

96-J-1

メカニズム・デザインのゲーム理論
—均衡の一意性と再交渉の可能性を巡って—

松島 齊
東京大学経済学部

1996年1月

このディスカッション・ペーパーは、内部での討論に資するための未定稿の段階にある論文草稿である。著者の承諾なしに引用・複写することは差し控えられる。

メカニズム・デザインのゲーム理論¹ —均衡の一意性と再交渉の可能性を巡って—

松島齊（東京大学経済学部）

1996年1月8日

1. はじめに：二人の労働者を伴うエイジェンシー問題の例

ある経営者が、二人の労働者を雇って事業を達成させようとしている。事業に成功すれば、労働者は二人あわせて総利得300を経営者から要求することができる。しかし、事業に失敗した場合は、彼らはなにも獲得できない。事業に成功した状態は、さらに二つの状態、成功1、成功2、に場合分けされる。成功したか失敗したかについては、公に観察可能な公共情報である。しかし、成功1であるか成功2であるかは、両労働者にしか観察できない内部情報であり、部外者はおろか経営者にも識別できない。

事業が成功する確率は、ひとえに両労働者の努力水準に依存して決まる。労働者1の努力水準が高ければ成功1の確率が高まり、労働者2の努力水準が高ければ成功2の確率が高まる。そして、両努力水準の向上は共に、失敗確率を引き下げる。また、努力水準自体は、当人にしか識別できない私的情報であると仮定する。

経営者が直面する困難は、いかに両労働者から高い努力水準を選ぶインセンティブを引き出すか、にある。努力水準そのものが当該労働者の私的情報である以上、経営者が直接命令して強制させることはできない。比較的可能性のあるやり方としては、経営者が、何らかのコストをかけて、あらかじめ成功1か成功2かを識別できるようにしておくことが考えられる。そして、成功1の場合は労働者1に、成功2の場合は労働者2に、各々有利に、総利得300の分

¹本編は、「経済学論集」（東京大学経済学部）に掲載予定の、執筆を依頼された原稿である。

配の仕方を約束しておくのである。そうすれば、両労働者は、努力水準を高めるコストを払って成功確率を引き上げたことに対して、経営者から十分な見返りをうけることを保証されたと解釈するであろう。以上の理由から、本論文では、分配の公正をも考慮した上で、成功1の場合は利得分配ベクトル(180、120)を、成功2の場合は(120、180)を実行することを、経営者の達成目標としたい。

しかしながら、成功1か成功2かを直接識別するコストが高い場合は、上述のようなやり方では満足いく解決案にならない。そこで、本論文では、代替案として、両労働者に、成功1と成功2のどちらが真の状態であるかについて、何らかの申告をすることを強制させて、その申告内容に基づいて、あらかじめ約束しておいたメカニズムのルールに則って、利得分配を決定する方法を模索することにしよう。しかしこの際、両労働者は、嘘の申告をして、自分に有利な分配を誘導するかもしれない。したがって、両労働者が自発的に正直な申告をするインセンティブを持つように、巧妙にメカニズム設計を工夫する必要がある。本論文の目的は、はたして、成功1、成功2双方の状態において常に、正直な申告が、唯一の、均衡戦略となるようにメカニズムを設計できるか、あるいは、どのようなメカニズムによってそれが可能となるか、を検討することにある。

本論文は、一般に、メカニズム・デザインのゲーム理論、とりわけ(完備情報下の)誘導可能性問題(implementation problem)、と呼ばれるアプローチのエッセンスを理解するのに役立つであろう。詳しい当該分野の研究成果の歴史的展開については、松島(1995)に紹介済みなので省略する²。

本論文において、今まで誘導可能性問題の文脈においては十分に検討されなかった新たな論点を紹介したい。当該分野のほとんどの研究では、メカニズムを通じて導かれた最終決定は、事後的に変更できない、強い拘束力を持つコミットメントとなることが仮定されてきた。第二節において、この仮定を踏襲した上で、メカニズム・デザインの可能性を説明する。そこでは、Abreu and Matsushima(1992, 1994)による設計案が、より直観的な理解に基づくように修正した上で、簡単に紹介される。

本論文の中心部は、第三節の議論にある。ここでは、メカニズムによる決定を、事後的に、労働者間の同意によって変更することができると仮定される。すなわち、メカニズムによるコミットメントが、労働者間の再交渉(renegotiation)決裂時を除いては拘束力を持たないケースを考察するので

²その他の展望論文は、Maskin(1985), Moore(1992), Palfrey(1992), Osborne and Rubinstein(1994, chapter 10), 等多数ある。また、松島(1996a, 1996b)も参照されたい。

ある。両労働者は再交渉を通じて、成功時には常に、経営者から総利得300の支給を請求する事に合意し、メカニズムが提示した分配分を越える額については、両労働者の相対的な交渉力に応じて、その分配が決定されると考えるのである。我々は、交渉力のあり方が、労働者の資産価値や社会的評価の相違など、社会状態を構成する諸要因に依拠しうることを、明示的に考慮したい。そして、再交渉の可能性が、経営者の達成目標の誘導可能性を著しく促進する事を説明する³。

再交渉の可能性の有無は、契約の経済理論と呼ばれる、誘導可能性問題とは異なる文脈においては、常に論議的とされてきた⁴。そこでは、再交渉の可能性を認めてしまうと、非効率な利得分配を事後的には強制できなくなるため、労働者の逸脱行為に対するペナルティーの実行能力が狭められると論じられてきた。そして、再交渉の可能性を、分権化と相互不可侵性を強化したり、あるいは全権集中化をはかるなどして、できるだけ制約する方がより望ましいと論じられたのである。本論文では、それとは対照的に、再交渉の可能性が一意的均衡による誘導可能性をむしろ促進すると論じる事によって、その積極的な役割のひとつを示したいと思う。

なお、本論文は、近作 Matsushima (1995) の主要部分を、例をつかって単純化して、解説したものである。

³深く関連する論文として、Maskin and Moore (1986)がある。

⁴Grossman and Hart (1986), Hart and Moore (1988), Segal (1995), Hart (1995), 等など。

2. 再交渉の可能性のない環境の考察

2-2. 例による説明

本節では、経営者が労働者に、申告内容の如何によっては、総利得300の請求をしないようにコミットさせることができることを仮定する。まず、次のような簡単なメカニズムから議論を出発させよう。成功1か成功2かが判明した時点で、各労働者は、独立に同時に、“成功1”か“成功2”かをメッセージとして経営者に申告する。もし二人ともに“成功1”を申告した場合は、それを真の状態とみなして、望ましい利得分配ベクトル(180, 120)を実行する。逆に、もし二人ともに“成功2”を申告した場合は、それを真の状態とみなして、(120, 180)を実行する。問題は二人が異なる申告をした場合で、この際は総利得300の支給をせず、利得ベクトル(0, 0)を実行することにコミットしておく。こうして、成功1、成功2のどちらの状態においても、両労働者は、[図1]に示される同一のゲームに直面することとなる。

	成功1	成功2
成功1	180, 120	0, 0
成功2	0, 0	120, 180

[図1]

このメカニズムでは、成功1、成功2のどちらが真の状態であっても常に、申告(戦略)ベクトル(成功1, 成功1)と(成功2, 成功2)が共に Nash 均衡となり、正直な申告のみならず、二人が同じ嘘をつく申告ベクトルをも正当化される。このような Nash 均衡複数存在のために、上述のような簡単なメカニズムでは、経営者の計画する分配ルールの実現が十分に保証されたとはいえないのである。

次に、経営者が、本来の達成目標である利得分配に付随させて、別の経営的決定をも、同じ申告ベクトルに依存させることによって、インセンティブの引出しを試みるやり方を考えよう。今、経営者の判断だけで、aかbかの選択と

いう経営的決定ができるものとする。この決定は、労働者1の効用に少なからず影響を与えるものとする。労働者1は、成功1のケースではaを、成功2のケースではbを、他方よりも利得M分だけより選好しているものと仮定する。

まず、Mが十分に大きいケース ($M > 180$) を考察しよう。上述のメカニズムを修正して、経営者が、利得分配に加えて、労働者1が”成功1”を申告したならばa、”成功2”を申告したならばb、の経営的決定を下すことにコミットしたとしよう。すると、成功1が真の状態である時、労働者1は、”成功1”を正直に申告しておけば、労働者2の申告内容に関係なく、常に180を越える効用が期待できることになる。よって、労働者1にとって、”成功1”が優位戦略、”成功2”が劣位戦略となる。成功2が真の状態である時も同様に、逆に、”成功2”が労働者1にとって優位戦略、”成功1”は劣位戦略となる。一方、労働者2は、以上の理由から、労働者1が必ず正直に申告することを前提とせざるをえないので、正直な申告ベクトルが Nash 均衡であることから、結局自分も正直に申告することが常に最適反応となる。こうして、上述の修正案においては、正直な申告が、唯一の Nash 均衡、あるいはより正確に、唯一の反復非劣位戦略ベクトル (iteratively undominated strategies) 、となり、とりあえずの解決策を得ることとなる。

上の議論は、Mが十分に大きいことを必要不可欠としている。これは、非常に制約的な条件である。経営者は、当初の問題に比べて相対的に大きな規模の経営的決定を付随させることになり、aかbかの決定に伴う経営者利益の変動額が当初の問題解決による利益増を上回ることも考えられるからである。我々は、小規模の経営的決定に範囲をしぼって、一意的誘導可能性を再度検討すべきである。

以下において、Mが非常に小さい額であることを仮定しよう。単純化のため、 $M=1$ として議論をすすめる。今までに紹介したメカニズムでは、経営者の目標を達成できない。そこで、次のような新たな設計案を紹介したい。経営者は、両労働者が”成功1”か”成功2”かに関する申告を終えた後、K分間を経てから、利得分配の決定、および付随された経営的決定をおこなうものとする。そして、両労働者は、決定までのK分間に、もし申告の変更をしたければ、1分間隔で変更が認められるとしよう。変更がなければ、最初の申告内容に従って、前述のメカニズムに照らして、利得分配と経営的決定がなされるものとする。

もし労働者のどちらかが途中で変更を申し出たならば、まず、経営的決定に関しては、労働者1のK分目の申告内容に従ってなされる。利得分配の決定は、全部で4通りある申告ベクトル、(成功1、成功1)、(成功1、成功2)、(成功2、成功1)、(成功2、成功2)、の各々について、K分間のうち何

分間その申告が維持されたかに比例して、対応する利得分配ベクトルにウエイト付けして査定されるものとする。また、両労働者のうち最後に変更した者は、相手労働者に対して、 $200/K$ のペナルティーを支払う。

さらに、労働者2は、 K 分目の申告内容が労働者1の K 分目の申告内容と食い違った場合には必ず、 $100/K$ のペナルティーを労働者1に支払う。以上が、メカニズム設計の定義である。

例えば、 $K=300$ として、労働者1は、最初の150分間は”成功1”、次の150分間は”成功2”を申告し、一方労働者2は、最初の100分間は”成功1”、次の100分間は”成功2”、最後の100分間は”成功1”を申告したとしよう。まず、労働者1の K 分目の申告内容が”成功2”なので、経営的決定 b が選択される。肝心の利得分配の決定に関しては、まず、申告ベクトル(成功1、成功1)は100分間、(成功1、成功2)は50分間、(成功2、成功2)は50分間、(成功2、成功1)は100分間アナウンスされたので、基礎となる利得分配ベクトルは

$$\begin{aligned} & \frac{100}{300}(180,120) + \frac{50}{300}(0,0) + \frac{50}{300}(120,180) + \frac{100}{300}(0,0) \\ & = (80,70) \end{aligned}$$

である。労働者2は、二人のうちで最後に変更したので、労働者1にペナルティー $200/K=2/3$ を支払う。労働者2は、 K 分目の申告内容が労働者1のそれと食い違っているので、労働者1にペナルティー $100/K=1/3$ を支払う。以上より、トータルして、利得分配ベクトル

$$(80,70) + \left(\frac{2}{3}, -\frac{2}{3}\right) + \left(\frac{1}{3}, -\frac{1}{3}\right)$$

が実行されることになる。

以下において、 K が十分に大きい場合、つまり、

$$K \geq 400$$

の時、”成功1”、”成功2”のどちらが真の状態であろうとも、両労働者にとって最初から正直に申告して一度も変更しないことが、唯一のNash均衡、より正確には唯一の反復非劣位戦略均衡、であることを示そう。まず、決定直前の K 分目では、労働者1が必ず正直な申告をすることが容易に示される。労働者1は、正直な申告をする事によって、経営的決定を自分に有利に導いて、

効用 $M=1$ の増加が見込まれる。その反面、労働者1が $K-1$ 分目に虚偽申告していた場合は、最終変更者に対するペナルティー $200/K$ が労働者1に課せられることとなる。さらに、相手労働者2が K 分目に虚偽申告した場合は、基礎となる利得分配ベクトルが $(0, 0)$ となり、労働者1は、自分も虚偽申告した時と比べて、“成功1”が真の状態である時には $180/K$ 、“成功2”が真の時には $120/K$ 、の利得減をこうむる。しかしながら、不等式

$$\frac{200}{K} + \frac{180}{K} < \frac{380}{400} < 1$$

が成立するから、正直な申告による便益が、常に損失を上回るのである。以上より、労働者1は、 K 分目に嘘をつくことを、劣位戦略として避けるのである。

次に、労働者2にとっても、 K 分目に正直な申告をすることが最適であることを示そう。労働者1が K 分目に正直な申告をすることが確実であるため、同様に正直に申告することによって、“成功1”が真の時には基礎となる利得分配 $120/K$ 、“成功2”が真の時には $180/K$ の増加が見込まれる。それのみならず、労働者1の申告との食い違いに対するペナルティー $100/K$ をも回避できる。その反面、正直に申告することで最後の変更者になる可能性が排除できないため、ペナルティー $200/K$ をこうむることも考慮しておく必要がある。しかし、不等式

$$\frac{120}{K} + \frac{100}{K} > \frac{200}{K}$$

が成立するため、正直な申告による便益が常に損失を上回るのである。以上より、労働者1が K 分目に正直な申告をすることを前提とすると、労働者2にとっても、 K 分目に嘘をつくことが劣位戦略になるのである。

K 分目では必ず、両労働者が正直に申告することがわかった。次に、 $K-1$ 分目においても、二人が正直に申告することを示そう。まず、 $K-1$ 分目で虚偽申告する者は必ず最終変更者となりペナルティー $200/K$ を課されることに気付かれよ。また、相手労働者が $K-1$ 分目に正直な申告をする場合には、自分も正直な申告をすれば総利得 300 の分配の恩恵にあずかれる一方、相手が虚偽申告する場合は、正直な申告によってむしろ上限 $180/K$ の損失を受けることになる。しかしその損失額も、最終変更者に対するペナルティー $200/K$ を常に下回るので、両労働者共に、 $K-1$ 分目にて虚偽申告をすることが劣位戦略となるのである。

$K-1$ 分目における上述の議論と全く同様にして、我々は、任意の時点にお

いても、その時点以後二人が共に正直な申告をすることを前提とすると、その時点で虚偽申告することが必ず劣位戦略になることを示す事ができるのである。

以上より、結局、すべての時点において正直に申告する事、すなわち、初めから正直な申告をして決して変更しない事が、真の状態が”成功1”であるか”成功2”であるかにかかわらず常に、唯一の反復非劣位戦略となるに至る⁵。達成目標とした利得分配の仕方は、上述のメカニズムによって、高々若干の経営的決定の犠牲だけで実行可能となる。経営者は、両労働者から高い努力水準を選択するインセンティブを引き出すことができるのである。

2-2. 一般化

以上の結果は、一般的な経済環境に拡張することができる。ある経済計画者が選択肢の集合 A の中からの選択に直面している。最適な選択は社会状態 ω に依存し、選択関数 $f: \Omega \rightarrow A$ にしたがうことを理想とする。ここで、 Ω は実現可能な社会状態 ω 全体の集合で、 A は加算無限個の点からなる集合上で定義されるすべてのくじ（確率的選択）を含むものとする。計画者は真の社会状態を直接知ることができないとする。一方、社会には n 人の経済主体が存在して、彼らは真の状態を知る事ができる。計画者は、各経済主体にメッセージを申告させ、あらかじめ設計したメカニズムに則って選択決定をする。メカニズムは、 $G = (M_1, \dots, M_n; g)$ と定義される。ここで、 M_i は経済主体 i のメッセージ（申告、あるいは戦略）全体の集合で、 $g: M \rightarrow A$ にしたがって実際の決定がなされる。なお、 $M = M_1 \times \dots \times M_n$ とする。社会状態 ω の下で、各経済主体 i の選択肢の集合 A 上で定義された効用関数は $u_i(\cdot, \omega)$ で表され、各経済主体はゲーム $(G, u_1(\cdot, \omega), \dots, u_n(\cdot, \omega))$ に直面することとなる。全経済主体は、ある戦略ベクトル $m^*(\omega) = (m_1^*(\omega), \dots, m_n^*(\omega)) \in M$ を申告することになるが、この時、まず、

$$g(m^*(\omega)) = f(\omega)$$

⁵K分間の変更可能期間中に、相手の申告内容はなにか、あるいは相手を変更したか否か、が観察できるものと仮定した場合は、劣位戦略の反復削除は、より正確には、すべての部分ゲームに対して貫徹させる必要がある。観察できないと仮定した場合は、通常の反復非劣位戦略の定義にしたがう。Abreu and Matsushima (1992a, 1992b), Glazar and Perry (1992) などを参照せよ。

が、すべての社会状態 $\omega \in \Omega$ について、成立する事が要求される。問題は、戦略ベクトル $m^*(\omega)$ を経済主体が実際に申告するインセンティブをもつか、である。そこで、 $m^*(\omega)$ はゲーム $(G, u_1(\cdot, \omega), \dots, u_n(\cdot, \omega))$ における唯一の反復非劣位戦略ベクトルであることが要求されるのである。以上の条件をみたすメカニズムが存在するならば、選択関数 f は誘導可能 (implementable) である、と定義される。

以下の条件は、選択関数 f の誘導可能性のために必要不可欠なものである。

条件 I : 任意の二つの社会状態 $\omega, \omega' \in \Omega$ について、もし等式

$$u_i(\cdot, \omega) = u_i(\cdot, \omega')$$

が、すべての経済主体 i について成立するときは、常に

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

が成立する⁶。

条件 I は、社会状態の相違は、それがどの経済主体の選好関係の相違にも反映されない場合には、社会的選択においても区別されないことを意味する。すなわち、選択関数は全経済主体の選好関係の集計ルールとみなされる、という条件である。

条件 I は、経済主体が三人以上存在する場合、あるいは二人しか存在しなくても正直な申告をすることが Nash 均衡のひとつになる場合には、必要条件であるのみならず充分条件でもあることが、Abreu and Matsushima (1992a) によって証明されている。すなわち、条件 I をみたすすべての選択関数は、唯一の反復非劣位戦略ベクトルによって (近似的に) 誘導可能である、というわけである。その証明のエッセンスは、第 2-1 節の例での議論に集約される。

反復非劣位戦略ベクトルの一意性は、Nash 均衡の一意性以上に、好都合なくいくつかの性質をみたしてくれる。まず、Nash 均衡のみならず、合理化 (rationalizable) 戦略ベクトルの一意性をも意味する⁷。そして、全プレー

⁶この等式は、正確には、なんらかのアフィン変換を施した後に一致することを意味している。

⁷Bernheim (1984), Pearce (1984) を参照のこと。

ヤーが合理的であることが周知の事実 (common knowledge) であると仮定すれば、その戦略ベクトルの実行を数学的帰納法だけから特定化することができる。

もっとも、現実の経済主体は、数学的帰納法に基づいて行動決定をするとは限らない。むしろ、学習や試行錯誤を繰り返して、満足水準の達成に近づこうとするだろう。重要な点は、反復非劣位戦略ベクトルの一意性が純粋戦略の反復削除のみによって導かれる場合、非常に一般的な範囲のどの学習や進化の調整過程に従おうとも、その反復非劣位戦略ベクトルの達成が長期的には約束されるという事である。この性質は、単なる Nash 均衡の一意性や、混合戦略の削除をともなう反復非劣位戦略ベクトルの一意性の場合に、調整過程の特定化如何によっては、リミット・サイクルなどが発生することとは、すこぶる対照的である⁸。

我々のメカニズムでは、純粋戦略の反復削除のみによって、誘導可能性が導かれている。したがって、誘導可能性をうながす均衡行動パターンは、単に合理的推論の産物としてではなく、広範囲の学習や進化のプロセスを通じて、社会の慣習として長期的に根付くことのできる、唯一の行動パターンであると解釈することができる。(しかしながら、このような「限定された」合理性に基づく解釈には、問題がないわけではない。第4節を参照のこと。)

⁸Milgrom and Roberts (1991), Dekel and Scotchmer (1992), Kandori (1995), などを参照のこと。

3. 再交渉の可能性のある環境の考察

3-1. 例による説明

前節での解決案は、条件 I に示されるように、労働者 1 が、経営的決定に対して、真の状態が”成功 1”であるか”成功 2”であるかに応じて、異なる選好関係をもつという前提条件に支えられている。本節では、条件 I が成立しないため、前節までの議論では誘導可能性を説明できない状況を、主に考察する。

前節では、もうひとつの重要な条件として、メカニズムを設計する段階において、あらかじめ、両労働者が事後的には総利得支給 300 の要求をしないことにコミットさせることができることが仮定されていた。本節では、この前提条件をも取り除くこととし、二人の労働者は協力して、事後的には常に総利得支給 300 を経営者に請求できるものとする。そしてその際には、労働者間の相対的な交渉力に応じて超過利得の分配がなされるものとする。

まず、第 2-1 節の最初に紹介した最も単純なメカニズムを再検討しよう。もし二人が異なる申告をした場合は、メカニズムは (0, 0) を提示するが、二人は権限を行使して、経営者から総利得 300 を取り立てる。労働者 1 は、労働者 2 との交渉を通じて、”成功 1”が真の状態である時は超過支給のうちの割合 h を、”成功 2”の時は割合 $1-h$ を獲得できるとする。したがって、両労働者の直面するゲームは、”成功 1”の時は [図 2-1]、”成功 2”の時は [図 2-2]、に示される 2 プレーヤー・ゼロ和ゲームになる。

	成功 1	成功 2
成功 1	180, 120	300 h , 300(1- h)
成功 2	300 h , 300(1- h)	120, 180

[図 2-1]

成功 1	180, 120	300(1- h), 300 h
成功 2	300(1- h), 300 h	120, 180

[図 2-2]

もし $h \geq \frac{3}{5}$ ならば、ミニ・マックス定理より、“成功1”、“成功2”の如何にかかわらず常に、正直な申告が唯一の Nash 均衡になる。つまり、労働者1の努力が実る“成功1”の状態において労働者1の交渉力が十分に強く、労働者2の努力が実る“成功2”の状態において労働者2の交渉力が十分に強い場合は、利得請求の権限を認める事がむしろ問題解決を促進させる働きをもつのである。

しかし、逆に $h < \frac{3}{5}$ ならば、両労働者が共に虚偽申告することが唯一の Nash 均衡となってしまう、新たに別のメカニズム設計案を模索する必要性がでてくるのである。そこで、以下では追加的に、内部情報を知っている第三者（スーパーヴァイザー、顧問弁護士）が存在すると仮定して、経営者は、その第三者にも“成功1”か“成功2”かを申告させる義務を課すことができるものとしよう。

まず、次のような比較的単純なメカニズムから出発しよう。二人の労働者と第三者の計三人が同時に申告をする。この時、必ず二人ないしは三人は同じ申告をすることになるので、その申告内容を真の状態とみなして、それに対応する利得分配を実行することにする。このメカニズムでは明らかに、三人が共に正直に申告する事が、常に Nash 均衡のひとつになる。なぜならば、他の二人が正直な申告をすることを前提とすると、残りの一人は、正直な申告をするか否かは、このメカニズムの下では、分配決定にもはや無影響なので、虚偽申告をする積極的な理由がないからである。しかしその反面、三人共に虚偽申告することをも、同じ理由で、別の Nash 均衡になってしまう。よって、このようなメカニズムでは問題の解決にならない。

以下において、第2-1節後半に紹介したメカニズム・デザインに類似した方法によって、ひとつの解決案を示したいと思う。単純化のため、 $h = 0$ のケースのみを考察する。不等式 $h < \frac{3}{5}$ をみたく一般的なケースについても、全く同様の議論が成立する。

経営者は、両労働者および第三者の三人が“成功1”か“成功2”に関する申告を終えた後、 K 分間を経てから、利得分配の決定をおこなうものとする。そして、三人は、決定までの K 分間に、もし申告の変更をしたければ、1分間隔で変更が認められるとする。

例えば、三人のうち二人以上が“成功1”を申告する状況がトータルで K' 分間、その一方で、やはり二人以上が“成功2”を申告する状況がトータルで $K - K'$ 分間あったとしよう。まず、基本給として、両労働者に利得分配ベク

トル

$$\frac{K'}{K}(170,110) + \frac{K-K'}{K}(100,180)$$

が支給される。基本給の総額は請求可能額300より20低い280であり、前者の状況では両労働者から10ずつ、後者では労働者1のみから20が、差し引かれていることに注意せよ。また、第三者に対する基本給はゼロとする。

もし労働者1が最後のK分目に”成功1”を申告した場合は必ず、両労働者に利得10が追加分として支給される。一方、労働者1が”成功2”を申告した場合には、追加支給されない。

労働者2は、K分目に労働者1のK分目の申告と異なる申告をした場合は、労働者1に利得2を支払う。同様に、第三者も、K分目に労働者1のそれと異なる申告をした場合、労働者1に利得2を支払う。

さらに、三人のうち最後に変更した者が、もし労働者1である場合は労働者1から労働者2へ、労働者2である場合は労働者2から労働者1へ、第三者である場合は第三者から経営者へ、各々利得1のペナルティーを支払う。以上がメカニズム設計の仕方である。両労働者には、再交渉の可能性が常に認められているものとする。

以下において、Kが十分に大きい場合、つまり、

$$K \geq 100$$

の時、”成功1”、”成功2”のどちらのケースにおいても、三人にとって、最初から正直に申告して一度も変更しないことが、唯一の反復非劣位戦略であることを示そう。

まず、K分目において労働者1は必ず正直な申告をすることが、容易に示される。労働者1が”成功1”を申告することによって、追加利得10が支給される一方、”成功2”を申告した場合は、労働者2と協力して、総利得20の追加請求をすることとなる。後者の場合、もし”成功1”が真の状態ならば、 $h=0$ であるため、労働者2との交渉を通じて、労働者1はなにも獲得することができない。もし”成功2”が真である時は、逆に、20すべてを獲得することができる。したがって、正直な申告をすることによって常に、追加的な支給増を通じて、労働者1の利得は10増加することになる。その反面、労働者1がK-1分目に虚偽申告していた場合は、最終変更者に対するペナルティー2が労働者1に課せられることとなる。さらに、労働者2と第三者のどちらか

一方のみがK分目に虚偽申告した場合は、労働者1は、正直に申告することによって、”成功1”が真である場合は $170/K - 100/K = 70/K$ の基本給の増額、”成功2”が真である場合は $70/K$ の減額をこうむる。しかし、不等式

$$2 + \frac{70}{K} < 2 + \frac{70}{100} < 10$$

が成立するため、正直な申告による便益が常に損失を上回るのである。以上より、労働者1は、K分目に嘘をつくことを、劣位戦略として避けるのである。

次に、労働者2にとっても、K分目に正直な申告をすることが最適であることを示そう。労働者1がK分目に正直な申告をすることが確実であるため、同様に正直に申告することによって、ペナルティー2を免れる事ができる。もっとも、正直に申告することで最後の変更者になる可能性が排除できないため、ペナルティー1をこうむることも考慮しておく必要がある。さらには、”成功1”が真である場合に、もし第三者が虚偽申告をするならば、労働者2も便乗して虚偽申告をすることによって、基本給を $70/K$ 増額させることができる。しかし、不等式

$$1 + \frac{70}{K} < 1 + \frac{70}{100} < 2$$

が成立するため、正直な申告による便益が、常に損失を下回るのである。以上より、労働者1がK分目に正直な申告をすることを前提とすると、労働者2にとっても、K分目に嘘をつくことが劣位戦略になるのである。全く同様にして、第三者もまた、K分目に正直な申告をすることが唯一の最適戦略となる。

K-1分目においても、三人は正直に申告する。まず、K-1分目で虚偽申告する者は必ず最終変更者としてペナルティー1を課されることに気付かれよ。その一方で、労働者1ないし2は、虚偽申告によって、基本給の増加 $70/K$ が見込まれるかもしれない。しかし、不等式

$$\frac{70}{K} < \frac{70}{100} < 1$$

が成立するため、三人にとって正直な申告が唯一の非劣位戦略になる。

K-1分目における上述の議論と全く同様にして、我々は、任意の時点においても、その時点以後三人が正直な申告をすることを前提とすると、その時点

で虚偽申告することが必ず劣位戦略になることを示す事ができるのである。

以上より、結局、すべての時点において正直に申告する事、そなわち、初めから正直な申告をして決して変更しない事が、常に唯一の反復非劣位戦略となるに至る。経営者の達成目標は、上述のメカニズムによって、再交渉の可能性を利用することによってはじめて、一義的均衡配分として誘導可能となるのである。

3-2. 一般化

以上の結果は、一般的な経済環境に拡張することができる。第2-2節にて定義された基本モデルを、以下のように修正する。メカニズム $G=(M_1, \dots, M_n; g)$ の下で、全経済主体が $m=(m_1, \dots, m_n) \in M$ を申告した際、まずとりあえずは選択肢 $g(m) \in A$ が提示される。しかし、全経済主体は、再交渉を通じて、 $g(m)$ とは別の選択肢に合意するかもしれない。ここで我々は、再交渉関数 $h: A \times \Omega \rightarrow A$ の存在をあらかじめ仮定し、彼らが $h(g(m), \omega) \in A$ に合意するものとする。全経済主体は、ある戦略ベクトル $m^*(\omega) = (m_1^*(\omega), \dots, m_n^*(\omega)) \in M$ を申告するが、この時、まず

$$h(g(m^*(\omega)), \omega) = f(\omega)$$

が、すべての社会状態 $\omega \in \Omega$ について、成立する事が要求される。そして、 $m^*(\omega)$ は、再交渉関数 h を伴うゲーム $(G, u_1(\cdot, \omega), \dots, u_n(\cdot, \omega))$ における、唯一の反復非劣位戦略ベクトルであることが要求されるのである。以上の条件をみたすメカニズムが存在するならば、選択関数 f は、再交渉関数 h の下で、誘導可能である、と定義される。

以下の二つの条件は、選択関数 f の誘導可能性のために必要不可欠なものである。

条件II： すべての社会状態 $\omega \in \Omega$ に対して、ある選択肢 $a \in A$ が存在して、

$$h(a, \omega) = f(\omega)$$

が成立する。

条件III: 任意の二つの社会状態 $\omega, \omega' \in \Omega$ について、もし等式

$$u_i(h(a, \omega), \omega) = u_i(h(a, \omega'), \omega')$$

が、すべての経済主体 i およびすべての $a \in A$ について成立するときは、常に

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

が成立する。

条件IIは、選択関数の指し示す選択肢は、再交渉の可能性を考慮した上で、交渉決裂時の選択を適切なものにコミットすることによって、到達可能であることを意味する。より本質的な条件である条件IIIは、社会状態の相違は、経済主体の選好関係の相違に反映されるか、あるいは相対的な交渉力の相違に反映されるかの、どちらかが成り立たなければ、社会的選択において区別されない事を意味する。

近年の Matsushima (1995) では、経済主体が三人以上存在する場合において、条件IIとIIIは、若干の追加的条件の下では、必要条件のみならず充分条件でもあることが証明されている。その証明のエッセンスは、上に示された例での議論に集約される。

4. まとめ

メカニズムを駆使すれば多くの選択関数を誘導可能にできるという本論の帰結は、Grossman and Hart (1986) に代表される契約の不完備性を前提とした経済分析と矛盾するものではない。もっとも、契約の不完備性は、単に内部情報の立証不可能性と再交渉の可能性だけから直接論理的に導かれるものではなく、明記しておく必要がある⁹。

また、Dewatripont and Maskin (1990) 等が述べるように、もし再交渉のプロセス自体がなんらかの単一のメカニズムで表現でき、その交渉行動が一意的な反復非劣位戦略ベクトルで記述されるならば、再交渉の可能性を認めることは、単に制約条件を一つ増やしたに過ぎなくなり、誘導可能性の余地を狭めこそすれども促進することはありません。我々は、実は、本論において暗黙に、再交渉プロセス自体が社会状態 ω に依存してきまると仮定しているのである。そしてさらに、我々は、そのプロセスの相違は外部に立証することができないことを、暗黙に仮定しているのである。そうすることによって、Dewatripont and Maskin 等の指摘による批判を回避することができる。

さて、反復非劣位戦略の一意的性の下では、次のような、提携活動の不可能性を意味する望ましい性質が成立する。すなわち、任意の社会状態 $\omega \in \Omega$ について、 $m^*(\omega) \in M$ を唯一の反復非劣位戦略ベクトルであるとした場合、すべての提携グループ $D \subset N$ に対して、そしてそのグループのすべての申告ベクトル $m_D = (m_j)_{j \in D} \neq (m_j^*(\omega))_{j \in D}$ に対して、ある当該グループのメンバー $i \in D$ およびそのメンバーの戦略 $m_i' \neq m_i$ が存在して、不等式

$$u_i(h(g(m_{N/D}^*(\omega), m_{D \setminus \{i\}}, m_i'), \omega), \omega) > u_i(h(g(m_{N/D}^*(\omega), m_D), \omega), \omega)$$

が成立する、ということである。つまり、どのような提携による、集団的な虚偽申告の逸脱行為の試みも、提携メンバーの、少なくとも一人によって、抜け駆けされる運命にあり、必ず失敗に終わるのである。

メカニズム・デザイン論は、計画者や当事者間の契約上の合意によって、物理的なコミットメントができることを前提としている。しかし、実際に人々が知覚するゲームの世界は、メカニズムが指し示す物理的条件だけから構成されるわけではない。例えば、経済主体が、実際には、メカニズムにおけるプレイ

⁹契約の不完備性に合理的根拠を与える研究がいくつかある。Anderlini and Felli (1994), Segal (1995) 等など。

を含んだ無限回繰り返しゲームに直面しているものと主観的に認識し、理論の予想と全く異なる行動パターンに長期的に安定化してしまうことは、大いに考えられるのである¹⁰。限定合理的なプレーヤーが、どのようなパターン認識の下でプレーするかを、メカニズム・デザインの文脈でも検討することが課題である。

第2-2節後半において、反復非劣位戦略による誘導可能性を、学習や進化の調整過程の一意安定状態として解釈するという提案をした。この際、我々は、経済主体は限定合理的である一方で、計画者自身は、調整過程の推移を十分に考慮した上で、合理的計算によって最適なメカニズムの詳細を設計できる程に sophisticated な存在であることを、暗黙のうちに仮定せざるをえない。単に経済主体のみならず、計画者自身も限定合理的であることをいかに明示的に取り扱うかが、メカニズム・デザイン論が今後すぐれた政策的意義を勝ち得るために答えていかなければならない問題である。

¹⁰Rubinstein (1991)などを参照のこと。

参考文献

- Abreu, D. and H. Matsushima (1992): "Virtual Implementation in Iteratively Undominated Strategies: Complete Information," *Econometrica* 60, 993-1008.
- Abreu, D. and H. Matsushima (1994): "Exact Implementation," *Journal of Economic Theory* 64, 1-19.
- Anderlini, L. and L. Felli (1994): "Incomplete Written Contracts: Undescribable States of Nature," *Quarterly Journal of Economics*.
- Bernheim, D. (1984): "Rationalizable Strategic Behavior," *Econometrica* 52, 1007-1028.
- Dekel, E. and S. Scotchmer (1992): "On the Evolution of Optimizing Behaviour," *Journal of Economic Theory* 57, 392-406.
- Dewatripont and Maskin (1990): "Contract Renegotiation in Models of Asymmetric Information," *European Economic Review* 34, 311-321.
- Glazer, J. and M. Perry (1992): "Virtual Implementation in Backwards Induction: A Note on Abreu and Matsushima Mechanism," mimeo.
- Grossman, S. and O. Hart (1986): "The Costs and Benefits of Ownerships: A Theory of Vertical and Lateral Integration," *Journal of Political Economy* 94, 691-719.
- Hart, O. (1995): *Firms, Contracts, and Financial Structure*. Oxford University Press.
- Hart, O. and J. Moore (1988): "Incomplete Contracts and Renegotiation," *Econometrica* 56, 755-786.
- Kandori, M. (1995): "Evolutionary Game Theory in Economics," invited papers for the Seventh World Congress of the Econometric Society.
- Maskin, E. (1985): "The Theory of Implementation in Nash Equilibrium," in *Social Goal and Social Organization*, ed. by L. Hurwicz, D. Schmeidler, and H. Sonnenschein. Cambridge: Cambridge University Press.
- Maskin, E. and J. Moore (1986): "Implementation and Renegotiation," mimeo.
- Matsushima, H. (1995): "Incomplete Contract Design," mimeo.
- 松島齊(1995): "A-Mメカニズム・デザインの合理性"、東京大学経済学部ディスカッション・ペーパー・95-J-13。「経済研究」(一橋大学経済研究所、岩波書店)掲載予定。
- 松島齊(1996a): "メカニズム・デザイン:いくつかの墓碑銘と再生"、執筆中。
- 松島齊(1996b): "キング・ソロモンの知恵3000年の旅"、執筆中。

Milgrom, P. and J. Roberts (1991): "Adaptive and Sophisticated Learning in Normal Form Games," *Games and Economic Behaviour* 3, 82-100.

Moore, J. (1991): "Implementation, Contracts and Renegotiation in Environments with Complete Information," in *Advances in Economic Theory: Proceedings of the Sixth World Congress of the Econometric Society*, ed. by J.-J. Lafont. Cambridge: 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: Cambridge University Press.

Pearce, D. (1984): "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica* 52, 1029-1050.

Rubinstein, A. (1991): "Comments on the Interpretation of Game Theory," *Econometrica* 59, 909-924.

Segal, I. (1995): "Complexity and Renegotiation: A Foundation for Incomplete Contracts," mimeo.