanism* $\Gamma$.

## 3 Discussions

In this section, I will propose two problems in front of the designer when the strategies of agents are of an action format:
1) In the direct mechanism $\Gamma_{direct}$, does the designer possess enough energy to carry out all actions that would be done by agents in the original indirect mechanism $\Gamma$? (Generally speaking, there are many factors that may be relevant to agents' actions, e.g., skill, energy, quality etc. For simplicity, here I only consider one indispensable factor, i.e., the energy required to carry out an action.)

According to Page 378, the 9th line to the last [2], "*... the mechanism designer is always at an informational disadvantage with respect to the agents, who, as a collective entity, know more about the true environment that does the designer*". Based on this idea, it looks somewhat "unreasonable" to assume that the designer is always at an energy advantage with respect to the agents, i.e., the designer possesses enough energy that is not less than the sum of all agents' energy.

As shown in Table 2, the energy condition is very weak when the strategies of agents are of a message format. However, when the strategies of agents are of an action format, the energy condition may be restrictive. The designer cannot take it for granted that he is always able to carry out all actions on behalf of all agents. When the power of the designer is restricted such that the energy condition does not hold, the revelation principle will not hold.

2) Furthermore, even if the energy condition is satisfied, there still exists another problem for the designer. As shown in Table 2, when the designer chooses the indirect mechanism $\Gamma$, he nearly spends zero energy; but if the designer chooses the direct mechanism $\Gamma_{direct}$, he has to spend $I \cdot E_a$ energy to make $\Gamma_{direct}$ work. Note that in the theory of mechanism design, the designer only care whether and how the social choice function $f(\cdot)$ can be implemented. Since $\Gamma$ and $\Gamma_{direct}$ implement the same $f(\cdot)$ in Bayesian Nash equilibrium, *why does the designer have incentives to work harder, i.e., choose $\Gamma_{direct}$ instead of $\Gamma$?* A possible answer is that the revelation principle may be not proper for a social choice function that is implemented by a "laborious" indirect mechanism in Bayesian Nash equilibrium.

## 4 Conclusion

In this paper, I propose that: 1) If an indirect mechanism is oral (i.e., the strategies of agents are of a message format), then there is no problem in the revelation principle (Note: this result holds under the assumption that $E_m$, $E_g$ and $E_{send}$ can be neglected). 2) If an indirect mechanism is laborious (i.e., the strategies of agents are of an action format), then an energy condition should be added to make the revelation principle hold in the real world. Furthermore, it is questionable to claim that the designer has incentives to work harder, but finally implement the same social choice function. Hence, the revelation principle is perhaps not proper for a social choice function that can only be implemented by a "laborious" indirect mechanism.

**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.

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

**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.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.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 $\theta_i \in \Theta_i$,

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

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 $\theta_i \in \Theta_i$,

$$E_{\theta_{-i}}[u_i(f(\theta_i, \theta_{-i})), \theta_i)|\theta_i] \geq E_{\theta_{-i}}[u_i(f(\hat{\theta}_i, \theta_{-i}), \theta_i)|\theta_i], \qquad (23.D.1)$$

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 $\theta_i \in \Theta_i$,

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

for all $\hat{s}_i \in S_i$. Condition (23.D.2) implies that for all $i$ and all $\theta_i \in \Theta_i$,

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

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 $\theta_i \in \Theta_i$,

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

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. Q.E.D.

## References


[1] A. Mas-Colell, M.D. Whinston and J.R. Green, Microeconomic Theory, Oxford University Press, 1995.

[2] R. Serrano, The theory of implementation of social choice rules, *SIAM Review* **46** (2004) 377-414.