



観測費用を伴う繰り返し囚人のジレンマゲームにおける評価形成 : ランダム・マッチング・モデル

宮原, 泰之

(Citation)

国民経済雑誌, 188(3):97-108

(Issue Date)

2003-09

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(JaLCD0I)

<https://doi.org/10.24546/00055881>

(URL)

<https://hdl.handle.net/20.500.14094/00055881>



観測費用を伴う繰り返し囚人の ジレンマゲームにおける評判形成： ランダム・マッチング・モデル

宮 原 泰 之

長期的な関係において、相手が過去に選択した行動を観測するには費用を伴うとした場合に、協力が達成可能かどうかを明らかにする。本稿では、ランダム・マッチング・モデルにおいて評判が協力達成のために有効であることを示す。そして、従来の評判効果に関する研究とは異なり、コミットメント・タイプではないという評判を形成することによって協力を達成することができることを明らかにする。

キーワード 観測費用，評判効果，無限繰り返しゲーム，モニタリング

1 序

ほとんどの生産や取引などの経済活動は多数の経済主体が関わる長期的な関係であると考えることができる。企業では、複数の従業員による協同作業が行われており、日々の努力投入が必要とされている。また、市場では企業間によるさまざまな競争が毎日繰り返されている。多くの関係においてお互いが協力することによってパレート効率性を達成することができる。企業については従業員達が協力して多くの努力を投入すれば、企業全体の利潤は最大化されるであろうし、また、寡占市場で価格競争が行なわれているのであれば、協力して高い価格を維持することができれば、その産業全体の利潤は最大化されるであろう。長期的な関係においては自分が協力的であることを示すこと、つまり、よい評判を形成することは、その関係におけるパレート効率性を達成するために有効である。

本稿では、社会における長期的な関係を無限繰り返しゲームとして定式化し、評判の形成に関する分析を行なう。従来の評判形成に関する研究では、ゲームが行われる前にプレイヤーのタイプが決定される。通常のタイプとコミットメント・タイプのいずれかがランダムに決定される。多くの研究ではコミットメント・タイプは協力的な行動を選択し続けるものと想定されている。例えば、Kreps, Milgrom, Roberts, and Wilson (1982) では有限繰り返し囚人のジレンマ・ゲームを分析しているが、コミットメント・タイプは仕返し戦略 (tit-for-

tat strategy) のみをプレーするものと想定されている。仕返し戦略は有限繰り返し囚人のジレンマ・ゲームでは均衡とはならない。仕返し戦略は、ある意味、利他的または協力的な戦略であると考えることができる。このような利他的な戦略にコミットするプレーヤーが存在する可能性があるとき、コミットメント・タイプが存在しない状況では達成可能ではない結果が均衡として達成される可能性があることを評判形成に関する研究は示している。本稿では従来の研究とは異なり、利他的ではなく、利己的なコミットメント・タイプを導入する。

無限繰り返しゲームに関する評判形成の分析は完全情報 (perfect information) と不完全情報 (imperfect information) の状況についてなされてきた。Schmidt (1993) では完全情報下、Fudenberg and Levine (1992) では不完全情報下における評判の形成について分析している。これらの研究では、協力的な戦略にコミットするプレーヤーをコミットメント・タイプとして導入している。それぞれの論文では通常のタイプのプレーヤーがコミットメント・タイプの戦略を真似るような均衡が存在することを示している。つまり、通常のタイプがコミットメント・タイプであるかのように振る舞い、協力的であるという評判を形成することによって協力的な結果が導出されることを示したのである。本稿では従来の研究とは異なる情報構造の下で通常のタイプがコミットメント・タイプではないという評判を形成することによって協力的な結果が導出されることを示す。

本稿で想定する情報構造は相手の行動を観察すれば相手を選択した過去一期間の行動を完全に知ることはできるが、観測費用を伴うというものである。さらに、観測は私的に行われるものであり、相手が観測しているかどうかは知ることはできないと仮定する。よって、観察すれば相手の選択した行動を完全に知ることができるとしても、情報の非対称性が生じることになる。このような状況を分析したものとして、Ben-Porath and Kahneman (2002), Miyagawa, Miyahara, and Sekiguchi (2003), Miyahara (2002) がある。これらの論文ではコミットメント・タイプを導入せずに、協力的な結果が導出されるかどうかを分析している。¹ Miyahara (2002) は無限繰り返し囚人のジレンマ・ゲームを分析している。本稿と同様の情報構造を仮定し、均衡戦略のクラスを限定した場合、観測費用が十分小さいとしても協力的な結果を導出することはできないことを明らかにしている。² つまり、観測費用を伴う状況において協力的な結果を導出することは簡単ではないことを明らかにしたのである。

以下では、ランダム・マッチングにおける無限繰り返し囚人のジレンマ・ゲームをコミットメント・タイプと観測費用を導入して分析する。基本モデルは Kandori (1992) と類似しているが、以下の二つの点で異なる。まず、すでに述べたようにコミットメント・タイプが導入されていることと観測費用が導入されている点で異なる。そして、本稿の設定では每期ランダム・マッチングが行われるわけではないことが大きな違いである。最初の期にマッチングが行われ二人からなるチームが構成され、プレーヤーのタイプが決定される。タイプは私

的情報となる。そして、囚人のジレンマ・ゲームがプレーされる。次の期は、両者が同意した場合にのみチームが解散され、再びランダム・マッチングが行われるのである。ひとりでも解散に反対した場合にはそのチームは継続されることになる。

本稿ではこの設定の下で、非協力的な行動を選択し続けるような戦略にコミットするタイプが存在するときに協力的な結果を近似するような均衡戦略が存在することを示す。これは評判形成に関する従来の研究とは異なる評判を形成することになる。本稿と同様の評判形成に着目した論文としては Mailath and Samuelson (1998) と Mailath and Samuelson (2001) がある。本稿で示す評判の形成は現実においてしばしば観察される現象である。「社会には非協力的なひとがいるが、自分はそうではない」というような評判を形成することによって協力を達成しようとすることはよくある。本稿で示す結果は十分意義があると言えよう。

本稿の構成は以下の通りである。2 節ではモデルの設定を行なう。3 節では協力的な結果を近似するような逐次均衡戦略を示す。4 節では均衡戦略の特徴と評判形成について解説をする。

2 モデル

以下のようなランダム・マッチング・モデルを考える。社会は $2n$ 人から構成され、その社会はグループ $N^1 = \{1, \dots, n\}$ とグループ $N^2 = \{n+1, \dots, 2n\}$ に分けられる。 $n \geq 2$ であると仮定する。グループ N^1 とグループ N^2 のメンバーがあるマッチング・ルールにしたがってランダムに組み合わせられ、 n 組のチームが形成される。それぞれのチームはふたりから構成され生産が行われる。それぞれのチームでは以下の対称的な囚人のジレンマ・ゲームがプレーされるものとする。

	C	D
C	$1, 1$	$-l, 1+g$
D	$1+g, -l$	$0, 0$

それぞれのチームはグループ N^1 と N^2 から選ばれたメンバーによって構成されているが、グループ N^1 出身のメンバーをチームではプレーヤー 1 と呼び、 N^2 出身のメンバーをプレーヤー 2 と呼ぶことにする。それぞれのプレーヤーは行動 C または D を選択することができる。

無限繰り返しゲームは以下の手順で行われる。第 1 期のはじめにランダム・マッチングが行われ、チームが形成される。それぞれのチームのプレーヤーはゲームが行われる最初の期に自然 (Nature) によってタイプが決定される。 $1-q$ の確率で通常のタイプとなり、 q の確率でコミットメント・タイプになる。そのチームが存続する限り、最初に決められたタイプ

は変更されないと仮定する。また、自分のタイプを知ることができるが、相手プレイヤーのタイプを観察することはできないと仮定する。

コミットメント・タイプになった場合は D のみを選択し続け、相手がどのような行動を選択したとしてもペイオフ $\bar{u}$ ($0 < \bar{u} < 1/2$) を得るものとする。一方、通常のタイプは各期において自分の意思で行動を選択できる。プレイヤーは每期同時に行動を選択する。プレイヤーが行なう意思決定は行動に関する決定、つまり、 C または D を選択することとモニタリングに関する決定からなる。モニターすると自分が属するチームの相手プレイヤーがその期に選択した行動のみを観察することができる。他のチームで何が起きているかを知ることにはできない。モニターした場合は相手が選択した行動を完全に知ることはできるが、観測費用 $\lambda > 0$ を伴うとする。モニターしなかった場合は、観測費用を支払う必要はないが、相手の行動に関する情報を全く得ることはできないとする。³ プレイヤー i のモニタリングに関する決定は私的情報であり、プレイヤー j ($j \neq i$) は自分がモニターされているかどうかを知ることにはできないと仮定する。また、議論の簡単化のため、行動とモニタリングの決定は同時に行なわれるものとする。⁴ プレイヤーが行なう意思決定は以下のように表される。例えば、 $C * \lambda$ と書いた場合は、 C を選択しモニターすることを表す。 $C * \phi$ と書いた場合は C を選択しモニターしないことを表す。 $D * \lambda$ 、 $D * \phi$ についても同様に解釈する。

プレイヤーはチームにおける生産が停止されたときにのみペイオフをまとめて受け取るものと仮定する。つまり、この仮定により、生産活動中にプレイヤーはペイオフを受け取らないので、モニターすることでしか相手プレイヤーの行動に関する情報を得ることはできないことを意味する。

生産活動が停止されるふたつの可能性があるとする。ひとつはある確率で強制的に社会全体における生産活動がすべて停止されるものである。 $1 - \delta$ の確率で停止するものと仮定する。つまり、 δ の確率で社会における生産活動は継続される。生産継続の確率 δ は通常の繰り返しゲームにおける割引因子 (discount factor) として解釈でき、プレイヤーのペイオフは通常の割引利得 (discounted payoff) で表される。つまり、ペイオフの流列 ($u_i^1, u_i^2, \dots, u_i^t, \dots$) が与えられたとき、プレイヤー i の割引利得は以下のように表される。

$$\sum_{t=1}^{\infty} \delta^{t-1} u_i^t.$$

各期の平均利得に注目するために、以下の分析ではプレイヤーの平均割引利得 (discounted average payoff) に注目する。 v_i をプレイヤー i の平均割引利得とすると次のように定義される。

$$v_i = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} u_i^t.$$

もうひとつの生産停止の可能性はそのチームに属するメンバーが生産の停止に同意した場合である。各期の終りに生産を続けるかどうかを決めることができる。行動と観測の意思決定を行なった後に生産を継続するか停止するかをプレーヤー達は同時に表明しなければならない。生産継続・停止の表明はコミットメント・タイプも自分の意思で決定できるものとする。生産継続・停止の表明はチーム内のメンバーのみが知ることができ、他のチームでどのような表明が行なわれたかを知ることとはできないと仮定する。チームの両プレーヤーが生産停止を表明した場合にのみ生産は停止される。それ以外は生産を継続しなければならない。チームによる自発的な生産の停止がなされた場合は、強制的な社会全体の生産停止がなされない限り、解散されたチームに属していたグループ N^1 出身のメンバーと N^2 出身のメンバーの新たなランダム・マッチングによって、新しいチームが形成される。⁵ 新しいチームで生産が行なわれる前に、プレーヤーの新たなタイプは $1-q$ と q の確率で通常のタイプかコミットメント・タイプに決定される。

各期の意思決定の後にに社会の構成員全員に観察可能であるサンスポットが存在すると仮定する。サンスポット m は区間 $[0, 1]$ から一様分布にしたがって実現するものとする。

3 均 衡 戦 略

この節では協力的な結果 $(v_1^*, v_2^*) = (1, 1)$ が逐次均衡 (sequential equilibrium) として近似されるような環境が存在することを示す。つまり、協力的な結果を近似するようなパラメータ λ , δ , q が存在することを示す。ここで取り上げる均衡戦略は次の手順にしたがってプレーされる。

均衡経路上：

通常の種類：

ステップ1： $C*\lambda$ を $1-q$ の確率で選択し、 q の確率で $C*\phi$ を選択する。

ステップ2：前の期の意思決定に関係なく、実現したサンスポットが $m \in (\hat{m}, 1]$ であれば、生産停止を表明し、 $m \in [0, \hat{m}]$ であれば継続を表明する。

ステップ3：(i) チームが継続される場合：前の期に自分が $C*\lambda$ を選択し、相手が C を選択したことを観察したときは、ステップ1に進む。相手が D を選択したことを観察した場合、または、自分が $C*\phi$ または D を選択した場合はステップ4に進む。(ii) 生産が停止された場合：新チームでステップ1からはじめる。

ステップ4： $D*\phi$ を選択し、ステップ2に進む。

コミットメント・タイプ：

実現したサンスポットが $m \in (\hat{m}, 1]$ であれば、生産停止を表明し、 $m \in [0, \hat{m}]$ であれば

継続を表明する。

均衡経路外：

過去にステップ1または3で逸脱した場合は、 $D*\phi$ を選択し続ける。ただし、ステップ2にはしたがう。ステップ2で逸脱した場合は、以後 $D*\phi$ を選択し続け、サンスポットに関わらず継続を表明し続ける。ステップ2での複数の逸脱は無視する。⁶

通常のタイプの戦略 $(\hat{\sigma}_1, \hat{\sigma}_2)$ は要約すると次のように行なわれる。 $C*\lambda$ を選択し続け、相手プレーヤーがCを選択し続けたことを観察した場合は、 $C*\lambda$ と $C*\phi$ をそれぞれ $1-q$ と q の確率で選択し続ける。 $C*\phi$ を選択した次の期以降は、チームが継続される限り、 $D*\phi$ を選択し続ける。過去にDを選択した場合は $D*\phi$ を選択し続ける。そして、サンスポットの値に応じて生産停止または継続の表明をする。新しいチームに参加した場合も同様の手順にしたがう。

命題：それぞれの $\varepsilon > 0$ に対して、以下を満たすような $\bar{\delta} \in (0, 1)$ 、 $\bar{q} > 0$ が存在する。

それぞれの $\delta \in [\bar{\delta}, 1)$ 、 $q \in (0, \bar{q}]$ に対して、すべての $\lambda \in (0, \bar{\lambda}(\delta, q)]$ 、すべての $i=1, 2$ について $|v_i - v_i^*| \leq \varepsilon$ を満たす逐次均衡戦略が存在するような $\bar{\lambda}(\delta, q) > 0$ が存在する。ここで、 v_i は逐次均衡における通常のタイプの平均割引利得を表している。

証明：まず、次の方程式を満たすような $\delta \in (0, 1)$ 、 $\lambda > 0$ 、 $q > 0$ 、 $\hat{m} = m(\delta, q, \lambda) \in [0, 1]$ が存在することを示す。

$$v_i = (1 - \hat{m}\delta)\{(1-q) - ql - \lambda\} + \hat{m}\delta(1-q)v_i \quad (1)$$

$$= (1 - \hat{m}\delta)\{(1-q) - ql\} + \hat{m}\delta(1 - \hat{m}\delta)(1-q)^2(1+g) \quad (2)$$

$$\geq 1 - \varepsilon. \quad (3)$$

最初の等号より、以下が得られる。

$$v_i = \frac{(1 - \hat{m}\delta)\{(1-q) - ql - \lambda\}}{1 - \hat{m}\delta(1-q)} \geq 1 - \varepsilon. \quad (4)$$

任意に $\hat{m} \in (0, 1)$ と $\delta \in (0, 1)$ を固定したとき、 $v_i > 1 - \varepsilon$ をみたすような $q > 0$ と $\lambda > 0$ が存在する。また、二番目の等号より、次が得られる。

$$(1 - \hat{m}\delta)(1-q)(1+g) = v_i - \frac{(1 - \hat{m}\delta)\lambda}{\hat{m}\delta(1-q)}. \quad (5)$$

このとき、(4)を満たすような q と λ について、(5)を満たすような $\delta \in (0, 1)$ と $\hat{m} \in (0, 1)$

が存在している。 $\bar{\delta} \in (g/(1+g), 1)$, $\bar{q} > 0$ は以下を満たすものとする。

$$\frac{(1-\hat{m}\bar{\delta})\{(1-\bar{q})-\bar{q}l\}}{1-\hat{m}\bar{\delta}(1-\bar{q})} > 1-\varepsilon.$$

このとき、それぞれの $\delta \in [\bar{\delta}, 1)$, $q \in (0, \bar{q}]$ について、(1)–(3)を満たすような $\bar{m} \in (0, 1)$ が存在する。また、それぞれの $\delta \in [\bar{\delta}, 1)$, $q \in (0, \bar{q}]$ について、以下で定義されるような $\bar{\lambda}(\delta, q)$ が存在する。つまり、 $\bar{\lambda}(\delta, q)$ は(1)–(3)を満たし、かつ、以下を満たす。

$$0 < \bar{\lambda}(\delta, q) \leq \left| \frac{(1-\hat{m}\delta)ql}{(1-\hat{m}\delta)(1-q)-\hat{m}\delta} \right|. \quad (6)$$

与えられた戦略が逐次均衡であることを示すためには、それぞれのステップにおいて、その戦略によって導出される信念 (belief) の下で、逸脱の誘因が存在しないことを示せばよい。

まず、ステップ1で逸脱の誘因がないことを示すには $C*\lambda$ と $C*\phi$ を選択することは無差別であり、 D を選択する誘因がないことを示せばよい。プレーヤー i がステップ1に到達するのはチームでの第1期目、または、 $C*\lambda$ を選択し続け、相手が C を選択し続けたことを観察した場合のいずれかである。どちらの場合についてもプレーヤー i はプレーヤー j が $1-q$ の確率で戦略 $\hat{\sigma}_j$ にしたがう、 q の確率で D を選択し続けるという信念を形成することになる。このとき、そのチームでの生産が継続された場合のプレーヤー i が $C*\lambda$ を選択したときに得られる平均割引利得は(1)で表され、 $C*\phi$ を選択したときの平均割引利得は(2)で表される。すでに上で示したように $\hat{m}$ を適切に決めることによって $C*\lambda$ と $C*\phi$ を無差別にすることができる。また、 D を選択したときの平均割引利得は $(1-\hat{m}\delta)(1-q)(1+g)$ となる。(5)より、 $(1-\hat{m}\delta)(1-q)(1+g) < v_i$ であるので、逸脱の誘因はない。よって、ステップ1における逸脱の誘因は存在しないことが示された。

ステップ2については相手プレーヤーが上記の戦略にしたがうとき、サンスポットの値が $m \in [0, \hat{m}]$ であれば、相手プレーヤーは継続を表明するので、プレーヤー i は停止を表明したとしても生産を停止させることはできない。停止の表明は逸脱行為と認識されるため、以後ゼロのペイオフを得ることになるので、継続を表明することによって損をすることはない。よって、通常のタイプもコミットメント・タイプも逸脱の誘因はない。また、 $m \in (\hat{m}, 1]$ については、プレーヤー j は停止を表明するので、生産が継続されるか停止されるかはプレーヤー i の表明によって決まる。停止を表明した場合、新しいチームが形成されることになるのでそれぞれのプレーヤーが均衡戦略にしたがうのであれば、新チームで通常のタイプになった場合は v_i の平均割引利得、コミットメント・タイプになった場合は $\bar{u}$ を得ることになる。つまり、停止の表明による将来の期待利得は $(1-q)v_i + q\bar{u} > 0$ である。一方、継続を表明した場合、現在のチームが継続されることになり、逸脱行為を働いたとして、将来得られる利得は通常のタイプはゼロ、コミットメント・タイプは $\bar{u}$ となる。よって、停止を表明すること

は最適である。つまり、それぞれのタイプについて生産継続・停止の決定の逸脱の誘因はない。

ステップ3については、次の2つのケースを示せばよい。

- (i) 前の期において $C * \lambda$ を選択し、相手プレーヤーが C を選択したことを知った場合はステップ1に戻る誘因がある。
- (ii) 前の期において $C * \phi$ を選択したときは、チームが継続する限り $D * \phi$ を選択し続ける誘因がある。

(i)については、前の期において $C * \lambda$ を選択し、相手プレーヤーが C を選択したことを知った場合は、プレーヤー j は $1-q$ の確率で戦略 $\hat{\sigma}_j$ を選択し、 q の確率で D を選択し続けるという信念が形成される。これはステップ1と同様の状況である。ステップ1における逸脱の誘因がないことはすでに示されているのでケース(i)は示されたことになる。

(ii)については次の期にはプレーヤー j は $(1-q)^2$ の確率で戦略 $\hat{\sigma}_j$ にしたがう、 $1-(1-q)^2$ の確率で D を選択し続けるという信念が形成される。このとき、 $D * \phi$ を選択したときの平均割引価値は $(1-\hat{m}\delta)(1-q)^2(1+g)$ となる。一方、 $D * \phi$ 以外を選択したときは高々 $(1-q)v_i - (1-\hat{m}\delta)ql + \lambda$ である。

$$\begin{aligned} & \{(1-\hat{m}\delta)(1-q)^2(1+g)\} - \{(1-q)v_i - (1-\hat{m}\delta)ql + \lambda\} \\ &= \frac{1}{\hat{m}\delta} [(1-\hat{m}\delta)ql - \{(1-\hat{m}\delta)(1-q) - \hat{m}\delta\}\lambda]. \end{aligned} \quad (7)$$

このとき、すべての $\lambda \leq \bar{\lambda}(\delta, q)$ について(7)は非負である。よって、 $C * \phi$ を選択した次の期には $D * \phi$ を選択することは最適である。

最後にステップ4について考える。プレーヤー i がある期において D を選択した場合、戦略 $\hat{\sigma}_j$ にしたがうのであれば、次の期以降はプレーヤー j は D を選択し続ける。よって、プレーヤー i は $D * \phi$ を選択し続けることが最適である。■

4 モニタリングの誘因と評判形成

この節では均衡戦略において、モニタリングの誘因がどのように導き出されているのか、そして、どのような評判効果が働いているのかを説明する。

4.1 モニタリングの誘因

前節では観測費用が十分小さく、コミットメント・タイプが実現する確率が十分小さい場合、協力的な結果を近似できることを示した。協力的な結果を近似するような均衡戦略を考

えるときに最も重要な点は、どのようにしてプレーヤーにモニタリングの誘因を与えるかということである。モニタリングの誘因を与えるためには将来の行動に対する不確実性が必要である。

まず、プレーヤーが純粋戦略にしたがうとしたとき、均衡においてどのような結果が導出されるか考えてみよう。もし、通常のタイプのプレーヤー 1 と 2 が第 1 期目に C を選択するような純粋戦略にしたがうと仮定すると、それぞれのプレーヤーは第 1 期目においてのみ、相手がコミットメント・タイプかどうかをチェックすることになる。相手がコミットメント・タイプであるとわかれば、2 期目以降は $D * \phi$ を選択し続けることが最適である。一方、相手プレーヤーが C を選択したことを観察したプレーヤーは、相手が通常のタイプであることを知ることになる。そして、 C を観察したプレーヤーは相手が純粋戦略にしたがうという信念を形成することになる。よって、2 期目以降はモニタリングしなくても相手が何を選択するかを知ることになる。このとき、それぞれのプレーヤーは相手は 2 期目以降モニタリングをしないことが最適となる。プレーヤー達がこのことを考慮すれば、2 期目以降は $D * \phi$ を選択し続けることが最適となる。以上により、プレーヤー達は 1 期目の行動に関係なく、2 期目以降は $D * \phi$ を選択することになる。1 期目の行動は将来の行動に影響を与えないので、それぞれのプレーヤーは 1 期目においても $D * \phi$ を選択することが最適である。1 期目に C を選択するような純粋戦略の均衡は存在しない。純粋戦略の均衡が存在するとすれば、均衡においてはすべてのプレーヤーが D を選択し続けるものしかない。

前節で構築した均衡戦略において、将来の行動に対する不確実性がどのように生み出され、モニタリングの誘因が与えられているのかを解説する。逐次均衡において通常のタイプのプレーヤーは高い確率 $(1-q)$ でモニターすることを選択する。 t 期まで、プレーヤー i が $C * \lambda$ を選択し続け、 C を観察し続けたとき、プレーヤー i は相手プレーヤー j が $t+1$ 期にあたかも C と D をランダムに選択してくるかのような信念を形成することになる。相手プレーヤーが C を選択したときと D を選択したときでは $t+2$ 期以降の行動は異なる。 $t+1$ 期にモニタリングをしなければ、観測費用 λ を節約できるが、 $t+2$ 期に C を選ぶべきか D を選ぶべきかが不確実になる。均衡戦略においては $t+1$ 期にモニタリングをすることとしないことは無差別になっているので、観測の誘因は保たれている。

観測費用が存在する場合、将来の行動に関する不確実性を内生的に作り出す必要がある。観測費用を導入した他の論文においても、将来の行動に関する不確実性を作り出すことによって観測の誘因を導出している。Ben-Porath and Kahneman (2002), Miyagawa, Miyahara, and Sekiguchi (2003) においても将来の行動に関する不確実性を利用して観測の誘因を導出している。

4.2 評判形成

本稿で導入したコミットメント・タイプは、これまでの評判効果 (reputation effect) の分析において導入されているコミットメント・タイプとは異なっている。従来の評判効果の分析ではコミットメント・タイプとしてターゲットとなるペイオフを達成するような行動にコミットするプレーヤーを導入していた。例えば, Kreps, Milgrom, Roberts, and Wilson (1982) (以下 KMRW) では、完全観測下の囚人のジレンマの有限回繰り返しゲームについて協力的な結果を近似できるかどうかを分析している。KMRW ではコミットメント・タイプとして仕返し戦略 (tit-for-tat strategy) にコミットするプレーヤーを導入している。もし、プレーヤー 1 が通常のタイプであり、プレーヤー 2 が仕返し戦略にコミットしているとし、プレーヤー 1 がそのことを知っていれば、プレーヤー 1 は最終期の一期前まで協力し続けることが最適である。KMRW では相手プレーヤーがコミットメント・タイプである確率がある程度あれば、有限繰り返し囚人のジレンマ・ゲームにおいて、通常のタイプが最終期近くまで協力を続けるような逐次均衡が存在することを示している。⁹

また、不完全公的観測下の分析では Fudenberg and Levine (1992) で分析されている。Fudenberg and Levine (1992) ではひとりの長期間生き続けるプレーヤー 1 がそれぞれの期に一期間だけ生きるプレーヤー 2 とゲームをプレーするという繰り返しゲームを分析している。そこでは、プレーヤー 1 の行動はプレーヤー 2 に観察可能ではないが、公的に観察可能なシグナルが存在し、シグナルの実現確率はプレーヤーが選択する行動に依存する。プレーヤー 1 にはある確率でコミットメント・タイプになる可能性があるとき、均衡で達成されるプレーヤー 1 の平均割引利得は割引率が 1 に近づくにつれてコミットメント・タイプがコミットする戦略によって得られる平均割引利得に近づくことを示している。

このように、完全観測下、不完全観測下においても、コミットメント・タイプを真似するような均衡が存在することを多くの研究は示してきた。¹⁰つまり、プレーヤーが自分はコミットメント・タイプであるかのような評判を形成する均衡が存在することを明らかにしてきたのである。一方、自分がコミットメント・タイプではないという評判を形成することによって協力的な結果が導出されるという研究はほとんどないといってよい。私の知る限り、そのような観点に注目した研究は Mailath and Samuelson (1998) と Mailath and Samuelson (2001) のみである。本稿では、そのような評判を形成することによって協力的な結果を達成することは可能であることを明らかにした。均衡戦略にしたがうとき、第 1 期が終わった後にはモニターすれば相手のタイプがわかることになる。観測しなかったプレーヤーは相手がどのタイプであるかを知らずにプレーを続ける。第 1 期目に自分が非協力的なタイプではないということを示すことによって、高い確率で協力が続けられるのである。つまり、砕けた表現をすれば、「社会には非協力的なひともいるが、自分はそうではない」という評判を形成

することになる。

本稿の拡張として、コミットメント・タイプが協力的な行動にコミットする場合に従来の研究と同様にコミットメント・タイプを真似するような均衡が存在するかどうかを明らかにすることは興味深い。例えば、コミットメント・タイプがトリガー戦略のみを選択するものと仮定するのである。このとき、プレーヤーにモニタリングの誘因を与えるためには、通常のプレーヤーが非協力的な行動 D を均衡において正の確率で選択する必要がある。つまり、モニタリングの誘因と D を選択する誘因のふたつの制約を満たすような均衡戦略を構築しなければならない。しかしながら、Miyahara (2002) で指摘されているように、これらを満たすような均衡戦略の構築は非常に困難であると予想される。この問題については将来の課題としたい。

注

本研究は部分的に全国銀行学術研究振興財団と科学研究費補助金（課題番号15730185）の援助を受けている。

本稿の作成にあたり、小林創講師（大阪府立大学）、末廣英生助教授、久本久男助教授（神戸大学）より貴重なコメントを頂きました。記して感謝します。

- 1 Ben-Porath and Kahneman (2002) ではコミュニケーションが可能である場合、フォーク定理が成立することを示している。Miyagawa, Miyahara, and Sekiguchi (2003) ではコミュニケーションが可能ではない場合について、Nash Threat フォーク定理が成立するための十分条件を示している。
- 2 Miyahara (2002) では均衡戦略のクラスとして、均衡経路上で将来の戦略 (continuation strategy) が、Brandenburger and Dekel (1987) の意味で、common-1-believed であるようなものに焦点を当てている。また、一回の観測で相手が選択した過去 2 期間の行動を知ることができる場合には、観測費用が十分小さければ、将来の戦略が common-1-believed となるような均衡戦略で協力的な結果が導出できることを示している。
- 3 この仮定については次の段落で解説される。
- 4 行動を選択した後にモニターするかどうかを決めるというように逐次的に意思決定をするものと仮定しても本稿での議論は変わらない。
- 5 ただし、ランダム・マッチングの直後に前の期のチームメイトとチームを構成することも起こりうる。
- 6 つまり、ステップ 2 での複数の逸脱の結果、生産が停止された場合は新しいチームでステップ 1 にしたがう、生産が継続された場合はステップ 3 にしたがう。
- 7 プレーヤー i は、第 1 期目については、確率 q でコミットメント・タイプが D を選択し続けるという信念を形成し、第 2 期目以降については前の期に $C * \phi$ を選択した通常のタイプが $D * \phi$ を選択し続けるという信念を形成する。
- 8 均衡戦略においては通常のタイプは実際に C と D をランダムに選択するわけではない。
- 9 他の完全観測下の有限繰り返しゲームの分析には Kreps and Wilson (1982), Milgrom and

Roberts (1982), Fudenberg and Maskin (1986), Fudenberg and Levine (1989), 無限繰り返しゲームについては Schmidt (1993) などがある。それぞれの論文において導入されているコミットメント・タイプは KMRW と同様に、ターゲットとするペイオフを達成するために都合のよい戦略にコミットするプレーヤーを考えている。例えば, Fudenberg and Maskin (1986) と Schmidt (1993) ではゲームの第 1 期目にミニマックス戦略にしたがうコミットメント・タイプが正の確率で生じるというモデル化を行なっている。

- 10 本稿の設定においても D を選択し続けるというコミットメント・タイプの戦略を真似するような逐次均衡が存在する。

参 考 文 献

- Ben-Porath, E., and M. Kahneman (2002): "Communication in Repeated Games with Costly Monitoring," *forthcoming in Games and Economic Behavior*.
- Brandenburger, A., and E. Dekel (1987): "Common Knowledge with Probability 1," *Journal of Mathematical Economics*, 16, 237-245.
- Fudenberg, D., and D. Levine (1989): "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, 57, 759-778.
- Fudenberg, D., and D. Levine (1992): "Maintaining a Reputation When Strategies Are Imperfectly Observed," *Review of Economic Studies*, 59, 561-579.
- Fudenberg, D., and E. Maskin (1986): "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54 533-556.
- Kandori, M. (1992): "Social Norms and Community Enforcement," *Review of Economic Studies*, 59, 63-80.
- Kreps, D., P. Milgrom, J. Roberts, and R. Wilson (1982): "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," *Journal of Economic Theory*, 27, 245-252.
- Kreps, D., and R. Wilson (1982): "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253-279.
- Mailath, G., and L. Samuelson (1998): "Your Reputation is Who You're Not, Not Who You'd," CARESS Working Paper No. 98-11, University of Pennsylvania.
- Mailath, G., and L. Samuelson (2001): "Who Wants a Good Reputation?," *Review of Economic Studies*, 68, 415-441.
- Miyagawa, E., Y. Miyahara, and T. Sekiguchi (2003): "Repeated Games with Observation Costs," mimeo.
- Milgrom, P., and J. Roberts (1982): "Predation, Reputation and Entry Deterrence," *Journal of Economic Theory*, 27, 280-312.
- Miyahara, Y. (2002): "Near Efficiency in Repeated Prisoner's Dilemma with Costly Monitoring," mimeo.
- Schmidt, K. (1993): "Reputation and Equilibrium Characterization in Repeated Games with Conflicting Interests," *Econometrica*, 61, 325-351.