

セキユア遂行

——理論と実験†——

西條辰義*・大和毅彦**

真の選好表明が支配戦略であることを要求する戦略的操作不能性は、社会選択理論において重要な概念として多くの研究がなされてきた。しかしながら、この概念は重大な欠陥を持っている。特に、ほとんどの戦略的操作不能なメカニズムには、たくさんのナッシュ均衡が存在し、悪い均衡結果を導きうる。この問題に対する一つの解決方法は、支配戦略均衡とナッシュ均衡での両方における二重遂行、すなわちセキユア遂行を要求することである。二つの戦略的操作不能なメカニズムに関する実験を行った。一つのメカニズムはセキユアではなく、多くの悪いナッシュ均衡を持ち、他のメカニズムはセキユアで、唯一の良いナッシュ均衡を持つ。被験者が支配戦略をとる比率は、セキユアでないメカニズムの方が、セキユアなメカニズムよりも有意に低いことを観察した。

1. 序

社会の制度そのものを与件とはせずに変数と扱うメカニズム・デザインや社会選択の分野においては、戦略的操作不能性(strategy-proofness)は基本概念である¹⁾。戦略的操作不能性は真の選好表明が支配戦略であることを要求する²⁾。合理的な個人は支配戦略をとらないはずがないという暗黙の前提がある。

一方で、この前提の可否が実験室で確認され始めている。たとえば、戦略的操作不能性を満たすセカンド・プライス・オークションの実験において Kagel-Harstad-Levin(1987), Kagel-Levin(1993), Harstad(2000)らは、被験者は真の値をほとんど表明しないことを報告している。さらには、セカンド・プライス・オークションと数学的には同値であるピボタル・メカニズムの実験において、Attiyeh-Franciosi-Isaac(2000)や Kawagoe-Mori(2001)は、被験者は半分以上しか支配戦略を選択しないことを観測している。このような実験結果をもとに Attiyeh-Franciosi-Isaac(2000)は、以下のような悲観的見解を表明するに至っている。

“we do not believe that the pivot mechanism warrants further practical consider-

ation... This is due to the fundamental failure of the mechanism, in our laboratory experiments, to induce truthful value revelation.”

戦略的操作不能性を満たさないしは動機整合的なメカニズムは20世紀後半から現在に至るまで数多くデザインされている。多くのメカニズム・デザイナーは彼らのデザインするメカニズムが社会の中で用いられることを期待したに違いない。ところが、戦略的操作不能性を満たす最も基本的なメカニズムであるピボタル・メカニズムとセカンド・プライス・オークションが機能しないとなると、この分野そのものの有効性に疑義が生ずることになる。

戦略的操作不能性のどこが問題なのだろうか。Saijo-Sjöström-Yamato(2006)は、戦略的に操作不能なメカニズムには支配戦略以外に数多くのナッシュ均衡が存在することに着目する。たとえば、ピボタル・メカニズムにおいては、次の節で示すように、戦略空間のほぼ半分程度がナッシュ均衡になってしまう。さらには、ピボタル・メカニズムにおける判断が公共財供給に「イエス」であるとしても、ナッシュ均衡のうち半分程度は公共財供給に「ノー」という結果を出してしまう。この意味で、「悪い」ナッシュ均

衡が数多く存在するのである。

そこで, Saijo-Sjöström-Yamato(2006)は, メカニズムのナッシュ均衡と支配戦略均衡が一致せねばならないという新たな遂行条件を提案し, これをセキュア遂行(secure implementation)と呼んでいる。セキュア遂行は二つの均衡概念に基づく戦略が一致することを要求しているが, 実は, 支配戦略均衡とナッシュ均衡の間に位置する均衡概念のすべてを含む多重遂行概念であることに注意したい。つまり, 彼らは, どのような行動原理で人々が行動するかを研究者があらかじめ選択するという方法論に異議を唱えていると考えてもよい。非協力的な環境のもとで, 強い合理性として支配戦略均衡を採用し, 弱い合理性としてナッシュ均衡を採用するのである。その間にある「合理性」をすべて含むように遂行概念を提案したのがセキュア遂行である。

彼らによると, ある社会選択関数がセキュア遂行可能であるためには, それが戦略的に操作不能であると共に長方形条件(rectangular property)を満たすことが必要十分である。もちろんピボタル・メカニズムは長方形条件を満たさないが, Groves-Clarke メカニズムにおいて選好が単峰性を満たすなら, 効率性の条件を満たす社会選択関数をセキュアに遂行できるのである。

セキュア遂行とそうでない場合の差をみるために, ピボタル・メカニズムと選好が単峰性を満たす場合の Groves-Clarke メカニズムの違いを実験室で検討したのが, Cason-Saijo-Sjöström-Yamato(2006)である。彼らの実験においては, メカニズムの複雑さや被験者の無理解を避けるために, メカニズムそのものを被験者に提示せず, メカニズムから構築できる利得表のみを提示するという手法を用いている。こうすることによって, ピボタル・メカニズムができるだけ機能する環境を整備することができる。残念ながら, このような環境でも被験者の選択のうち支配戦略であるのは半分であった。一方, 同じような環境のもとでの Groves-Clarke 実験における被験者の選択のうち支配

戦略であったのは81%であった。

Cason-Saijo-Sjöström-Yamato(2006)および Saijo-Sjöström-Yamato(2006)の結果の意味するところは重要である。新たに戦略的操作不能なメカニズムをデザインするのであれば, それが長方形条件を満たしているのかどうか確認せねばならない。満たさないのであれば, そのメカニズムが被験者を用いる実験で機能するのかどうか検証せねばならない。

本論文の構成は以下の通りである。第二節で Saijo-Sjöström-Yamato(2006)のセキュア遂行の理論の骨格を提示する。これを受けて第三節では Cason-Saijo-Sjöström-Yamato(2006)の実験結果を要約し, セキュア・メカニズムとノン・セキュア・メカニズムの違いを検討する。第四節では今後の課題を展望する。

2. セキュア遂行の理論

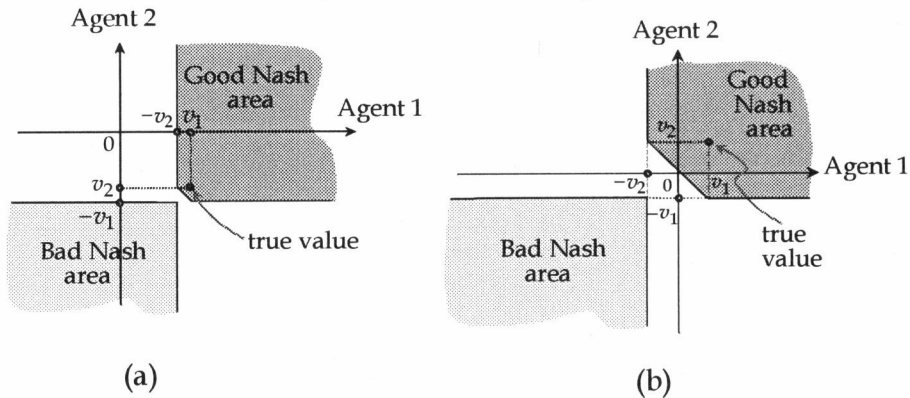
2.1 戦略的操作不能なメカニズムの問題点

メカニズム・デザインの文献においては, 戦略的操作不能なメカニズムがさかんに研究されてきた。しかしながら, それらの多くにおいて, 支配戦略均衡以外に多数のナッシュ均衡が存在する。さらには, これらのナッシュ均衡では, 社会的に望ましくない「悪い」結果をしばしば導く。これらの点を, いくつかの実験が行われてきた以下の二つの戦略的操作不能なメカニズムを例としてみよう。

1) ピボタル・メカニズム(Clarke(1971))。

二人の主体1と2が, ある公共プロジェクトを実施するか否か(ある一定量の非排除的な公共財を生産するか否か)を決める問題に直面しているとしよう。各主体は準線形の効用関数を持ち, 主体 i の公共プロジェクトに関する純便益の真の値は, もしプロジェクトが実施されたならば v_i , 実施されなければ0であるとしよう($i=1, 2$)。ピボタル・メカニズムでは, 各主体 i は自分の純便益の値 $\bar{v}_i$ を報告する。もちろん, この値は真の便益である必要はない。プロジェクトを実施するか否かと各主体が支払う税金は以下のように決まる。

図1. ピボタル・メカニズムの均衡



ルール1: もし $\bar{v}_1 + \bar{v}_2 \geq 0$ ならば, プロジェクトを実施する. $\bar{v}_1 + \bar{v}_2 < 0$ ならば, プロジェクトは実施しない.

ルール2: 各主体 i は以下のピボタル税 t_i を支払う.

$$\begin{aligned} t_i &= -\bar{v}_j & \text{if } \bar{v}_j < 0 \text{ and } \bar{v}_1 + \bar{v}_2 \geq 0 \\ &= \bar{v}_j & \text{if } \bar{v}_j > 0 \text{ and } \bar{v}_1 + \bar{v}_2 < 0 \\ &= 0 & \text{otherwise} \quad (j \neq i). \end{aligned}$$

いま (v_1, v_2) を真の純便益の値とする. まず最初に $v_1 > 0, v_2 < 0, v_1 + v_2 > 0$ と仮定しよう. この場合, 真の純便益の値の和は正なので, プロジェクトは実施されるべきである. ピボタル・メカニズムでは, 各主体 i は自分の真の値 v_i を報告することが支配戦略となる. しかしながら, 図1-(a)はこの支配戦略均衡以外に数多くのナッシュ均衡が存在し, ナッシュ均衡の集合はほぼ2次元実数空間の半分の大きさであることを示している. ナッシュ均衡の右上の部分は, プロジェクトが実施されるという意味で, 「良い」均衡の集合である. しかしながら, ナッシュ均衡の左下の部分は, プロジェクトが実施されないという意味で, 「悪い」均衡の集合である.

次に $v_1 \geq v_2 > 0$ と仮定しよう. この場合, 二人の主体ともプロジェクトを実施したいと思っている. しかしながら, 図1-(b)で示されているように, プロジェクトの実施されない悪いナッシュ均衡の集合は依然として大きい. Saijo-Sjöström-Yamato(2006)は, この否定的な結論を一般化して, 任意の有限個のプロジェクト

と主体が存在するケースでも成立することを示している.

2) セカンド・プライス・オークション (Vickrey(1961))

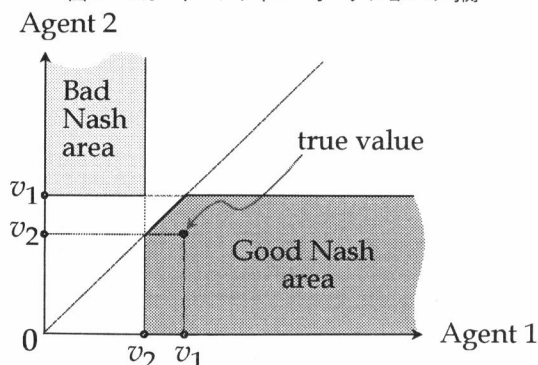
二人の主体1と2が, 分割できない財をどちらが受け取るかを決める問題に直面しているとしよう. 主体 i が財を手にとり得た場合に得られる便益の真の値は $v_i \geq 0$, 財を手にとり得なかった場合には0である ($i=1, 2$). 主体 i の入札額 $\bar{v}_i$ をとしよう. セカンド・プライス・オークションは以下の二つのルールからなる.

ルール1: $\bar{v}_i > \bar{v}_j$ のとき, 主体 i が財を受け取り, $\bar{v}_j$ を支払う ($i, j=1, 2; i \neq j$).

ルール2: $\bar{v}_1 = \bar{v}_2$ のとき, 主体1が財を受け取り, $\bar{v}_2$ を支払う.

(v_1, v_2) を真の便益の値としよう. いま, $v_1 > v_2 > 0$ と仮定する. この時, 主体1の便益が2の便益より大きいので, 主体1が財を受け取るべきである. セカンド・プライス・メカニズムでは, 各主体 i が自分の真の値 v_i を入札することが支配戦略となる. しかしながら, 図2が示すように, ナッシュ均衡の集合はきわめて大きい. ナッシュ均衡の右下の部分は, 主体1が財を受け取るという意味で, 「良い」均衡の集合である. しかし, ナッシュ均衡の左上の部分は, 主体2が財を受け取るという意味で, 「悪い」均衡の集合である.

図2. セカンド・プライス・オークションの均衡



セカンド・プライス・オークションの実験で、なぜ被験者が支配戦略をプレイしないかについては、それが明確でなく、ルールをよく理解できないで混乱したからという理由ももちろん考えられる(例えば Harstad(2000)を参照)。しかしながら、図2に示されているように、セカンド・プライス・オークションにナッシュ均衡が数多くある事実は、支配戦略均衡からのランダムな乖離というよりは、むしろ体系的な乖離を生じる可能性を示唆していると言えよう。

同様の問題を抱える他の戦略的操作不能なメカニズムとしては、排除可能な公共財(excludable public good)供給のためのシリアル・コスト・シェアリング・メカニズム、投票環境において選好が単峰性を満たす場合でのコンドールセ・スキーム(メディアン・ボーター・スキームの一種)、交換経済の環境において選好が単峰性を満たす場合におけるユニフォーム配分ルール(固定価格取引ルールの一環)、および非分割財市場におけるトップ・トレーディング・サイクル・ルールなどがある。これらの代表的な戦略的操作不能なメカニズムの各々においても、多数のナッシュ均衡が存在し、さらには、次の意味で「悪い」結果になるナッシュ均衡が存在しうる。シリアル・コスト・シェアリング・メカニズムについては、公共財に対する純便益の合計が正であるときでも、公共財が供給されない。また、コンドールセ・スキーム、ユニフォーム配分ルールとトップ・トレーディング・サイクル・ルールに関しては、パレート効率的な配分が達成されないのである(Saijo-

Sjöström-Yamato(2003)参照)。

2.2 公共財経済におけるセキユア遂行

前節で見たように、戦略的操作不能メカニズムの多くが、支配戦略均衡以外に数多くの「悪い」ナッシュ均衡が存在しうる。これらのメカニズムは、社会選択関数を支配戦略で遂行できているものの、ナッシュ均衡では遂行できていないのである。この問題を解決するために、Saijo-Sjöström-Yamato(2006)は「セキユア遂行(secure implementation)」という新しい概念を提唱した。本節では、準線形の選好を持つ二人の主体、一つの私的財、一つの公共財からなる経済におけるセキユア遂行について述べる。まず、実現可能な配分の集合を

$$A = \{(y, t_1, t_2) | y \in Y, t_1, t_2 \in \mathbb{R}\}$$

と表そう。ただし、ここで $Y \subseteq \mathbb{R}$ は生産可能性集合、 $y \in Y$ は公共財の生産量水準、 t_i は主体 i に対する私的財の移転(transfer)である。ここでは、公共財の生産コストはゼロであると仮定する。各主体 i の効用関数 $u_i: A \rightarrow \mathbb{R}$ は利己的で準線形である：

$$u_i(y, t_1, t_2) = u_i(y, t_i) = v_i(y) + t_i, i = 1, 2$$

主体 i のとりうる全ての評価関数 $v_i: Y \rightarrow \mathbb{R}$ のクラスを V_i で表す。また、評価プロファイル $v = (v_1, v_2) \in V \equiv V_1 \times V_2$ としよう。

社会選択関数(social choice function)は、各評価プロファイル $v \in V$ に対して、ただ一つの実現可能な配分 $f(v) \in A$ を割り当てる関数 $f: V \rightarrow A$ である。社会選択関数は社会の目標を表すもので、 $f(v)$ を v のもとで f -最適な結果と呼ぶ。

各主体 i の戦略の集合を S_i とする。メカニズム(もしくはゲーム・フォーム)は、各戦略プロファイル $(s_1, s_2) \in S_1 \times S_2$ に、 A に属するただ一つの結果を割り当てる関数 $g: S_1 \times S_2 \rightarrow A$ である。戦略プロファイル $s = (s_1, s_2) \in S_1 \times S_2$ のとき、メカニズム g の結果を $g(s) = (y^g(s), t^g(s))$ と書く。ここで、 $y^g(s)$ は公共財の水準、 $t^g(s) = (t_1^g(s), t_2^g(s))$ は移転ベクトルである。

戦略プロファイル $s = (s_1, s_2) \in S_1 \times S_2$ が、メカニズム g の評価プロファイル $v \in V$ におけ

るナッシュ均衡であるのは,

$$v_1(y^g(s_1, s_2)) + t_1^g(s_1, s_2) \geq v_1(y^g(s'_1, s_2)) + t_1^g(s'_1, s_2), \quad \forall s'_1 \in S_1$$

$$v_2(y^g(s_1, s_2)) + t_2^g(s_1, s_2) \geq v_2(y^g(s_1, s'_2)) + t_2^g(s_1, s'_2), \quad \forall s'_2 \in S_2$$

が成立するときである。\$N^g(v)\$ を \$g\$ の \$v\$ におけるナッシュ均衡の集合としよう。

戦略プロファイル \$s=(s_1, s_2) \in S_1 \times S_2\$ が, メカニズム \$g\$ の評価プロファイル \$v \in V\$ における支配戦略均衡であるのは,

$$v_1(y^g(s_1, s_2)) + t_1^g(s_1, s_2) \geq v_1(y^g(s'_1, s_2)) + t_1^g(s'_1, s_2), \quad \forall s'_1 \in S_1, \quad \forall s_2 \in S_2$$

$$v_2(y^g(s_1, s_2)) + t_2^g(s_1, s_2) \geq v_2(y^g(s_1, s'_2)) + t_2^g(s_1, s'_2), \quad \forall s_1 \in S_1, \quad \forall s'_2 \in S_2$$

\$DS^g(v)\$ を \$g\$ の \$v\$ における支配戦略均衡の集合としよう。

定義 1. 以下の条件を満たすとき, メカニズム \$g\$ は社会選択関数 \$f\$ を支配戦略均衡で遂行する (*implement in dominant strategy equilibria*) という: すべての \$v \in V\$ について, (i) \$g(s)=f(v)\$ であるような支配戦略均衡 \$s \in DS^g(v)\$ が存在する. (ii) 任意の支配戦略均衡 \$s \in DS^g(v)\$ について, \$g(s)=f(u)\$。社会選択関数 \$f\$ を支配戦略均衡で遂行するメカニズムが存在するとき, \$f\$ は支配戦略遂行可能である (*dominant strategy implementable*) という。

定義 2. 以下の条件を満たすとき, メカニズム \$g\$ は社会選択関数 \$f\$ をセキュア遂行する (*securely implement*) という: すべての \$v \in V\$ について, (i) \$g(s)=f(v)\$ であるような支配戦略均衡 \$s \in DS^g(v)\$ が存在する. (ii) 任意のナッシュ均衡 \$s \in N^g(v)\$ について, \$g(s)=f(u)\$。社会選択関数 \$f\$ をセキュア遂行するメカニズムが存在するとき, \$f\$ はセキュア遂行可能である (*securely implementable*) という。

支配戦略均衡での遂行は, すべての可能な選好プロファイルについて, 全ての支配戦略での配分が \$f\$-最適であることを要求する。この要請に加えて, セキュア遂行は, 支配戦略均衡以外の全てのナッシュ均衡でも, 配分が \$f\$-最適

であることを要求する。

Saijo-Sjöström-Yamato(2006)は, 以下の二つの条件を用いて, セキュア遂行可能な社会選択関数のクラスの特徴づけを行った。評価プロファイル \$v=(v_1, v_2)\$ のもとで社会選択関数 \$f\$ が指定する配分を \$f(v)=(y^f(v), t^f(v))\$ で表そう。ただし, ここで \$y^f(v)\$ は公共財の水準, \$t^f(v)=(t_1^f(v), t_2^f(v))\$ は移転ベクトルである。

定義 3. 社会選択関数 \$f\$ が戦略的操作不能 (*strategy-proof*) であるのは,

$$\begin{aligned} & v_1(y^f(v_1, \bar{v}_2)) + t_1^f(v_1, \bar{v}_2) \\ & \geq v_1(y^f(\bar{v}_1, \bar{v}_2)) + t_1^f(\bar{v}_1, \bar{v}_2), \\ & \quad \forall \bar{v}_1 \in V_1, \quad \forall \bar{v}_2 \in V_2 \\ & v_2(y^f(\bar{v}_1, v_2)) + t_2^f(\bar{v}_1, v_2) \\ & \geq v_2(y^f(\bar{v}_1, \bar{v}_2)) + t_2^f(\bar{v}_1, \bar{v}_2), \\ & \quad \forall \bar{v}_1 \in V_1, \quad \forall \bar{v}_2 \in V_2 \end{aligned}$$

が成立するときである。

支配戦略遂行可能性と戦略的操作不能性については, 以下の結果がよく知られている。

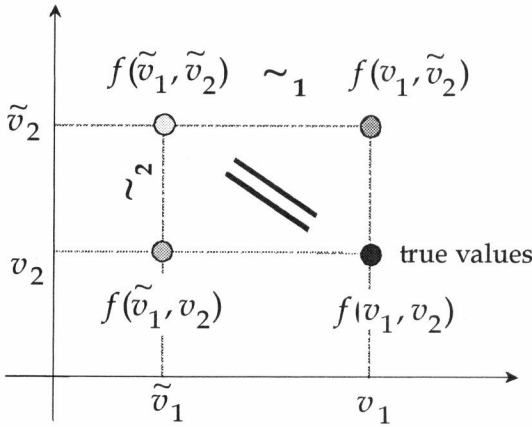
命題 1(支配戦略遂行に関する表明原理 (Revelation Principle, Gibbard(1973)))。もし社会選択関数 \$f\$ が支配戦略遂行可能ならば, \$f\$ は戦略的操作不能である。

戦略的操作不能性は, 支配戦略遂行の必要条件であるので, セキュア遂行の必要条件にもなっている。しかしながら, もう一つ別の条件がセキュア遂行のために必要となる。このことを直感的に理解するために, いま, 直接表明メカニズム \$g=f\$ が社会選択関数 \$f\$ をセキュア遂行するとしよう。\$n=2\$ として, \$(v_1, v_2)\$ を真の評価プロファイルとしよう。図 3 を見よ。\$u_1(f(v_1, \bar{v}_2))=u_1(f(\bar{v}_1, \bar{v}_2))\$ と仮定しよう, つまり,

$$\begin{aligned} & v_1(y^f(v_1, \bar{v}_2)) + t_1^f(v_1, \bar{v}_2) \\ & = v_1(y^f(\bar{v}_1, \bar{v}_2)) + t_1^f(\bar{v}_1, \bar{v}_2). \end{aligned} \quad (1)$$

すなわち, 主体 2 が \$\bar{v}_2\$ を表明する時, 主体 1 にとって, 真の選好 \$v_1\$ を表明することと \$\bar{v}_1\$ を表明することが無差別であるとしよう。戦略的操作不能性より, \$v_1\$ を表明することが支配戦略な

図3. 長方形条件



ので, (1)から

$$\begin{aligned} & v_1(y^f(\tilde{v}_1, \tilde{v}_2)) + t_1^f(\tilde{v}_1, \tilde{v}_2) \\ &= v_1(y^f(v_1, \tilde{v}_2)) + t_1^f(v_1, \tilde{v}_2) \\ &\geq v_1(y^f(v'_1, \tilde{v}_2)) + t_1^f(v'_1, \tilde{v}_2), \quad \forall v'_1 \in V_1. \end{aligned}$$

つまり, 主体2が $\tilde{v}_2$ を表明したとき, $\tilde{v}_1$ を表明することが主体1にとって最適反応となっている. 次に, $u_2(f(\tilde{v}_1, v_2)) = u_2(f(\tilde{v}_1, \tilde{v}_2))$, すなわち,

$$\begin{aligned} & v_2(y^f(\tilde{v}_1, v_2)) + t_2^f(\tilde{v}_1, v_2) \\ &= v_2(y^f(\tilde{v}_1, \tilde{v}_2)) + t_2^f(\tilde{v}_1, \tilde{v}_2) \end{aligned} \quad (2)$$

と仮定しよう. 上と同様の議論により, 任意の $v'_2 \in V_2$ について,

$$\begin{aligned} & v_2(y^f(\tilde{v}_1, \tilde{v}_2)) + t_2^f(\tilde{v}_1, \tilde{v}_2) \\ &= v_2(y^f(\tilde{v}_1, v_2)) + t_2^f(\tilde{v}_1, v_2) \\ &\geq v_2(y^f(\tilde{v}_1, v'_2)) + t_2^f(\tilde{v}_1, v'_2), \end{aligned}$$

つまり, 主体1が $\tilde{v}_1$ を表明するとき, $\tilde{v}_2$ を表明することが主体2にとって最適反応となっている. これゆえ, $f(\tilde{v}_1, \tilde{v}_2) = (y^f(\tilde{v}_1, \tilde{v}_2), t^f(\tilde{v}_1, \tilde{v}_2))$ はナッシュ均衡配分である. さらに, $f(v_1, v_2) = (y^f(v_1, v_2), t^f(v_1, v_2))$ は支配戦略均衡配分なので, セキユア遂行可能性より, 支配戦略均衡配分 $f(v_1, v_2)$ とナッシュ均衡配分 $f(\tilde{v}_1, \tilde{v}_2)$ は一致する. 以上のことをまとめると, (1)と(2)が成立するとき, $f(v_1, v_2) = f(\tilde{v}_1, \tilde{v}_2)$ でなければならないことになる.

上で説明した条件を定式化すると以下のようになる.

定義4. 以下の条件を満たすとき, 社会選択関

数 f は長方形条件(rectangular property)を満たすという. 任意の $v, \tilde{v} \in V$ について, もし

$$\begin{aligned} & v_1(y^f(v_1, \tilde{v}_2)) + t_1^f(v_1, \tilde{v}_2) \\ &= v_1(y^f(\tilde{v}_1, \tilde{v}_2)) + t_1^f(\tilde{v}_1, \tilde{v}_2) \text{ かつ} \\ & v_2(y^f(\tilde{v}_1, v_2)) + t_2^f(\tilde{v}_1, v_2) \\ &= v_2(y^f(\tilde{v}_1, \tilde{v}_2)) + t_2^f(\tilde{v}_1, \tilde{v}_2) \end{aligned}$$

ならば, そのとき $f(v_1, v_2) = f(\tilde{v}_1, \tilde{v}_2)$ である.

Saijo-Sjöström-Yamato(2006)は, 戦略的操作不能性と長方形条件が必要かつ十分条件になることを示した.

定理1. 社会選択関数 f がセキユア遂行可能である時かつその時のみ f は戦略的操作不能性と長方形条件を満たす.

以下の公共財供給に関する効率性の条件を満たす社会選択関数 f について考えよう.

$$y^f(v_1, v_2) \in \arg \max_{y \in Y} [v_1(y) + v_2(y)]$$

$$\text{for all } (v_1, v_2) \in V. \quad (3)$$

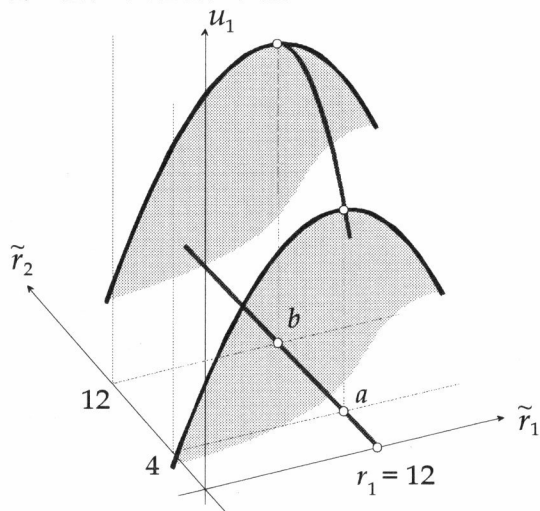
命題2(Clarke(1971), Groves(1973), Green-Laffont(1979), Holmstrom(1979)). 効率性条件(3)を満たす社会選択関数 f が支配戦略均衡で遂行可能である時かつその時のみ f は以下の条件(4)を満たす.

$$\begin{aligned} & t_1^f(v_1, v_2) = v_2(y^f(v_1, v_2)) + h_1(v_2), \\ & t_2^f(v_1, v_2) = v_1(y^f(v_1, v_2)) + h_2(v_1), \\ & \forall (v_1, v_2) \in V \end{aligned} \quad (4)$$

ここで, h_i は v_i に依存しないある任意の関数である.

(3)と(4)を満たす直接表明(direct revelation)メカニズムは, **Groves-Clarkeメカニズム**とよばれる. 命題2によると, 効率性条件(3)を満たす社会選択関数を支配戦略均衡で遂行については, Groves-Clarkeメカニズムのクラスだけを考察すればよいことになる. しかし, Saijo-Sjöström-Yamato(2006)は, 評価関数のクラス V に制約をおかず, Y が有限集合のとき, 効率性条件(3)を満たす社会選択関数を支

図4. 選好が単峰性を満たす場合の Groves-Clarke メカニズム



配戦略で遂行するどんなメカニズムについても、ナッシュ均衡での結果の集合は、支配戦略均衡での結果の集合より厳密に大きくなり、セキュア遂行が不可能であることを示している。

ただし、もし V が単峰性を満たす選好のクラスに限られ、公共財の供給水準 y が連続変数であるならば、戦略的操作不能で効率性条件(3)を満たす社会選択関数が、Groves-Clarkeメカニズムでセキュア遂行できる。いま、 $Y = \mathbb{R}$ で、 $i=1, 2$ について、

$$V_i = \{v_i: \mathbb{R} \rightarrow \mathbb{R} | v_i(y) = -(y - r_i)^2, r_i \in \mathbb{R}\}$$

と仮定しよう。ここで、 r_i は主体 i のピーク、つまり最も好ましい公共財の水準である。このような単峰性を満たす選好は、 v_i の代わりに r_i で表すことが可能である。効率性条件(3)を満たす公共財供給水準は、 $y(r_1, r_2) = (r_1 + r_2)/2$ で与えられる。このケースでは、(3)と(4)を満たす社会選択関数は長方形条件も満たすので、セキュア遂行可能である(Saijo-Sjöström-Yamato(2006))。

次節で説明する実験デザインでは、 $h_i = 0$ のケースを考える。この場合、

$$\begin{aligned} u_1(\tilde{r}_1, \tilde{r}_2) &= v_1(y(\tilde{r}_1, \tilde{r}_2)) + t_1(\tilde{r}_1, \tilde{r}_2) \\ &= -((\tilde{r}_1 + \tilde{r}_2)/2 - r_1)^2 - ((\tilde{r}_1 + \tilde{r}_2)/2 - \tilde{r}_2)^2 \\ &= -\{(\tilde{r}_1 - r_1)^2 + (\tilde{r}_2 - r_1)^2\}/2 \end{aligned}$$

ここで、 r_1 は主体1の真のピークの値、 $(\tilde{r}_1, \tilde{r}_2)$ は主体の報告したピークの値である。主体

1の利得は r_1 で最大化され、利得を最大化する値はただ一つしかない。さらに、主体2の報告値 $\tilde{r}_2$ に関係なく、主体1の利得は r_1 で最大化される。図4は、 $r_1 = 12$ における主体1の利得を表している。もし $\tilde{r}_2 = 4$ ならば、1の利得は a で最大化され、 $\tilde{r}_2 = 12$ ならば、1の利得は b で最大化される。両方のケースで、1の利得は $r_1 = 12$ で最大化される。これゆえ、主体1の最適反応曲線は、 $\tilde{r}_2$ 軸に平行となる。つまり、本当のことを報告することが、ただ一つの支配戦略で、しかも、強い意味の支配戦略である。しかしながら、真の選好表明が強い意味の支配戦略になるのは、公共財の供給水準が連続変数の場合である。我々の実験では、公共財の供給水準と利得の値は離散値をとるので、真の選好表明が強い意味の支配戦略とはならない。それにも関わらず、ただ一つの弱い意味での支配戦略は存在して、ナッシュ均衡はただ一つなので、セキュア遂行は達成可能である(次節でトリートメントSと呼ぶ)。選好が単峰性を満たさなければ、ナッシュ均衡は数多くあり、遂行はセキュアではない(トリートメントP)。

3. 実験：セキュア vs. ノン・セキュア・メカニズム

3.1 実験のデザイン

本節においては Cason-Saijo-Sjöström-Yamato(2006)におけるピボタル・メカニズムと選好が単峰性を持つ環境における Groves-Clarke メカニズムの比較実験を紹介したい。20人の被験者を一組とする実験をセッションと呼ぼう。ピボタル・メカニズムでは2つのセッションを実施し、これをトリートメントPと呼ぶ。一方、Groves-Clarkeメカニズムでも2つのセッションを実施し、これをトリートメントSと呼ぶ。1998年6月、当時の東京都立大学においてトリートメントPとトリートメントSを各々1セッションずつ実施した。一方、2003年2月、パーデュー大学においても同様の実験を実施した。

各セッションにおける被験者は2つのタイプのうちどちらかのタイプになる。各タイプの被

験者は10人ずつである。なお、被験者は2つのタイプがあることは知らされていない。被験者は自分とは異なったタイプの被験者と対戦することになるが、毎回ご破算で異なったタイプの被験者と対戦するので、同じ実験の繰り返しは高々10回である。被験者の参加時間はおよそ1時間程度であった³⁾。

トリートメントPのピボタル・メカニズムの実験においては、公共財に対する真の値を $(v_1, v_2) = (-6, 8)$ とし、公共財がない場合の値を $(v_1, v_2) = (0, 0)$ とする。 $v_1 + v_2 > 0$ なので、真の値が表明されるならば公共財は供給されることになる。タイプ1の被験者の戦略集合を-22から2までの整数の集合とし、タイプ2のそれを-4から20までの整数の集合としている。各タイプの戦略の個数は24なので、これをもとに各々のタイプで 24×24 の利得表を作成することができる。ただし、実際の実験では、2つの点でこの利得表を変換している。まず、戦略の名前を変更している。タイプ1の-4は1、-3は2という具合に戦略の名前を1から24に変えている。次に、利得の線形変換を行っている。タイプ1の場合の変換は $14v_1 + 294$ 、タイプ2の場合の変換は $14v_2 + 182$ である。表1、表2は、実験で用いた各々のタイプの利得表を示している。

以上のような利得表を用いる利点と問題点を整理しよう。Attiyeh-Franciosi-Isaac(2000)のピボタル・メカニズム実験が示しているように、メカニズムのルールのみでは真の選好を表明する被験者は稀である。そのため、ルールに加えて利得表の情報を与えらるとどうなるのかを検証したのがKawagoe-Mori(2001)である⁴⁾。ルールのみの場合だと真の選好を表明するのは20%弱、ルールに加え利得表を見せると半分近くが支配戦略をとったのである。Cason-Saijo-Sjöström-Yamato(2006)は、ピボタル・メカニズムのルールは示さず、利得表のみを用いてピボタル・メカニズムの性能の検証をしている。つまり、ゲームのルールそのものの複雑さを回避し、その利得の構造のみを抽出し、そのメカニズムのパフォーマンスを検証しようとしてい

るのである。換言するなら、ピボタル・メカニズムが最も機能しやすい環境を与え、そのパフォーマンスを見ようというアプローチである。もしこの環境で機能しないのなら、そのメカニズムは他の環境でもほぼ機能しないといってよいであろう。

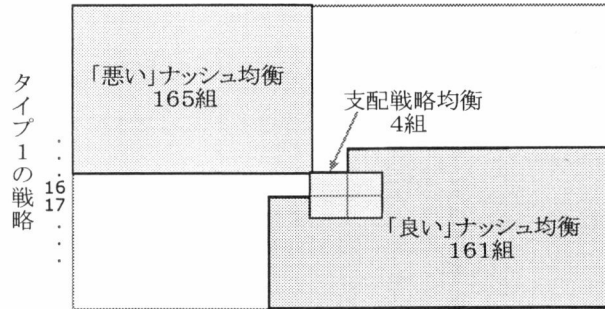
一方、この環境であるメカニズムがよいパフォーマンスを示すとしても、それは他の環境でよいパフォーマンスを示すとは必ずしも言い得ない。次のステップは、ルールを含む複雑さに関するテストのクリアである。たとえば、我々の実際の生活においては、表1のような利得表を用いて意思決定をすることはほとんどないからである。

表1、2からわかるように、タイプ1の支配戦略は16、17、タイプ2のそれは12、13である。二つの支配戦略は他の被験者がどのような選択をとっても利得が同じという意味で戦略的には同値である。もちろん、他の主体の選択によって自己の利得が変わる場合もある⁵⁾。利得表の作成プロセスからわかるように、表1、2のように変換されてしまうとピボタル・メカニズムが当初保有していた真の値という情報が消えてしまうことにも注意したい。

表1において、タイプ2の戦略をたとえば12に固定しよう。このとき、タイプ2の利得は196か210のいずれかである。ピボタル・メカニズムの場合、相手の戦略を与えるとき、自己の利得は高い(210)ほうか、低い(196)ほうのいずれしかない。そうでない場合は、利得が同じになってしまう。このため、表3が示すようにナッシュ均衡はかなりの数になる。表1、2の場合、「悪い」ナッシュ均衡の数は165、「良い」ナッシュ均衡は162個である。全てのセルの数は625(25×25)なので、そのうち52.3%がナッシュ均衡となる。

よく知られているように、ピボタル・メカニズムのアウトカムはパレート効率とはならず、余剰が発生する。政府ないしはメカニズムの実行主体がピボタル税を集めるからである。表4が示すように、タイプ1が17、タイプ2が12という支配戦略を選ぶなら、各々の利得は210

表3. トリートメントPにおけるナッシュ均衡の割合
タイプ2の戦略
... 12, 13, ...



で、その合計は420であるが、一方、(8, 5)という「悪い」ナッシュ均衡の場合、利得合計は476となり、支配戦略均衡の利得を超えてしまう。「悪い」ナッシュ均衡165個のうち、支配戦略均衡利得である(210, 196)ないしは(210, 210)と比してパレート優越されないのは150個(91%)、一方、「良い」ナッシュ均衡162個のうちそれらと比してパレート優越されないのは150個(93%)となる。さらには、「悪い」ナッシュ均衡のうち、パレート効率な均衡の個数は45.5%、「良い」ナッシュ均衡のうち、パレート効率なそれは28%である。以上のように、ピボタル・メカニズムにおいて支配戦略が選ばれなくなる「圧力」が存在するといつてよい。

表4. タイプ1が8か17、タイプ2が5か12を選ぶときの利得表

		タイプ2	
		5	12
タイプ1	8	(294, 182)	(194, 182)
	17	(294, 182)	(210, 210)

ピボタル・メカニズムは被験者の公共財の評価の和が最大になるように設計されている。この実験例の場合、公共財は供給されるべきなのに、公共財が供給されない「悪い」均衡の半数近くがパレート効率なのである。このことは、戦略的操作不能性を満たすメカニズム・デザインの理論そのものの再考を迫る可能性があることを指摘しておこう。

トリートメントSはトリートメントPと利得表のみが異なる。タイプ1, 2の真の評価関数を $v_1(y) = -(y-12)^2$, $v_2(y) = -(y-17)^2$

とする。ただし、 y は公共財の水準である。各タイプが選好表明としてピーク(たとえばタイプ1の真のピークは12)を表明するとし、 $(\tilde{r}_1, \tilde{r}_2)$ を表明値とし、公共財の水準 $y(\tilde{r}_1, \tilde{r}_2)$ およびタイプ i へのトランスファー $t_i(\tilde{r}_1, \tilde{r}_2)$ は以下の Groves-Clarke メカニズムによって決定される。

$$y(\tilde{r}_1, \tilde{r}_2) = (\tilde{r}_1 + \tilde{r}_2)/2 \text{ および}$$

$$t_i(\tilde{r}_1, \tilde{r}_2) = -((\tilde{r}_1 + \tilde{r}_2)/2 - \tilde{r}_j)^2, i, j = 1, 2; j \neq i.$$

よって各タイプの利得関数は次のようになる。

$$v_1(y(\tilde{r}_1, \tilde{r}_2)) + t_1(\tilde{r}_1, \tilde{r}_2)$$

$$= -((\tilde{r}_1 + \tilde{r}_2)/2 - 12)^2$$

$$-((\tilde{r}_1 + \tilde{r}_2)/2 - \tilde{r}_2)^2 \text{ および}$$

$$v_2(y(\tilde{r}_1, \tilde{r}_2)) + t_2(\tilde{r}_1, \tilde{r}_2)$$

$$= -((\tilde{r}_1 + \tilde{r}_2)/2 - 17)^2$$

$$-((\tilde{r}_1 + \tilde{r}_2)/2 - \tilde{r}_1)^2.$$

各タイプの戦略を0から24までの整数とする。利得表を作る際には、戦略0を戦略1などとし、1だけずらし、さらには、上記の利得関数について $10v_i/14 + 218.5$ と線形変換を施す⁶⁾。

各々の利得表ともにユニークな支配戦略が存在する。タイプ1のそれは13、タイプ2のそれは18である。なお、戦略そのものが整数という離散型なので、理論とは異なり、共に狭義の支配戦略ではない。この意味では、トリートメントPとSの利得表には差がない。ただし、トリートメントSには支配戦略以外にはナッシュ戦略はない。つまり、セキユア遂行に成功している。

「進化」的な観点から二つのトリートメント

図5. トリートメントPの実験結果

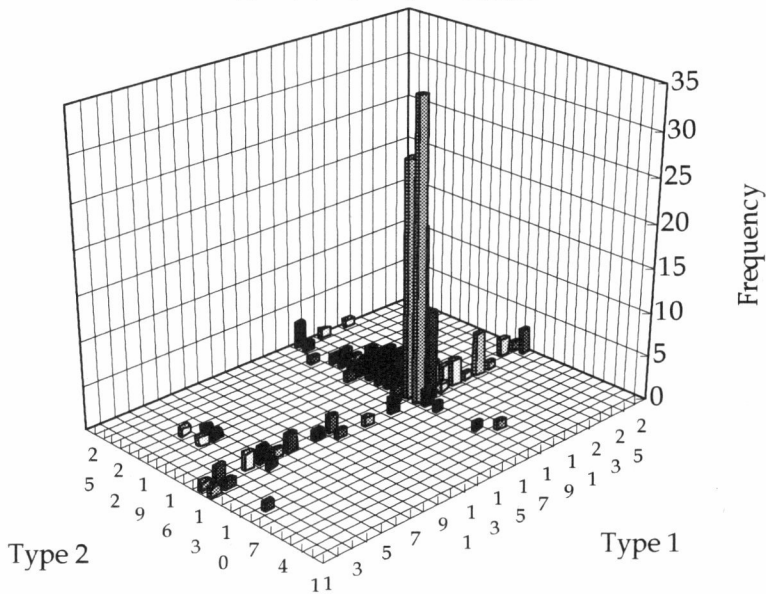
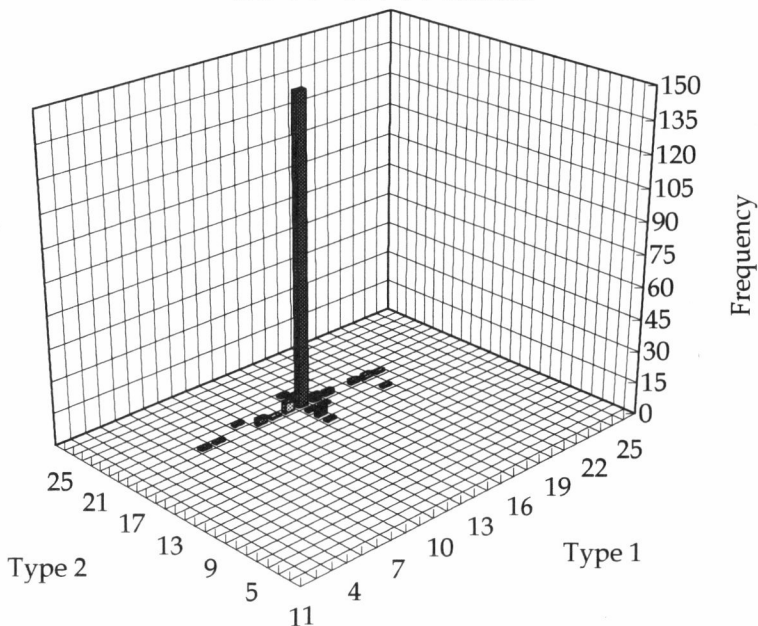


図6. トリートメントSの実験結果



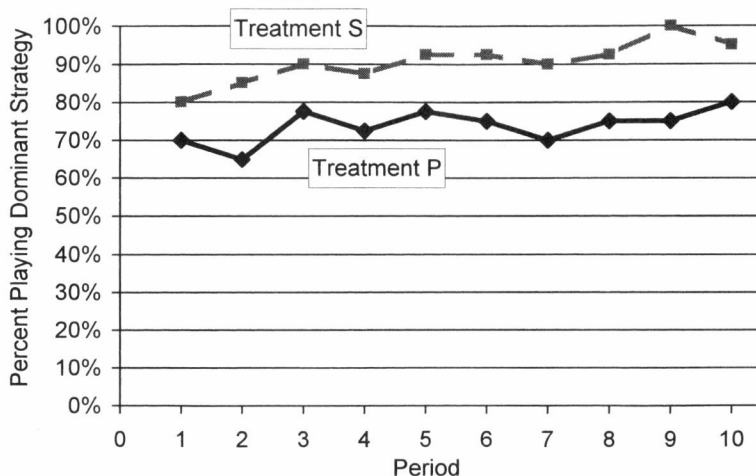
を見よう。トリートメントSの場合、支配戦略の隣り(タイプ1の場合12と14, タイプ2の場合17と19)以外の戦略は狭義に支配されている。さらには、タイプ2が17から19を選ぶ場合、タイプ1の最適反応は13となる。同様にタイプ1が12から14を選ぶ場合、タイプ2の最適反応は18である。そのため、(13, 18)への

収束は早いといえよう。一方、トリートメントPの場合、ナッシュ均衡が収束点になるいかなる学習動学を用いても、「悪い」ナッシュ均衡から抜け出せなくなる可能性が残ってしまう。

3.2 実験結果

各トリートメントの各回には10ペアのデー

図7. 被験者が支配戦略をプレイした割合



タがあり、日本の繰り返しは8回、アメリカの繰り返しは10回なので、合計180組のデータを得ることになる。図5はトリートメントP、図6はトリートメントSの頻度を示している⁷⁾。

トリートメントPの場合、4つの支配戦略均衡(16, 12), (16, 13), (17, 12), (17, 13)の頻度の合計は90であった。つまり、ちょうど半数の組が支配戦略をとったことになる。残りのうち、61組がナッシュ均衡だった。このうち、「良い」ナッシュ均衡は60個、「悪い」ナッシュ均衡は1個だった。つまり、ナッシュ均衡で公共財を供給できなかったのは1個のみである。残りの29個はナッシュ均衡ではなく、すべてにおいて公共財は供給されていない。よって、30個、つまり1/6の組において公共財が供給されなかったのである。また、ナッシュ均衡以外の組である29個は、支配戦略均衡以外の組のうち約1/3である。一方、戦略の組のすべての可能性を示す625個のセルのうち支配戦略以外の数は621個でそのうち298がナッシュ均衡ではない。つまり半分弱がナッシュ均衡ではないので、支配戦略からの乖離はランダムではなくナッシュ均衡と何らかの関係があるかもしれない。

トリートメントSの場合、支配戦略均衡は(13, 18)でこの頻度は146個であり、81%が支配戦略をとったことになる。支配戦略以外のどの組をみても頻度は3以下であった。

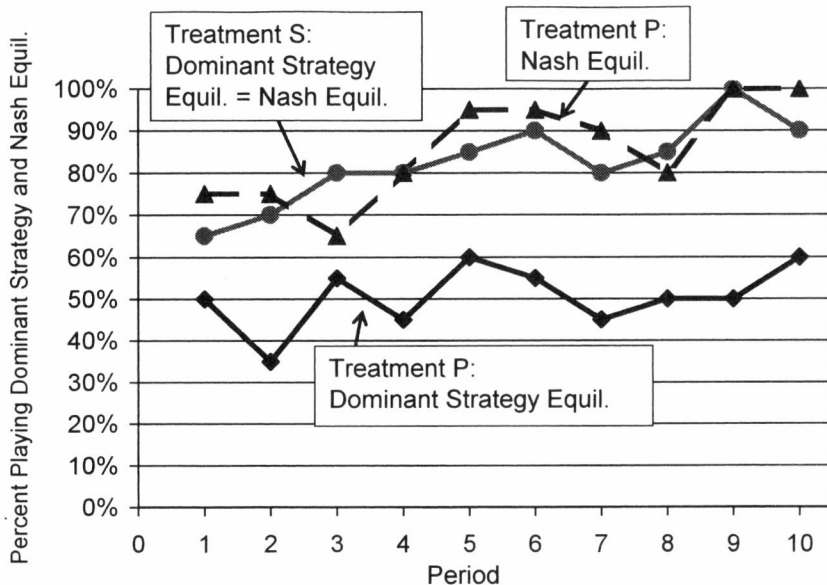
以下では二つのメカニズムを支配戦略をとった被験者の頻度と支配戦略をとった組の頻度の比較をしたい。図7は支配戦略をとった被験者の頻度の割合である。フィッシャーの直接確率検定を用いると、10回のうち7回(2, 6, 7, 8, 9回は有意水準5%, 4, 5回は有意水準10%)は、トリートメントSの被験者の方がトリートメントPの被験者よりもより多く支配戦略をとっている。

図8は組で支配戦略をとった被験者の頻度の割合である。フィッシャーの直接確率検定を用いると、10回のうち8回(2, 4, 6, 7, 8, 9回は有意水準5%, 3, 5回は有意水準10%)は、トリートメントSの被験者の組の方がトリートメントPの被験者の組よりもより多く支配戦略をとっている。さらには、ピボタル・メカニズムは支配戦略均衡で遂行するメカニズムであるのにもかかわらず、後半の回になると約9割の組がナッシュ均衡戦略を選んでいる。さらには「悪い」ナッシュ均衡は180組のうち1組である。この意味では、実験データを見る限り、ピボタル・メカニズムはナッシュ遂行に成功しているのである。さらなる研究が必要となろう。

4. 結語

本稿では、セキユア遂行の理論とその実験結果を検討し、少なくとも実験環境では、ピボタル・メカニズムやセカンド・ブライズ・オーク

図 8. 組で支配戦略をプレイした被験者の頻度の割合



ションなど非セキュアなメカニズムは機能しないものの、選好の単峰性を満たす環境におけるGroves-Clarkeメカニズムは機能することをみた。

本稿は、伝統的なメカニズム・デザインの枠組みを採用しているが、同じ枠組みでPattanaik-Suzumura(1996)およびGotoh-Suzumura-Yoshihara(2005)らは、メカニズム(ないしはゲーム形式)そのものの手続的公平性を問う研究を開始している。Saijo-Sjöström-Yamato(2006)は、彼らの提案するセキュア遂行において制度の帰結である財配分のみ注目しているが、メカニズムそのものの手続的公平性は視野に入れていない。あるべきメカニズムの要件の一つとして手続的公平性がセキュア遂行の可能性にどのような影響を与えるのかは残された課題である。

本稿の枠組みは、経済的環境のみに限定した議論であって、公共財の供給という政治的な課題であるのにもかかわらずそのようなプロセスを排除している。

現実の政治・経済システムにおいては、各消費者の公共的サービスに対する需要を「政府」の供給活動に媒介するのは非人格的

メカニズムではなく、選挙によって民意を具体化する代表者として議会に送り出された「政治家」である。かような間接民主制ないしは代表民主制のもとでは、公共財に対する虚偽選好表明の可能性と限界は直接民主制のもとにおける場合とは自ら異なるものにならない。ただし「政治家」は選挙権者の多数による支持を確保するという自己自身の目的に従って行動するのであるから、彼らが支持者に委託された公共財に対する選好を虚偽表明し彼らのフリー・ライディングを助けるよう行動することに自らの利益を見出すか否かは決して自明のことではないからである(鈴木(1982), pp. 228-229)。

鈴木村の重要な指摘から四半世紀たつのだが、政治・経済システムにおける制度設計の理論は未だそのスタンダードに相当する理論すらない状況である。

一方で、経済学における人々の評価関数である効用関数そのものに疑念が出始めている。理論家が想定するような評価行動をとっていないことを多くの実験研究者、ニューロエコノミストたちが報告し始めている。さらには、生物としてのヒトの行動の枠組みから経済活動も逃れ

ることはできないという意味で人間行動の進化論的な起源を探る研究も重要である。

以上のような様々な課題が山積しているが、これらは今後の我々の課題としたい。

(*大阪大学社会経済研究所・大阪大学サステイナビリティサイエンス研究機構・東京工業品取引所市場構造研究所・CASSEL at UCLA/**東京工業大学大学院社会理工学研究科)

注

† 本研究の一部は、文部科学省科学研究費補助金(課題番号 16203012, 15310023)の補助を受けた。

1) 戦略的操作不能性は経済的な環境では動機整合性(incentive compatibility)と呼ばれることが多い。戦略的操作不能性およびメカニズム・デザインに関する優れたテキストとして、鈴木興太郎『経済計画論』筑摩書房、1982年を参照されたい。

2) 真の選好表明が支配戦略となることと真の選好表明がナッシュ戦略になることは同値である。Dasgupta-Hammond-Maskin (1979)を参照されたい。

3) 繰り返しの回数は、日本では8回、アメリカでは10回であった。以下の9回目、10回の目のデータはアメリカのデータのみである。

4) 詳細な利得表を被験者に見せるという手法は1990年代からよく使用されている。たとえば、Saijo-Nakamura(1995)など。

5) タイプ1の利得はタイプ2が12, 13のいずれをとっても210で変化しないが、タイプ2の利得は、タイプ1が16の場合196, 17の場合210である。

6) Cason-Saijo-Sjöström-Yamato(2006)の223-4ページを参照されたい。

7) 詳細な統計分析はCason Saijo-Sjöström-Yamato(2006)を参照されたい。

参考文献

- 鈴木興太郎(1982)『経済計画理論』筑摩書房。
- Attieyeh, G., Franciosi, R. and Isaac, R. M. (2000) "Experiments with the Pivotal Process for Providing Public Goods," *Public Choice*, Vol. 102, No. 1-2, pp. 95-114.
- Cason, T., Saijo, T., Sjöström, T. and Yamato, T. (2006) "Secure Implementation Experiments: Do Strategy-Proof Mechanisms Really Work?" *Games and Economic Behavior*, Vol. 57, No. 2, pp. 206-235.
- Clarke, E. H. (1971) "Multipart Pricing of Public Goods," *Public Choice*, Vol. 2, No. 1, pp. 19-33.
- Dasgupta, P., Hammond, P. and Maskin, E. (1979) "The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility," *Review of Economic Studies*, Vol. 46, No. 143, pp. 185-216.
- Gibbard, A. (1973) "Manipulation of Voting Schemes," *Econometrica*, Vol. 41, No. 4, pp. 587-601.
- Gotoh, R., Suzumura, K. and Yoshihara, N. (2005) "Extended Social Ordering Functions for Rationalizing Fair Game Forms in the sense of Rawls and Sen," *International Journal of Economic Theory*, Vol. 1, No. 1, pp. 21-41.
- Green, J. R. and Laffont, J.-J. (1979) *Incentives in Public Decision Making*, Amsterdam: North-Holland.
- Groves, T. (1973) "Incentives in Teams," *Econometrica*, Vol. 41, No. 4, pp. 617-631.
- Harstad, R. M. (2000) "Dominant Strategy Adoption and Bidders' Experience with Pricing Rules," *Experimental Economics*, Vol. 3, No. 3, pp. 261-280.
- Holmstrom, B. (1979) "Groves Scheme on Restricted Domain," *Econometrica*, Vol. 47, No. 5, pp. 1137-1144.
- Kagel, J. H., Harstad, R. M. and Levin, D. (1987) "Information Impact and Allocation Rules in Auctions with Affiliated Private Values: A Laboratory Study," *Econometrica*, Vol. 55, No. 6, pp. 1275-1304.
- Kagel, J. H. and Levin, D. (1993) "Independent Private Value Auctions: Bidder Behavior in First-Second-and Third-Price Auctions with Varying Number of Bidders," *Economic Journal*, Vol. 103, No. 419, pp. 868-879.
- Kawagoe, T. and Mori, T. (2001) "Can the Pivotal Mechanism Induce Truth-Telling? An Experimental Study," *Public Choice*, Vol. 108, No. 3-4, pp. 331-354.
- Pattanaik, K. and Suzumura, K. (1996) "Individual Rights and Social Evaluation: A Conceptual Framework," *Oxford Economic Papers*, Vol. 48, No. 2, pp. 194-212.
- Saijo, T., Sjöström, T. and Yamato, T. (2003) "Secure Implementation: Strategy-Proof Mechanisms Reconsidered," Working Paper 4-03-1, Department of Economics, the Pennsylvania State University.
- Saijo, T. and Nakamura, H. (1995) "The 'Spite' Dilemma in Voluntary Contribution Mechanism Experiments," *Journal of Conflict Resolution*, Vol. 39, No. 3, pp. 535-560.
- Saijo, T., Sjöström, T. and Yamato, T. (2006) "Secure Implementation," mimeo.
- Vickrey, W. (1961) "Counterspeculation, Auctions, and Competitive Sealed Tenders," *Journal of Finance*, Vol. 16, No. 1, pp. 8-37.