慶應義塾大学学術情報リポジトリ
Keio Associated Repository of Academic resouces

Title	自発的繰り返し囚人のジレンマにおける確率的受け入れと漸次協力の効果について
Sub Title	On stochastic acceptance and gradual cooperation in voluntarily repeated prisoner's dilemma with no information flow
Author	グレーヴァ, 香子(Fujiwara-Greve, Takako)
Publisher	慶應義塾経済学会
Publication year	2008
Jtitle	三田学会雑誌 (Keio journal of economics). Vol.101, No.3 (2008. 10) ,p.465(69)- 489(93)
JaLC DOI	10.14991/001.20081001-0069
Abstract	近年, 自発的繰り返しゲームの研究が進んできた。自発的繰り返し囚人のジレンマにおいて協力を達成するには, 常に協力したときの利得より生涯利得を低くしなくてはならない。その方法として, 出会ってもパートナーとして受け入れない確率を導入する(確率的受け入れ)か, プレイはするが最初の何回かは協力しない(漸次協力)戦略が均衡として示されてきた。本稿ではこれらの均衡の利得を比較することで, どちらの戦略が相対的に望ましいかを考える。 Lately, progress has been made in the research of voluntarily repeated games. To achieve cooperation in a voluntarily repeated prisoner's dilemma, the lifetime payoff must be set lower than the total payoff from repeated cooperation. Two kinds of equilibria are proposed to generate this structure: stochastic acceptance equilibria where players accept a newly matched partner only with some probability less than one to start the game, and gradual cooperation equilibria where players accept each other for sure but do not cooperate in some initial periods. In this study, I compare the equilibrium payoffs of these equilibria to determine which strategy is better.
Notes	小特集 : 経済の数理 : 非線形動学と経済の変動を中心に
Genre	Journal Article
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00234610-20081001-0069

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

自発的繰り返し囚人のジレンマにおける確率的受け入れと漸次協力の効果について

On Stochastic Acceptance and Gradual Cooperation in Voluntarily Repeated Prisoner's Dilemma with No Information Flow

グレーヴァ　香子(Takako Fujiwara-Greve)

近年，自発的繰り返しゲームの研究が進んできた。自発的繰り返し囚人のジレンマにおいて協力を達成するには，常に協力したときの利得より生涯利得を低くしなくてはならない。その方法として，出会ってもパートナーとして受け入れない確率を導入する（確率的受け入れ）か，プレイはするが最初の何回かは協力しない（漸次協力）戦略が均衡として示されてきた。本稿ではこれらの均衡の利得を比較することで，どちらの戦略が相対的に望ましいかを考える。

Abstract

Lately, progress has been made in the research of voluntarily repeated games. To achieve cooperation in a voluntarily repeated prisoner's dilemma, the lifetime payoff must be set lower than the total payoff from repeated cooperation. Two kinds of equilibria are proposed to generate this structure: stochastic acceptance equilibria where players accept a newly matched partner only with some probability less than one to start the game, and gradual cooperation equilibria where players accept each other for sure but do not cooperate in some initial periods. In this study, I compare the equilibrium payoffs of these equilibria to determine which strategy is better.

「三田学会雑誌」101 巻 3 号（2008 年 10 月）

自発的繰り返し囚人のジレンマにおける確率的受け入れと漸次協力の効果について*

グレーヴァ香子

要　　旨

近年，自発的繰り返しゲームの研究が進んできた。自発的繰り返し囚人のジレンマにおいて協力を達成するには，常に協力したときの利得より生涯利得を低くしなくてはならない。その方法として，出会ってもパートナーとして受け入れない確率を導入する（確率的受け入れ）か，プレイはするが最初の何回かは協力しない（漸次協力）戦略が均衡として示されてきた。本稿ではこれらの均衡の利得を比較することで，どちらの戦略が相対的に望ましいかを考える。

キーワード

自発的繰り返し，囚人のジレンマ，相関戦略，漸次協力，効率性

JEL classification

C 73

1.　はじめに

1.1　自発的繰り返しゲームの定義とその研究意義

標準的な非協力ゲーム理論において，同じような戦略的相互関係が何回も続くという状況は，主として 2 つのモデルでこれまで分析されてきた。一つのモデルは繰り返しゲーム（例えば， Aumann (1981)， Fudenberg and Maskin (1986) を参照）であり，もう一つのモデルはランダム・マッチングゲーム（例えば， Matsushima (1990)， Kandori (1992)， Ellison (1994) を参照）である。繰り返しゲームでは，同じプレイヤーたちが同じ段階ゲームを繰り返し行う。ランダム・マッチングゲームでは，一回ごとにランダムにプレイヤーたちが出会って同じ段階ゲームを行う。これらのモデルはプレイヤーの組み合わせの決め方としては両極端のケースであって，繰り返しゲームでは相手は最初から最後まで変わらず，ランダム・マッチングゲームでは，毎回相手が変わる。

＊　本稿をまとめるにあたり，中山幹夫先生，武藤滋夫先生に有用なコメントをいただいたことをここに記して感謝いたします。

現実はこれらのモデルの中間であり，何らかのプロセスで出会ってゲームをした相手と，また同じゲームをプレイしてもいいし，それきりということも可能であることが多い。また，相手がどう決まるかということは，出会いや別れを決める外的要因と共に，当事者たちの選択によるはずである。例えば，経済取引において，最初に取引をするかどうかは，出会いのプロセスなど外的な要因も大きいであろうが，同じ相手と繰り返し同じ取引をすることはよくあり，それは売り手と買い手の意思で決まってくる。

このように，長期的ゲームにおいて，ゲームのルールによって，プレイヤーの組み合わせが決められているのは，応用上必ずしも妥当でない。それが本稿につながる筆者の一連の研究の基本的問題意識である。

本稿では，ランダムに出会い，出会った相手とゲームをするかを選び，一回ゲームをする毎にお互いこの相手とまたプレイするかを選ぶオプションがあるモデルを考察する。一人でも継続しないことを選択すると関係は終わり，その場合はまたランダムに相手が決まる。これを「自発的繰り返しゲーム」(Voluntarily Repeated Games) と呼ぶ[(1)]。このモデルでは，プレイヤーたちの選択により，ずっと同じ相手と同じ段階ゲームをプレイすることもできるし，毎回やめてランダムな相手と再スタートすることもできる。従って，繰り返しゲームとランダム・マッチングゲームの形が当事者によって選べるモデルとなっている。また，同じ相手と繰り返し続ける場合はお互いの行動は完全に観察可能であるとするが，新しい相手とゲームを始める際には，お互いの過去の行動の履歴がわからないという情報の切断があると仮定する。もし，新しい相手の過去の行動がわかるのであれば，それを利用して通常の繰り返しゲームにおける戦略と似たような戦略をとることができ，あまり新しい知見は得られないからである。相手を変えることができるというだけでなく，過去から逃げることができるということになると，通常の繰り返しゲームでフォーク定理を成立させているトリガー戦略などは使えないので，このモデルの意義がある。

1.2　既存研究と本稿の研究の関係および本稿の研究の意義

特に，段階ゲームが囚人のジレンマタイプのケースについて，近年いくつかの研究が蓄積されてきた[(2)]。囚人のジレンマタイプのゲームでは，どのようにゲームが続くかによって，非常に明暗が別れることが特に興味のある問題であるからである。

例えば，1 回限り，あるいは有限回の繰り返し囚人のジレンマでは，効率的な行動（これを「協力」

---

(1)　Fujiwara-Greve and Okuno-Fujiwara (2008) では Voluntarily Separable Repeated Prisoner's Dilemma としているが，彼らのモデルでは出会った当初に相手を受け入れるオプションがないので，厳密には本稿のモデルと異なり，繰り返しをどこでやめるかが問題となるためこのような名前になっている。

(2)　以下で紹介されている論文以外には，例えば Datta (1996)， Ghosh and Ray (1996)，Kranton (1996) などがある。

と呼ぶことにする）は均衡においてプレイできない。しかし，終わりが明確にならない繰り返し囚人のジレンマになると，十分に将来が重要であれば，トリガー戦略などを使って均衡経路で協力することができる (Fudenberg and Maskin (1986))。ランダム・マッチングゲームでは，現在の相手以外との記憶をまったく利用せずに行動を決めるとすれば実質上 1 回限りの囚人のジレンマになるので，やはり協力できないが，自分の過去を覚えていて，それに依存して行動を決めるという戦略が可能であれば，過去に裏切られたプレイヤーは今後ずっと全ての相手を裏切るという戦略を取れば，プレイヤーの数が有限の場合，最初に裏切りを発生させたプレイヤーにも必ず「裏切りの連鎖」が戻ってくる，という因果応報のメカニズムで協力的な均衡が成立する (Kandori (1992)，Ellison (1994))。

では，自発的繰り返しゲームにしたらどうか。通常の繰り返しゲームと異なって，利己的な行動をとって一回限りの高い利得を得た上で逃げてしまうことができる場合は，裏切りを知っているプレイヤーとは二度と会わないので，トリガー戦略は使えない。また，プレイヤーの人数が非常に多ければ「裏切りの連鎖」も意味をなさないし，ここでは情報の切断を仮定するので，自分の過去も利用しない戦略のみを考えたい。従ってこれらの戦略とは異なる形で協力のインセンティヴを与えなくてはならない。

これまでの研究では, 主として 2 種類[(3)]の均衡で, 効率的行動をさせることに成功している。Fujiwara-Greve (2002) では，相関戦略を用い，新たに出会った相手とは必ずしも確実にゲームをプレイしない，という「確率的受け入れ」を導入した。すると，裏切った後，新たな相手とゲームを始めるにはある程度の期間待つことが予想され，現在の協力関係を維持し続けるときの利得より低くなれば，このような戦略は協力的な均衡となる。これに対し Fujiwara-Greve and Okuno-Fujiwara (2008) では，出会うということはゲームを少なくとも 1 回はプレイすると解釈するので必ずパートナーシップは形成されるのであるが，最初の数回は協力せず，利得を低くしておき，それでもゲームが続いたら協力を開始するという「漸次協力」戦略[(4)]が均衡となることを示した。

これらの均衡はいずれも，相手を変えると利得が下がるという構造になっているのがポイントである。そのため，協力せず自分だけが高い利得を得ると，相手がゲームを降りてしまうので自分も新たなパートナーを探さなくてはならなくなり，それが「罰」として働くということである。しかも，このような形の利得構造でなければ，1 回限りのゲームにおいて合理的ではない協力行動を誘導することはできない。ところで，各プレイヤーの生涯利得は，当初相手がいない時点からの利得であるから，実は「罰」に入ったときの利得と同じである。従って，通常の無限回繰り返し囚人のジレンマとは根本的に異なって，自発的繰り返し囚人のジレンマにおいては，毎回お互いに協力すると得られる効率的な利得は，均衡では得られないのである。

---

(3)　この他に，プレイヤーたちが異なる戦略を取る非対称均衡でも，効率的な行動をある期以降ずっとさせることができることが Fujiwara-Greve and Okuno-Fujiwara (2008) で示されている。

(4)　あるいは，「信頼構築」(trust building) 戦略とも呼ばれる。

つまり，上記の 2 種類の均衡はいずれも効率的ではない。そこで，本稿ではどちらが相対的に効率的であるかを調べてみる。 Fujiwara-Greve (2002) では，出会ったときに確率的にお互いを受け入れるのであるが，裏切って得られる利得との兼ね合いで，受け入れる確率があまり高くてはならない。そのような確率にうまく調整するには二人が共通に観察できる確率的なシグナルが必要であるが，うまく存在するかはわからない。これに対し， Fujiwara-Greve and Okuno-Fujiwara (2008) の均衡では，そのような事前の調整メカニズムは必要ないが，パートナーシップの最初の何回かは協力できない。離散的時間のモデルでは，意思決定はある間隔をおいてしかできないので，初期の非協力の期間が不必要に長い可能性がある。

このように，2 種類の均衡にはそれぞれ異なった理由で利得のロスが存在している。その比較をすることは，純粋理論的な興味のみならず，どのような状況ではどちらのタイプの均衡のロスが大きいかを調べることで，状況に応じて適切な均衡をとるように誘導する方策を考えることができ，政策的な含意も得られる。

2. モデル

同質のプレイヤーが $[0,1]$ 区間に連続的に存在する大きな社会を考える。各プレイヤーは，離散時間 $\tau = 1, 2, \ldots$ において意思決定を行い，将来利得を $\delta \in (0,1)$ で割り引いて評価するとする。

長期ゲームは以下のように行われる。$\tau = 1$ の期初には全てのプレイヤーが「パートナーなし」の状態であるとする。$\tau = 2, 3, \ldots$ 期の状態は外生的な要因とプレイヤーの戦略的退出という内生的な要因の両方によって決まる。

各期初において，パートナーがいない状態のプレイヤーたちはランダムに出会うプロセスに参加する。（市場のようなものを想定すればよい。）ここで，確率 p で他のプレイヤーと出会えるとする。もし，他のプレイヤーと出会えなければ，今期は何もせず，$\underline{u}$ という利得だけを得て，パートナーなしの状態のまま，次期になるまで待つしかないとする。

出会ったプレイヤーのペアは，お互いの過去の行動の履歴がわからない下で，この相手とゲームに入る（行動 a，これは Accept を意味する）か，拒絶する（行動 r，これは Reject を意味する）かを同時に，あるいはお互いの選択を知らずに選ぶ。二人とも行動 a を選択した場合，二人はパートナーとなり，囚人のジレンマ（表 1）をプレイする。一人でも行動 r を選択したら，二人ともパートナーなしの状態に戻り，誰とも出会えなかったのと同じことになる。（つまり，今期これ以降は何もせず $\underline{u}$ を得て次期にランダムプロセスに参加する。）

パートナーとなった二人がプレイする囚人のジレンマ（表 1）は通常の 2 行動のもので，C は協力的行動，D は利己的行動と解釈する。利得の大小関係は $g > c > d > \ell$ とする。これによって，

表 1　囚人のジレンマ

P1\ P2	C	D
C	c, c	ℓ, g
D	g, ℓ	d, d

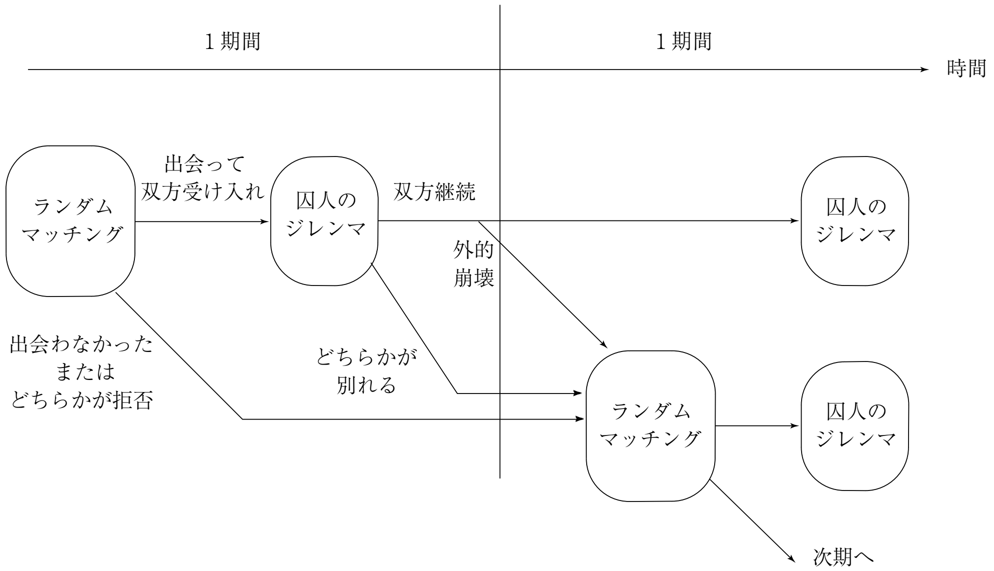
1 回だけこのゲームをプレイする場合は利己的行動 D が各プレイヤーにとって支配戦略であるが，二人とも協力的行動 C をした方が，二人とも利己的な行動 D をするより効率的であることになる。この他にも利得にいくつかの仮定をおく。まず，$2c > g + \ell$ であるとする。この仮定は，協力的行動 C を二人ともすることが効率的であることを保証している。また，$c > \underline{u}$ を仮定する。そうでないと，パートナーを得て協力する意味がない。さらに $\underline{u} > d$ を仮定する。これにより，お互い利己的な行動しかしないのであれば別れた方がいいという状況を設定する。

段階ゲーム（囚人のジレンマ）での行動はパートナー同士にしか観察されないとする。今期の囚人のジレンマが終ったら，二人はそれぞれ現在の相手と継続する（行動 m，これは Maintain を意味する）か，別れる（行動 e，これは End を意味する）かを同時に決定するとする。この意思決定はこれまで観察された過去のお互いの行動の履歴に依存してよい。この後，外的な理由により，パートナーシップが壊れる確率を $1 - q$ とする。（転居や景気の悪化によるビジネスチャンスの喪失などを想定するとよい。）したがって，次期に同じ相手とプレイできるのは，二人とも行動 m を選び，かつ外的な理由でパートナーシップが壊れないときである。また，このように外的要因でもパートナーシップが壊れることがあるので，ランダムマッチング・プロセスに参加するということが即ち誰かを裏切ってきた，ということにはならず，過去の履歴がわからないという仮定と整合する。外的要因または選択によってパートナーシップが終了した場合，二人とも次期の期初にパートナーなしの状態として始めることになる。このように，パートナーシップの長さ $t = 1, 2, \ldots$ は二人の戦略および外的要因によって決まる。図 1 はこの長期ゲームのアウトラインを示したものである。

次に戦略を定義する。まず，$H_1 = \{\emptyset\}$ を，空な履歴のみからなる集合と定義し，これは新しい相手と出会ったとき（$t = 1$ のとき），お互いの過去の行動の情報がないことを示している。各 $t = 2, 3, \ldots$ について，$H_t = (\{C, D\} \times \{C, D\})^{t-1}$ を t 期以前のパートナーシップ内の行動の履歴の集合とする[(5)]。このとき，一人のプレイヤーの純戦略は以下の性質を満たす関数の列の組み合わせ $s = (f^A, \{f_t^G\}_{t=1}^{\infty}, \{f_t^C\}_{t=1}^{\infty})$ である。

(5)　観察可能な行動の列としては最初の受け入れと各期の継続の意思決定もあるが，双方が受け入れ，毎期継続を選ぶことだけがその後の意思決定に結びつくので，戦略の定義には必要ない。

図１　ゲームのアウトライン



[受け入れ] $f^A : H_1 \to \{a, r\}$ は，新しい相手と出会ったとき行動 a または r を選ぶルール，

[囚人のジレンマ] 各 $t = 1, 2, \ldots$ について，$f_t^G : H_t \to \{C, D\}$ はこれまでのパートナーシップ内の行動の履歴に応じて今期の囚人のジレンマでの行動を選ぶ関数，

[継続] 各 $t = 1, 2, \ldots$ について，$f_t^C : H_t \times (\{C, D\} \times \{C, D\}) \to \{m, e\}$ は，今期を含んだパートナーシップ内の行動の履歴に応じて継続するかどうかを決める関数。

この定義では，戦略はゲームそのものの時間 $\tau = 1, 2, \ldots$ には依存しない。（ただし，パートナーシップが何期続いているかという t には依存してもよい。）また，ある戦略 s を採用するということは，全ての新しい相手に対して同じ s を用いるということも含まれている。これらは一見，本来の戦略集合を限定するように見える仮定であるが，たとえゲーム内の時間 τ に依存するような戦略を考えても，相手が変わるごとに情報が失われてしまう以上，時間だけを利用しても利得を上げることはできない[(6)]。出会った相手によってその後の継続戦略を変えようとしても，情報の切断の仮定により，それは不可能である。しかも，この「やり直し」の構造は，生涯利得の構造に再帰性をもたらす（(1)式参照）ので分析が簡単になるとともに，本稿では考察されていないが戦略の選択や観察に不確実性がある場合，利得のロスを少なくするものであり，現実的にも望ましいと言える。

S を上記の定義の純戦略全体の集合とし，$S \times S$ 上の確率分布全体の集合，$\Sigma = \Delta(S \times S)$ を相関戦略プロフィール全体の集合とする。各戦略 s について，パートナーシップ内の行動の履歴を一

(6)　また，「はじめに」で述べたように，自己の記憶を利用して裏切りの連鎖を考えても，非常に大きい人口を仮定しているので「罰」として働かない。

つ決めたとき，その後の s の継続戦略とは，その履歴を含んだ履歴の集合に s を制限した戦略（即ち定義域を制限し，行動の選び方は s と同じ戦略）のことである。同様に，戦略の組み合わせについても，その継続戦略の組み合わせを定義することができる。

全てのプレイヤーの戦略の組み合わせ（あるいは相関戦略プロフィール）と，ランダムな出会いと崩壊の確率過程とによって，各プレイヤーに確率的な利得の列が与えられる。各プレイヤーは，その期待割引総利得

$$U(\sigma) = E\sum_{\tau=1}^{\infty}\delta^{\tau-1}u(\tau;\sigma)$$

を最大にするよう行動すると仮定する。ここで，σ は戦略の組み合わせ，または相関戦略プロフィール，$u(\tau;\sigma)$ は社会全体の戦略分布が σ であるとき τ 時点において一人のプレイヤーが得られる利得である。本稿では対称戦略のみを考えるので，プレイヤーの名前は必要ない。期待値は，σ に含まれる相関確率，出会いの確率，外生的崩壊確率についてを合わせて計算するものとする。以上で長期ゲームの定式化が完成した。ゲームは完備情報とする。

このゲームは展開形ゲームであり，しかも新たに出会ったときに過去がわからないので，不完全情報のゲームである。そこで，均衡概念としては逐次均衡 (sequential equilibrium) を考える[7]。まず，ゲーム全体にゆきわたる信念の体系 (belief system) を定義する。X を展開形ゲームの意思決定点全体の集合とし，$\mathcal{P}$ を X の情報分割とする[8]。$\mathcal{P}$ の元 I がいずれかのプレイヤーの情報集合である。

定義：信念の体系 (belief system) とは，関数 $\mu : X \to [0,1]$ で，各情報集合について足し合わせると 1 になっているもの，即ち，任意の $I \in \mathcal{P}$ について $\sum_{x\in I}\mu(x) = 1$ となるものである。

定義：ある戦略プロフィール σ と信念の体系 μ の組み合わせ (σ,μ) が逐次均衡 (sequential equilibrium) であるとは，以下の 2 つの条件が成立することである。

[逐次合理性] 任意のプレイヤーとそのプレイヤーの任意の情報集合 I について，これまで起こったことを μ で予想し，今後起こることを σ の継続戦略の組み合わせで予想したとき，σ のこのプレイヤーの部分の戦略が最適である；

[整合性] σ に収束する戦略の組み合わせ $\{\sigma^k\}_{k=1,2,\ldots}$ として，各 $k = 1,2,\ldots$ について，σ^k は全ての意思決定点に正の確率を付けるものが存在し，σ^k からベイズルールによって定義される $\mu^k : X \to [0,1]$ の収束先が μ である。

---

(7)　逐次均衡の最初の定義は Kreps and Wilson (1982) にある。不完全情報の繰り返しゲームにおける逐次均衡の分析としては， Abreu et al. (1986), Kandori and Matsushima (1998) などがある。

(8)　展開形ゲームにおける概念である，意思決定点，情報分割，情報集合などについては 岡田 (1997) などゲーム理論の教科書に詳しく述べてある。

整合性は，均衡戦略がわずかな戦略の揺れについて頑健であることを保証するものである。しかし，本稿では明示的には信念の体系を構築しないで，逐次合理性のみをチェックする方法を取る。なぜなら，(1) このモデルでは，各プレイヤーの利得に影響を及ぼすのは，現在の相手のみで他のペアの行動の部分についてどのような信念をもっていようとも，戦略の最適性の計算にはまったく関係がない，かつ，(2) パートナーシップ内では完全に行動が観察されるので信念の部分は明らかである，からである。

また，本稿では対称戦略のみに着目するので，以下の形の戦略が逐次均衡であるとする。全てのプレイヤーがある（相関）戦略 s をしているとき，その戦略プロフィールを σ_s と書く。σ_s が逐次均衡であるとは，各プレイヤーとその人が直面するパートナーシップ内の任意の行動の履歴について，パートナーが s の継続戦略をとり，ペアが崩壊した場合，ランダムに出会う新しい相手も s をしているとき，s の継続戦略をとることがこのプレイヤーにとっても最適であることである。

対称戦略にしぼることについては，少なくとも 2 つの正当化ができる。一つは，全てのプレイヤーが同質であるので，同じ思考をすると仮定するのは自然であることである。2 つ目は，対称戦略の均衡は一つの「社会規範的行動」として解釈しやすいということである。(例えば Okuno-Fujiwara and Postelwaite (1995) も同じ議論をしている。)「A さんはこれこれ，B さんはこれこれをするのがよい」というものより，「皆，正直に取引しなさい」という方が社会的規範として理解しやすい。

以上で分析の準備ができた。ちなみに，s として，全てのプレイヤーが新しい相手と出会ったとき拒否（行動 r）を選び，万一パートナーシップが形成されたら D を選んで，お互い何をしようと一回で別れるという戦略を考えると，σ_s は逐次均衡である。従ってこのゲームに逐次均衡は常に（任意のパラメターについて）存在する。

3. 確率的受け入れ均衡

3.1 確率シグナルが連続体で存在する場合

この節では，プレイヤーたちが出会ったときに相手を受け入れてゲームを始めるかどうかを相関させるための装置として，0 から 1 の間の任意の確率をもったシグナルが使用できると仮定する。例えば，出会ったときにルーレットを回し，ある範囲の数値が出ればお互いに相手を受け入れ，そうでなかったら拒否するというような方法である。このルーレット盤が連続的に分けられれば，お互いを受け入れる確率を任意の数値にすることができるわけである。経済学的には，例えば太陽黒点均衡と似たようなことを考えればよい。ゲームそのものとは関係ないが，プレイヤーたちが気にする何らかの確率過程が存在し，それを見て新たな相手とゲームをするかどうかを判断するということである。しかも，実際にそのシグナルに合わせて行動を調節することが均衡となれば，ゲームと

はまったく関係ないとわかっていても，皆がそのシグナルに従う以上，自分も従うことになる。

具体的には，シグナルの出る確率 α をパラメターとして，以下の「確率的受け入れ」戦略 $s(\alpha)$ を全てのプレイヤーが行うことが逐次均衡となるような条件を求める。

- ランダムマッチング・プロセスで誰かと出会った場合，α の確率で発生するシグナルが観察されたら，またそのときのみ新しい相手を受け入れる行動 a をとる。
- 囚人のジレンマでは，パートナーシップ内の行動の履歴が空であるか，あるいは (C, C) のみがプレイされてきたときだけ C をとる。そうでないときは D をとる。
- 今期の囚人のジレンマで (C, C) が観察されたら，またそのときのみ，継続の行動 m を選ぶ。

全てのプレイヤーが $s(\alpha)$ を行うとき，パートナーなしの状態でこれからランダムマッチング・プロセスに入る時点の一人のプレイヤーの長期利得を U とし，ある期初に既にパートナーがいる状態のとき以降の長期利得を V とすると，U と V は以下の連立方程式を満たす。

$$U = p\{\alpha V + (1-\alpha)(\underline{u} + \delta U)\} + (1-p)(\underline{u} + \delta U), \quad (1)$$

$$V = c + \delta\{qV + (1-q)U\}. \quad (2)$$

まず，（1）式の説明であるが，p の確率で出会いがあり，さらに α の確率でシグナルを観察したときだけ囚人のジレンマに入ることができる。そこから先の利得は定義により V である。シグナルを観察しなかった場合お互いに受け入れないのでパートナーシップが形成されず，今期は $\underline{u}$ をもらい，次期はまたパートナーなしの状態なので割引将来利得は δU となる[(9)]。$1-p$ の確率で出会えないので，そのときも今期の利得は $\underline{u}$，次期以降の割引将来利得は δU である。

次に，（2）式であるが，パートナーがいるので，すぐに囚人のジレンマをプレイできる。全員が $s(\alpha)$ に従っていてパートナーがいるということは，これまでの行動の履歴は空か (C, C) のみから成り立っているはずである。したがってお互いに C を取るので，今期の利得は c となる。次期以降は，q の確率でこの関係が継続したときは再びパートナーありの状態となり，V が継続利得となる。$1-q$ の確率で外的要因で関係が解消された場合，パートナーなしの状態となるので，U が継続利得となる。

（1）式と（2）式を連立して解くと，明示的に

$$U = \frac{\alpha pc + (1-\delta q)(1-\alpha p)\underline{u}}{(1-\delta)\{1-\delta q(1-\alpha p)\}}, \quad (3)$$

$$V = \frac{\{1-\delta(1-\alpha p)\}c + \delta(1-q)(1-\alpha p)\underline{u}}{(1-\delta)\{1-\delta q(1-\alpha p)\}} \quad (4)$$

---

(9)　ここの U と，左辺の U が同じであるのは戦略の定義の中に，新たな相手と出会ったとき，まったく同じ戦略をとるということが入っているからである。

と出せる。(3) 式と (4) 式から，任意の利得パラメター (g, c, d, ℓ) とその他のパラメター (p, q, δ, α) について

$$V - (\underline{u} + \delta U) = \frac{c - \underline{u}}{1 - \delta q(1 - \alpha p)} > 0 \tag{5}$$

と

$$V - U = \frac{(1 - \alpha p)(c - \underline{u})}{1 - \delta q(1 - \alpha p)} > 0 \tag{6}$$

が導かれる。これらを使って以下の命題が証明できる。

命題 1：任意の $p \in (0, 1]$ に対し $\alpha \in (0, 1]$ が存在して，全てのプレイヤーが確率的受け入れ戦略 $s(\alpha)$ をとることが逐次均衡である必要十分条件は

$$\frac{g - c}{(g - \underline{u})\delta q} < 1 \tag{7}$$

である。

証明：動的計画法により，各意思決定点について，$s(\alpha)$ から 1 回だけ逸脱し，その後は $s(\alpha)$ に戻るとすると，ずっと $s(\alpha)$ をしているときより利得が高まらないことを調べる。

1. 受け入れ時点：シグナルが観察されなければ，相手は拒否してくるので，こちらは何をしても利得は変わらない。従って弱い意味で $s(\alpha)$ に従うのが最適である。シグナルが観察された場合，二人は受け入れるはずであるが，逸脱して拒否行動 r をとると，継続利得は $\underline{u} + \delta U$ となる。逸脱しなければ，V が継続利得であるから，(5) より逸脱しない方がよい。
2. 囚人のジレンマの時点：経路上では，囚人のジレンマの段階に来ているということは，これまで誰も D をしていないということであり，$s(\alpha)$ に従うと今後の利得は V である。ここで一回だけ逸脱して D をとると，今期は g を得るが，この相手とは終わりになるので，今後の利得は $g + \delta U$ となる。従って，逸脱しない条件は

$$V - (g + \delta U) \geqq 0 \iff \delta q(1 - \alpha p) \geqq \frac{g - c}{g - \underline{u}} \tag{8}$$

   である。
3. 継続の時点：今期 D を観察した場合，パートナーがやめるので，弱い意味で行動 e は最適である。今期 (C, C) を観察した場合，パートナーは継続を選ぶが，逸脱してこちらが行動 e を選ぶと，継続利得は U，$s(\alpha)$ に従って継続すれば V であるから，(6) より逸脱しても利得は高くならない。

以上のことから，$s(\alpha)$ を全員がすることが逐次均衡であるための必要十分条件は（8）式であることがわかる。これを書き換えると，

$$\alpha \leqq \Big\{1-\frac{g-c}{(g-\underline{u})\delta q}\Big\}\frac{1}{p}$$

となる。この条件を満たす $\alpha > 0$ が，任意の $p \in (0,1]$ に対して存在するためには，右辺の中かっこ内が正であることが必要十分である，即ち $\frac{g-c}{(g-\underline{u})\delta q} < 1$ が必要十分条件である。 □

系 1 (Fujiwara-Greve (2002))**：**戦略 $s(\alpha)$ を全員がする戦略プロフィールが逐次均衡になるような $(p,\delta,q,\alpha) \in (0,1] \times (0,1) \times (0,1) \times (0,1]$ が存在する。
証明：（8）式の右辺は厳密に 0 と 1 の間にあるので，δq が十分大きく，αp が十分小さければこの不等式は成立する。 □

以下では δq が十分に小さく，（7）式が常に成立すると仮定し，確率的合意均衡が存在する範囲で議論をする。

出会いの確率は経済構造に関係するので[(10)]，p の変化に注目し，各 p について，もっとも利得の高い均衡を調べる。これを後に，漸次協力戦略均衡の中でもっとも利得の高いものと比較することになる。任意の p について，確率的受け入れ戦略均衡の利得を最大にする受け入れ確率は

$$\alpha^*(p) := \min\Big\{\Big(1-\frac{g-c}{(g-\underline{u})\delta q}\Big)\frac{1}{p}, 1\Big\}$$

で表される。p が 0 から 1 へと増えていくと，$\alpha^*(p)$ は 1 から $1-\frac{g-c}{(g-\underline{u})\delta q}$ へと減少していく。これは，出会いの確率が高まると，それに応じて受け入れ確率を低くしないと，裏切って逃げても十分高い確率で新たなパートナーとゲームを始めることができてしまい，「罰」にならないからである。逆に言うと，出会いの確率 p が低ければ，$\alpha = 1$ ということも均衡になり，とにかく出会ったらすぐ協力関係を始めることができるようになる。かつての日本で，新卒以外の労働市場がほとんどなかった時代に，企業と労働者が非常に協力的な長期関係を保っていたのもこの論理で説明できる。

p の関数として，確率的受け入れ均衡の最大利得を求める。まず，$\alpha^*(p) < 1$ となる境目の p は

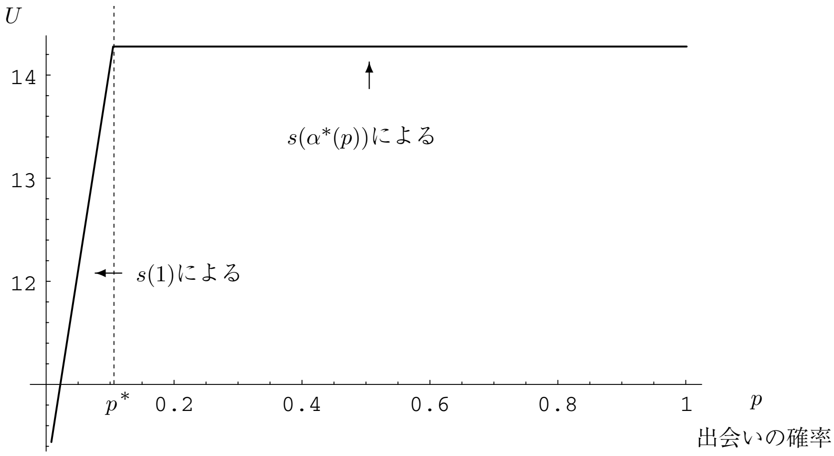
$$p^* = 1-\frac{g-c}{(g-\underline{u})\delta q}$$

である。$p \leqq p^*$ のときは，$\alpha = 1$ が最大の利得を与え，そのときの利得は

$$U(s(1)) = \frac{pc + (1-\delta q)(1-p)\underline{u}}{(1-\delta)\{1-\delta q(1-p)\}}$$

（10） 割引率も利子率と関連して議論されるが，心理的要因もあると思われるので，ここでは細かく分析していない。

図２　確率的受け入れ均衡の最大利得



で表される。これは

$$\frac{\partial U(s(1))}{\partial p} = \frac{(1-\delta q)(c-\underline{u})}{(1-\delta)\{1-\delta q(1-p)\}^2} > 0$$

より p の増加関数である。p^* より大きい p の場合，$\alpha^*(p)$ がうまく調整して，ぎりぎり逸脱しない条件をもたらすので，均衡利得は一定で

$$U(s(\alpha^*(p))) = \frac{\delta q g - (g-c)}{\delta q(1-\delta)} \tag{9}$$

となる。U が $p \leqq p^*$ において増加関数であることから，これが最大利得であることがわかる。（図2は $g=10, c=6, d=2, \ell=0, \delta=0.8, q=0.7$ の数値例でこのことを示している。これ以降の図もすべてこれらのパラメター値による。）ところで，通常の繰り返し囚人のジレンマの場合，対称均衡の最大利得は $\frac{c}{1-\delta}$ であり，

$$\frac{c}{1-\delta} - \frac{\delta q g-(g-c)}{\delta q(1-\delta)} = \frac{(1-\delta q)(g-c)}{(1-\delta)\delta q} > 0$$

より，自発的繰り返しにすると最大利得は効率的でなくなることがわかる。

また，政策により p をいくら大きくしても，均衡利得は $U(s(\alpha^*(p)))$ より大きくならない。囚人のジレンマをプレイしているのであれば，協力させるためにはしかたのないことである。従って，コストをかけてマッチングを改善することはある程度以上の効果を持たず，囚人のジレンマの構造を改善するような政策，あるいは勝手にゲームをやめられないようにする政策などの方が望ましい。もう一つの政策としては，情報構造の改善がある。 Fujiwara-Greve et al. (2008) では，転居や工場閉鎖などのやむを得ない事情でゲームをやめた場合「紹介状」が出るというモデルを使って，多少なりとも情報の切断が解消され，利己的な行動による関係解消かそうでないかの情報が伝達されれば，より高い利得をもたらす均衡が存在することが示されている。

3.2　有限個の確率シグナルが存在する場合

これまでは，任意の α について，プレイヤーたちが相関戦略をとることができると仮定して，均衡利得の構造を調べた。本節では，限られた確率シグナルしか使えないと，均衡利得の構造がどうなるかを調べる。使用可能な確率の集合を有限集合 $A \subseteq (0,1)$ とする。もし，ある確率 α で現れ，二人が観察可能なシグナルが存在するならば，逆にそのシグナルが出ない，ということも確率的受け入れに使えるので，$\alpha \in A$ ならば $1-\alpha \in A$ である。このことから，A は，小さい確率から順に $A=\{\alpha_1,\alpha_2,\ldots,\alpha_K\}$ のように書くことができ，各 k について $\alpha_K=1-\alpha_1$, $\alpha_{K-1}=1-\alpha_2$ となっている。さらに，A の要素が一つであるなら，それは 0.5 でなくてはならないこともわかる。もし，A が 0.5 と異なる要素 α を持つのであれば，$1-\alpha$ も使えるので，A が非空であることは即ち 0.5 以上の確率をもつシグナルが少なくとも 1 つ存在することを意味する。これは後の分析で有用となる。

均衡の必要十分条件（8）式を満たすためには，p が増加したら，だんだんに低い α のシグナルを使用する戦略にしなければならない。その中で最大の利得を得るためには出来る限り $\alpha^*(p)$ に近く，かつそれより低い確率をもつシグナルを使用することになる。これを数式で表現すると以下のようになる。各 $k=1,2,\ldots,K$ について，p_{α_k} を

$$\alpha_k=\alpha^*(p_{\alpha_k})$$

で定義し，α_M を，A の要素で $p_{\alpha_M} \leqq 1$ となる最小のものとする。最後に

$$\hat{\alpha}^*(p)=\begin{cases}1 & \text{if } p \leqq p^* \\ \alpha_1 & \text{if } p^* < p \leqq p_{\alpha_1} \\ \alpha_k & \text{if } p_{\alpha_{k-1}} < p \leqq p_{\alpha_k} \qquad \text{for } k=2,3,\ldots,M\end{cases}$$

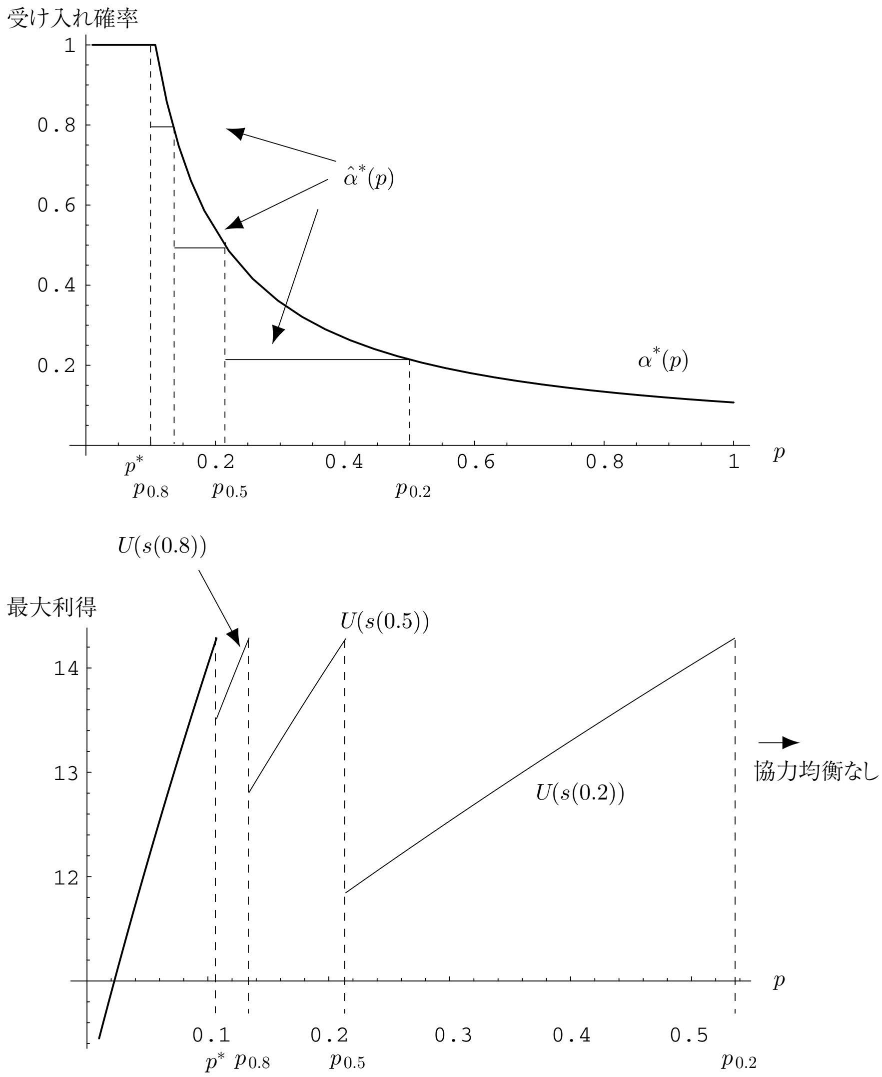
とすると，各 $p \leqq p_{\alpha_k}$ について，受け入れ確率 $\hat{\alpha}^*(p)$ は $\alpha^*(p)$ を超えない範囲で最大である。（図 3 上のグラフを参照。）

また，p_{α_M} を超えた p については（8）式を満たす α が存在しないため，確率的合意受け入れ均衡は存在しないことになる。（ただし，2 節の終わりに述べた拒否均衡は常に存在する。）区間 $(p_{\alpha_{k-1}},p_{\alpha_k}]$ 内では同じ戦略 $s(\alpha_k)$ が均衡となるので，その利得は p の増加関数となり，ちょうど $p=p_{\alpha_k}$ となるところで（8）式が等式で満たされ，均衡利得は最大となる。（図 3 の下のグラフ参照。）

4.　漸次協力均衡

本節では，出会った相手と必ずプレイするが，当初の T 期間は D をプレイし，その後 C をプレイするという，以下で定義される漸次協力戦略 $s'(T)$ による均衡を分析する。

図３　シグナルの集合が $\{0.2, 0.5, 0.8\}$ の場合の最適戦略と最大利得



1. ランダムマッチング・プロセスで出会った相手は必ず受け入れる（行動 a をとる）。
2. パートナーシップが $t \leqq T$ 期間目である場合，D をプレイし，今期 (D, D) を観察したときだけ継続の行動 m を選ぶ。
3. パートナーシップが $t > T$ 期間目である場合，C をプレイし，今期 (C, C) を観察したときだけ継続の行動 m を選ぶ。

社会全体が $s'(T)$ をする対称戦略プロフィールが逐次均衡であるためには，当初の非協力期間 T が調節されなければならない。社会全体が $s'(T)$ をしているとき，一人のプレイヤーがパートナーなしの状態で，これからランダムに相手がみつかる時点からの長期利得を U'，既にパートナーシップ

に入っていて t 期目の期初からの長期利得を $V'(t)$ とすると，これらは以下の連立方程式を満たす。

$$\begin{aligned}
U' &= (1-p)(\underline{u}+\delta U') + pV'(0) \\
V'(0) &= \{1+\delta q+\cdots+\delta^{(T-1)}q^{(T-1)}\}d + (\delta^T q^T+\cdots)c \\
&\quad +\delta(1-q)\{1+\delta q+\delta^2 q^2+\cdots\}U', \\
&= \frac{1-\delta^T q^T}{1-\delta q}d + \frac{\delta^T q^T}{1-\delta q}c + \frac{\delta(1-q)}{1-\delta q}U'.
\end{aligned}$$

これらを解くと，

$$U' = \frac{(1-\delta q)(1-p)\underline{u} + p\{(1-\delta^T q^T)d + \delta^T q^T c\}}{(1-\delta)\{1-\delta q(1-p)\}} \tag{10}$$

が得られる。また，$V'(t)$ は，$t \leqq T$ については

$$\begin{aligned}
V'(t) &= \{1+\delta q+\delta^2 q^2+\cdots+\delta^{(T-t-1)}q^{(T-t-1)}\}d + \{\delta^{(T-t)}q^{(T-t)}+\cdots\}c \\
&\quad +\delta(1-q)\{1+\delta q+\delta^2 q^2+\cdots\}U' \\
&= \frac{1-\delta^{(T-t)}q^{(T-t)}}{1-\delta q}d + \frac{\delta^{(T-t)}q^{(T-t)}}{1-\delta q}c + \frac{\delta(1-q)}{1-\delta q}U',
\end{aligned}$$

$t \geqq T+1$ については

$$\begin{aligned}
V'(t) &= V'(T+1) = (1+\delta q+\delta^2 q^2+\cdots)c + \delta(1-q)(1+\delta q+\delta^2 q^2+\cdots)U' \\
&= \frac{1}{1-\delta q}c + \frac{\delta(1-q)}{1-\delta q}U'
\end{aligned}$$

と書ける。任意の $t \leqq T+1$ について，時間が経てば (D,D) をする期間が少なくなるので，明らかに $V'(t) > V'(t-1)$ である。$T+1$ 期目以降は，ずっと (C,C) を続けるだけなので，$T+1$ 期目以降の長期利得 $V'(t)$ は t について一定である。

さらに，

$$V'(0) - (\underline{u}+\delta U') = \frac{\delta^T q^T(c-d) - (\underline{u}-d)}{1-\delta q(1-p)},$$

また，

$$V'(0) - U' = \frac{(1-p)\{\delta^T q^T(c-d) - (\underline{u}-d)\}}{1-\delta q(1-p)}$$

が成立する。従って，$V'(0) \geqq \underline{u}+\delta U'$ かつ $V'(0) \geqq U'$ が成立する必要十分条件は

$$(\delta q)^T(c-d) \geqq (\underline{u}-d) \tag{11}$$

である。このとき $V'(t) > V'(t-1)$ から，任意の $t \geqq 1$ について $V'(t) > \underline{u}+\delta U'$ かつ $V'(t) > U'$ も言える。これらの準備により，以下の命題が導かれる。

命題2：任意の自然数 T について，全てのプレイヤーが $s'(T)$ をとることが逐次均衡である必要十分条件は（11）および

$$(\delta q)^T(c-d) \leqq (\underline{u}-d)+\frac{1-\delta q(1-p)}{\delta pq}\{\delta q(g-\underline{u})-(g-c)\} \tag{12}$$

が同時に成立することである。

証明：各意思決定点について逐次合理性を調べる。

1. 受け入れ時点：新たに出会った相手は必ず受け入れてくれるので，$s'(T)$ に従って受け入れれば，今後の利得は $V'(0)$ である。逸脱して拒否すると今後の利得は $\underline{u}+\delta U'$ である。（11）より $V'(0)>\underline{u}+\delta U'$ が成立するので，逸脱の方が利得が低い。
2. 囚人のジレンマの時点：$t \leqq T$ のときは $s'(T)$ に従って D をとると今後の利得は $V'(t)$ であるが，逸脱して C をとると今後の利得は $\ell+\delta U'$ となり，（11）より逸脱の方が利得が低い。$t \geqq T+1$ のときは，$s'(T)$ に従って C をとると今後の利得は $V'(T+1)$ であるが，逸脱して D をとると $g+\delta U'$ が利得となる。これらを比較すると，

$$\begin{aligned}&\{V'(T+1)-(g+\delta U')\}(1-\delta q)\{1-\delta q(1-p)\}\\ =\ &\{1-\delta q(1-p)\}\{\delta q(g-\underline{u})-(g-c)\}+\delta pq(\underline{u}-d)-\delta pq(\delta q)^T(c-d)\end{aligned}$$

   より，（12）が成立すれば，上式が正となり，逸脱の利得の方が低い。
3. 継続の時点：$t \leqq T$ 時点で (D,D) を観察した場合，$s'(T)$ に従って継続を選択すれば今後の利得は $V'(t+1)$ であるが，逸脱して別れると今後の利得は U' となり，（11）より逸脱の利得の方が低い。$t \geqq T+1$ で (C,C) を観察した場合も同様である。その他の（経路外の）行動の組み合わせを観察した場合，パートナーが別れを選択してくるので，何をしても利得は変わらない。

このように全ての情報集合において一回限りの逸脱の利得は $s'(T)$ の利得を上回らないので，$s'(T)$ は逐次合理的であり，全てのプレイヤーが $s'(T)$ をとることは逐次均衡である。 □

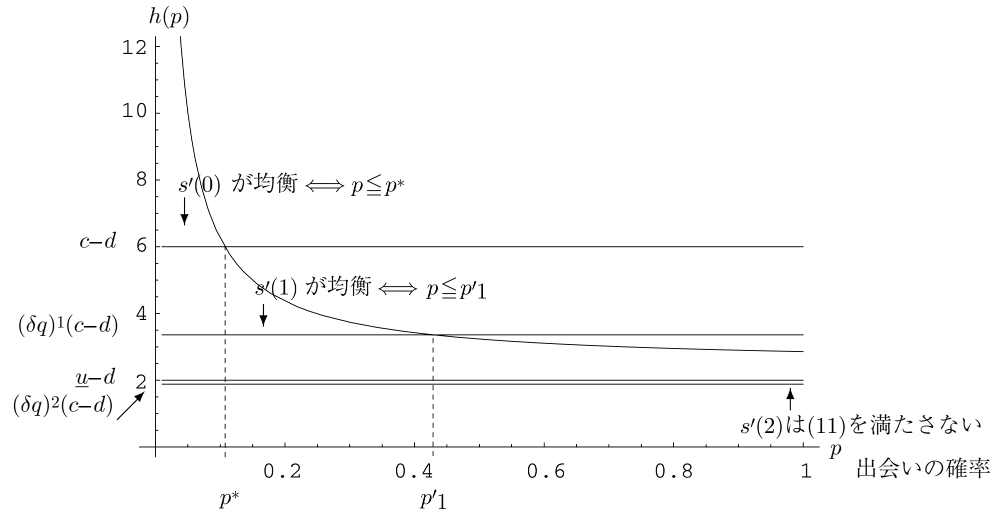
そこで，各 $p \in (0,1)$ について，どのような T が上記の十分条件（11）と（12）を満たすのかを調べる。（12）の右辺を

$$h(p) := (\underline{u}-d)+\frac{1-\delta q(1-p)}{\delta pq}\{\delta q(g-\underline{u})-(g-c)\}$$

と置くと，（12）は $(\delta q)^T(c-d) \leqq h(p)$ と書ける。

$$\frac{\partial h}{\partial p} = -\frac{(1-\delta q)\{\delta q(g-\underline{u})-(g-c)\}}{\delta p^2 q}$$

図 4　$0 < p \leqq 1$ についての漸次協力均衡の存在



より，(7) の下では h は p の厳密な減少関数であることがわかる。p が増加すると $h(p)$ は $\underline{u} - d$ に上から漸近する。(図 4 を参照。) $T = 0$ から順に見て行くと，$T = 0$ のときの (11) と (12) の左辺は $c - d(> \underline{u} - d)$ であり，p^* の定義から

$$p \leqq p^* \iff c - d \leqq h(p)$$

であるから，$0 < p \leqq p^*$ の範囲では (11) と (12) が同時に成立し，$s'(0)$ は逐次均衡であることがわかる。($s'(0)$ は $s(1)$ とまったく同じ戦略であるから当然である。)

$p'_1 \in (0, 1)$ を

$$\delta q(c - d) = h(p'_1)$$

で定義すると，$s'(1)$ が逐次均衡となるのは，$\delta q(c - d) \geqq \underline{u} - d$ かつ $p \leqq p'_1$ が成立するときである。ただし前者は一般には必ずしも成立しないことに注意が必要である。

一般には，$\overline{T}$ を (11) が成立する最大の T としたとき，各 $T = 0, 1, \ldots, \overline{T}$ について $p'_T \in (0, 1)$ を

$$(\delta q)^T (c - d) = h(p'_T)$$

で定義すると，任意の $p \leqq p'_T$ について $s'(T)$ を全員がプレイすることは逐次均衡である。(また，$p'_0 = p^*$ である。)

h は厳密な減少関数であるので，図 4 からも明らかなように，任意の T について $p'_T < p'_{T+1}$ が成立する。$s'(T)$ が逐次均衡ならば $s'(T+1)$ も均衡となるが，$p'_T < p \leqq p'_{T+1}$ の範囲で最も利得が高いのは，D をする期間が最も少ないものであるから $s'(T)$ 戦略による均衡である。だんだんに T を長くしていくと，$\overline{T}$ に達して (11) が満たされなくなるか，あるいは p'_T が 1 を超えるので，ど

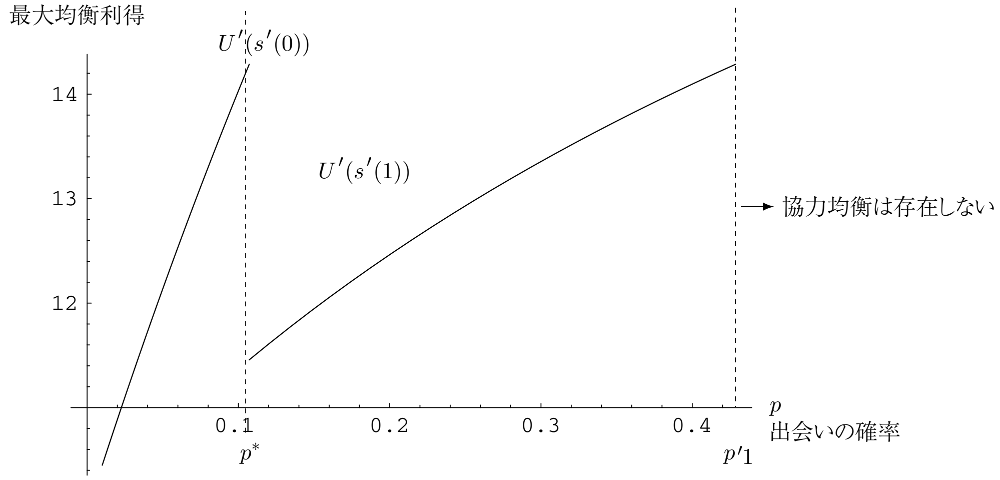


図５　漸次均衡の最大利得

こかで均衡が存在しなくなる。これらをまとめると，T^* を $p'_T < 1$ となる最大の T と定義すると，各 p について最も効率的な均衡利得は以下のようになる。

$$
\begin{aligned}
U'(s'(0)) &\iff 0 < p \leqq p'_0 (= p^*) \\
U'(s'(T)) &\iff p'_{T-1} < p \leqq p'_T \quad \forall T = 1, 2, \ldots, T^*, \\
U'(s'(T^*+1)) &\iff p'_{T^*} < p \text{ and } \ T^* + 1 \leqq \bar{T}.
\end{aligned}
$$

ここから導かれる漸次協力均衡の最大利得のグラフは例えば図５のようになる。各 T について，ちょうど $p = p'_T$ であるときに最大利得が達成され，それは

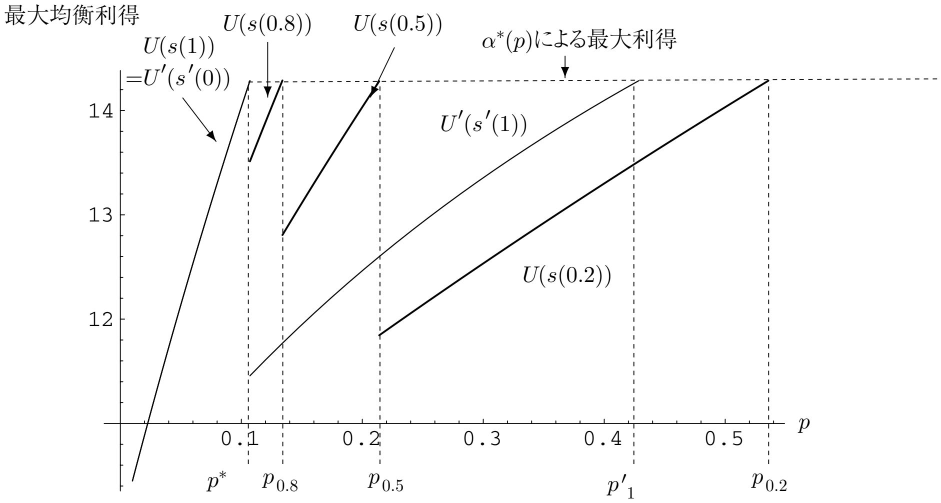
$$
\begin{aligned}
U'(s'(T)) &= \frac{1}{(1-\delta)\{1-\delta q(1-p)\}}\Big[(1-\delta q)(1-p)\underline{u} \\
&\qquad + p\Big\{d + (\underline{u} - d) + \frac{1-\delta q(1-p)}{\delta p q}\{\delta q(g - \underline{u}) - (g - c)\}\Big\}\Big] \\
&= \frac{\delta q g - (g - c)}{\delta q (1-\delta)}
\end{aligned}
$$

で与えられる。これはちょうど確率的受け入れ均衡の最大利得（9）と等しい。p'_{T-1} と p'_T の間の p については，（11）が満たされる限り

$$
\frac{\partial U'(s'(T))}{\partial p} = \frac{(1-\delta q)\{(1-\delta^T q^T)d + \delta^T q^T c - \underline{u}\}}{(1-\delta)\{1-\delta q(1-p)\}^2} > 0
$$

より，均衡利得は p の増加関数である。

図 6　最大利得の比較



5.　均衡利得の比較

以上の分析から，まず，任意の確率シグナルが存在する場合，確率的受け入れ均衡の利得は必ず漸次協力均衡の利得以上になることがわかる。この理由が離散時間による意思決定にあることは明らかである。漸次協力均衡が離散的にしか初期の非協力期間を調整できないのに対し，任意の確率シグナルで受け入れ確率を調整できるのであれば，確率的受け入れ均衡の方が効率的になる。ただし，出会いの確率が非常に低く，$p < p^*$ であるときは，新しい相手を見つけるのが難しいので別れることが十分な罰となるため，最適な均衡は，常に新しい相手を受け入れ，最初から協力し，もし相手が逸脱したら別れる，というものである。つまり，最適な確率的受け入れ均衡 $s(1)$ と最適な漸次協力均衡 $s'(0)$ は一致する。

確率的受け入れ均衡に使用できるシグナルが限られてくると，どちらの均衡利得が高いかはパラメターに依存する。しかし，シグナルが存在するのであれば，少なくとも 0.5 以上の確率をもつシグナルが一つは存在することから，そのシグナルを利用した確率的受け入れ均衡が $s'(1)$ を用いた漸次協力均衡より高い利得をもたらす十分条件を求めることができる。

命題 3：A が非空で $\delta q(c-d) > \underline{u} - d$ かつ

$$\delta q \leqq \frac{(g-\underline{u})(\underline{u}-d)+(c-\underline{u})(c-d)}{(g-\underline{u})(c-d)} \tag{13}$$

が成立するとき，$\alpha \in A$ と $p_\alpha \in (p^*, p'_1)$ が存在して，任意の $p \in (p^*, p_\alpha)$ について，全員が $s(\alpha)$ をする戦略の組み合わせが逐次均衡となり，しかもその利得は，この区間の p において漸次協力均衡

の中で最大利得を与える $s'(1)$ による均衡の利得以上になる。

証明：まず，$s(\alpha)$ の利得が $s'(1)$ の利得より小さくならないための α の条件を求めると

$$U(s(\alpha)) \geqq U'(s'(1)) \iff \alpha \geqq \frac{\delta q(c-d)-(\underline{u}-d)}{\delta pq(c-d)+(c-\underline{u})} =: \underline{\alpha}(p)$$

となる。明らかに $\underline{\alpha}$ は p の減少関数である。

次に，任意の $p \geqq p^*$ について

$$\alpha^*(p) \geqq \underline{\alpha}(p) \iff p \leqq p_1' \tag{14}$$

であることを示す。

p_1' をその定義から明示的に求めると

$$\begin{aligned}
& h(p_1') = (\underline{u}-d) + \frac{1-\delta q(1-p_1')}{\delta p_1' q}\{\delta q(g-\underline{u})-(g-c)\} = \delta q(c-d) \\
\iff & \delta p_1' q(\underline{u}-d) + \{1-\delta q(1-p_1')\}\{\delta q(g-\underline{u})-(g-c)\} = \delta^2 q^2 p_1'(c-d) \\
\iff & p_1'\{\delta q(\underline{u}-d) - \delta^2 q^2(c-d) + \delta^2 q^2(g-\underline{u}) - \delta q(g-d)\} = -(1-\delta q)\{\delta q(g-\underline{u})-(g-c)\} \\
\iff & p_1' = \frac{\delta q(g-\underline{u})-(g-c)}{\delta q\{(g-\underline{u})-(c-d)\}}
\end{aligned}$$

となる。

$B := \delta q(g-\underline{u})-(g-c)$ とおき，計算すると

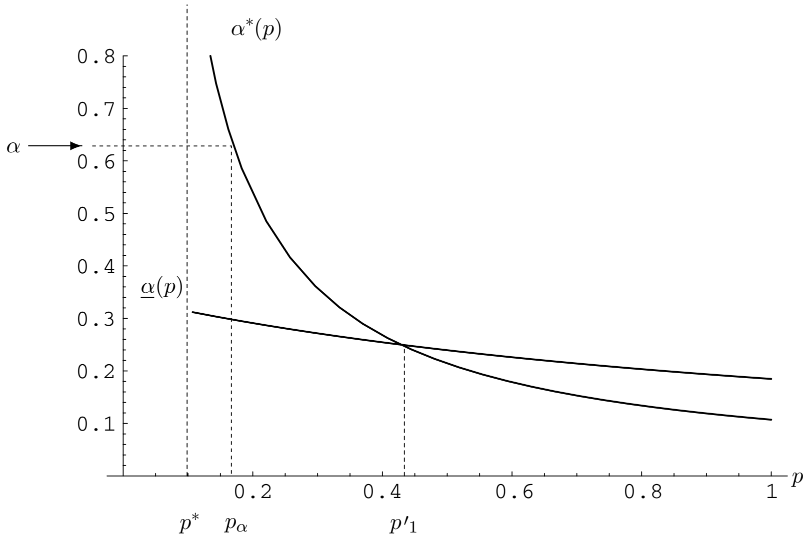
$$\begin{aligned}
& \alpha^*(p) \geqq \underline{\alpha}(p) \\
\iff & \frac{B}{\delta q(g-\underline{u})} \geqq p \cdot \frac{\delta q(c-d)-(\underline{u}-d)}{\delta pq(c-d)+(c-\underline{u})} \\
\iff & B\delta pq(c-d) + B(c-\underline{u}) \geqq \delta pq(g-\underline{u})\{\delta q(c-d)-(\underline{u}-d)\} \\
\iff & \delta pq\Big[(g-\underline{u})\{\delta q(c-d)-(\underline{u}-d)\} - B(c-d)\Big] \leqq B(c-\underline{u})
\end{aligned}$$

$$\begin{aligned}
\iff & \delta pq\Big[(g-\underline{u})\delta q(c-d) - (g-\underline{u})(\underline{u}-d) - \delta q(g-\underline{u})(c-d) + (g-c)(c-d)\Big] \leqq B(c-\underline{u}) \\
\iff & \delta pq\Big[\{g-\underline{u}) - (c-\underline{u})\}(c-d) - (g-\underline{u})(\underline{u}-d)\Big] \leqq B(c-\underline{u}) \\
\iff & \delta pq(c-\underline{u})\{(g-\underline{u})-(c-d)\} \leqq B(c-\underline{u}) \\
\iff & p \leqq \frac{\delta q(g-\underline{u})-(g-c)}{\delta q\{(g-\underline{u})-(c-d)\}} = p_1'
\end{aligned}$$

となるので（14）が示された。

第三に，（13）の下では $\underline{\alpha}(p^*) \leqq 0.5$ であることを示す。

図 7　α の存在



$$\begin{aligned}
&\underline{\alpha}(p^*) = \frac{(g-\underline{u})\{\delta q(c-d)-(\underline{u}-d)\}}{(g-\underline{u})\{\delta q(c-d)-(c-d)\}+(c-\underline{u})\{(g-\underline{u})+(c-d)\}} \leqq \frac{1}{2},\\
\Longleftrightarrow\ & 2(g-\underline{u})\{\delta q(c-d)-(\underline{u}-d)\} \leqq (g-\underline{u})\{\delta q(c-d)-(c-d)\}+(c-\underline{u})\{(g-\underline{u})+(c-d)\}\\
\Longleftrightarrow\ & \delta q(g-\underline{u})(c-d) \leqq (g-\underline{u})\{2(\underline{u}-d)-(c-d)+(c-\underline{u})\}+(c-\underline{u})(c-d)\\
\Longleftrightarrow\ & \delta q \leqq \frac{(g-\underline{u})(\underline{u}-d)+(c-\underline{u})(c-d)}{(g-\underline{u})(c-d)}.
\end{aligned}$$

A の中には 0.5 以上の α が存在するので，これを使用すると，上記より $\alpha \geqq 0.5 \geqq \underline{\alpha}(p^*)$ である。従って，任意の $p > p^*$ について $\alpha > \underline{\alpha}(p)$ となる。ここで，p_α を $\alpha = \alpha^*(p_\alpha)$ で定義すると，$\underline{\alpha}$ が減少関数であることと，$\underline{\alpha}(p'_1) = \alpha^*(p'_1)$ より，$p_\alpha < p'_1$ である。（図 7 参照。）p_α 以下の p について，$\alpha \leqq \alpha^*(p)$ が成り立つので $s(\alpha)$ は逐次均衡であり，しかも $\alpha > \underline{\alpha}(p)$ も成立するので，$s'(1)$ の利得以上が保証される。

最後に，$\delta q(c-d) > \underline{u} - d$ より，$s'(1)$ が $p \in (p^*, p'_1)$ については最大の均衡利得を与えることに注意すると，命題の主張が全て証明された。□

命題 3 の条件は複雑そうに見えるが，以下のような簡単な十分条件を求めることもできる。つまり，協力の利得 c が十分に大きければよいのである。

注意 1： $c - d \geqq g - \underline{u}$ ならば，任意の $\delta \in (0,1)$ と $q \in (0,1)$ について（13）が成立する。

証明：計算すると，

$$
\begin{aligned}
&\frac{(g-\underline{u})(\underline{u}-d)+(c-\underline{u})(c-d)}{(g-\underline{u})(c-d)} \geqq 1 \\
\Longleftrightarrow\ & (g-\underline{u})(\underline{u}-d)+(c-\underline{u})(c-d) \geqq (g-\underline{u})(c-d) \\
\Longleftrightarrow\ & (c-\underline{u})(c-d) \geqq (g-\underline{u})(c-d-\underline{u}+d) \\
\Longleftrightarrow\ & c-d \geqq g-\underline{u}
\end{aligned}
$$

となる。従って，$c-d \geqq g-\underline{u}$ が成立すれば（13）の右辺は 1 以上となり，δq は 1 より小さいので（13）が成立する。 □

このとき，（7），$\delta q(c-d) > \underline{u}-d$，（13）のすべてを満たす δq が存在することも言える。

図のパラメター値 ($g = 10$, $c = 6$, $\underline{u} = 2$, $d = 0$, $\delta = 0.8$, $q = 0.7$) においては，$\underline{\alpha}(p)$ は 0.31（$p = p^*$ のとき）と 0.25（$p = p'_1$ のとき）の間の値なので，$\alpha = 0.8$ と $\alpha = 0.5$ を用いた確率的受け入れ均衡は $s'(1)$ による均衡よりも利得が高くなっており，$\alpha = 0.2$ を用いた確率的受け入れ均衡は利得が低くなっていることがわかる。

6.　結論的覚え書き

6.1　贈り物交換モデルとの比較

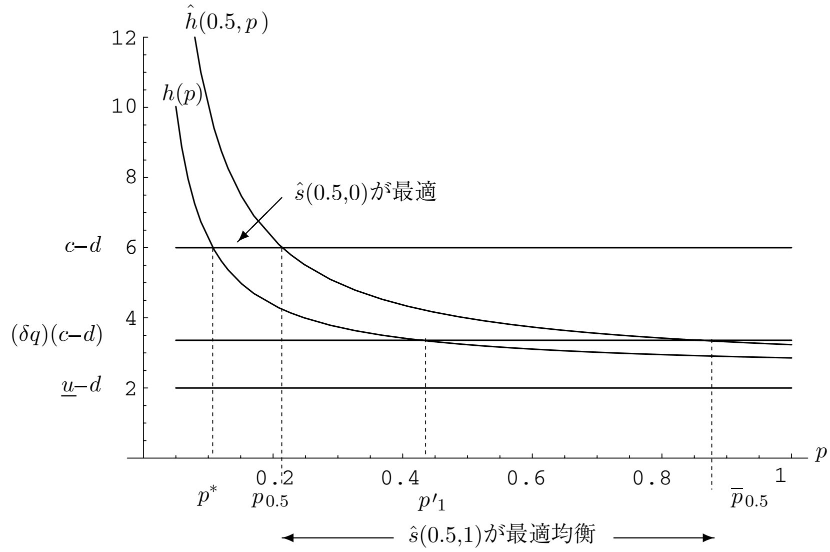
Carmichael and MacLeod (1997) は本稿のモデルと似たモデルを考察している。ただし，彼らのモデルでは出会いの確率は常に 1 であり，受け入れるかどうかの意思決定はないので，一つの関係が解消されると必ず新しい相手とゲームを始めなくてはならないとされている。そこで彼らは，新たな相手とプレイに入る直前に，贈り手にとっては費用がかかるが，受け手にとっては価値のない贈り物を交換するかどうかという意思決定を導入した。この贈り物の交換が，確率的受け入れと似たような働きをし，裏切って新しい相手とゲームを始めようとすると利得が下がるという構造になっている。贈り物にかかる費用を連続的に変化させることができるので，これまでの分析と同じ議論により，最適な贈り物交換均衡は最適な漸次協力均衡より効率的であると言える。

6.2　確率シグナルと信頼構築を併用する均衡

相関戦略に使える確率シグナルが限られているとき，確率的受け入れと漸次協力を同時に用いる戦略を考えることもできる。例えば以下のような戦略 $\hat{s}(\alpha, T)$ を全てのプレイヤーがすることを考えよう。

1. 確率 α で現れるシグナルが観察されたとき，またそのときのみ新しい相手を受け入れる。
2. パートナーシップが $t \leqq T$ 期間目である場合，D をプレイし，今期 (D, D) を観察したときだ

図８　確率的受け入れと漸次協力を共用することにより，効率性が高まる例



け継続の行動 m を選ぶ。

3. パートナーシップが $t>T$ 期間目である場合，C をプレイし，今期 (C,C) を観察したときだけ継続の行動 m を選ぶ。

このような $\hat{s}(\alpha,T)$ を全員がする戦略プロフィールが逐次均衡であるためには，4 節の分析で p を αp に変えた条件が成立すればよい。従って（11）は同じであり，（12）は

$$(\delta q)^T(c-d) \leqq (\underline{u}-d) + \frac{1-\delta q(1-\alpha p)}{\delta\alpha pq}\{\delta q(g-\underline{u})-(g-c)\} =: \hat{h}(\alpha,p) \tag{15}$$

となる。h が p の減少関数であったことから，$\hat{h}$ は α, p 両方について減少関数である。ゆえに α を小さくすれば，（15）は満たされやすくなり，T も小さくできるか，あるいは，より広い p の範囲で協力的な均衡を存在させられる。

例えば，これまでの全ての図に使用してきたパラメターの下で，$\alpha=0.5$ だけが使用可能であるとしてみると，図 6 より，p'_1 より大きい p の範囲では，協力的均衡は存在しないのに対し，$\hat{s}(0.5,1)$ を全員がとる戦略の組み合わせは均衡となる $p>p'_1$ の範囲がある。（図 8 参照。）したがって，確率的受け入れと漸次協力を併用することにより，効率性が高まると言える。

6.3　結　語

まず，本論文の結果の解釈を述べる。出会いの確率が低い $(p \leqq p^*)$ 場合，別れが十分な罰となるので，最適な確率的受け入れ均衡と漸次協力均衡は一致し，それは，新しい相手を常に受け入れ，最初から協力し，裏切られたら別れる，という戦略を全員が行うということである。日本の 1960 年

代，70 年代の労働市場では，このような均衡に似た状況になっていたと思われる。当時は新卒以外の雇用市場がほとんどなく，その代わり企業と労働者は互いに長期的な強い協力関係を築いていた。

マッチングがあまり難しくない状況になると，均衡利得は，出会いの確率と比較して，いかにうまく初期関係を調節できるかにかかってくる。p が p^* より大きくても，さほどではない場合，1 期間 (D, D) をしてしまうことは，出会った相手を少しの確率で受け入れない戦略より低い利得をもたらすことがあり得る。このように，離散時間であることは非常に影響が大きいが，実際，プレイヤーたちが一定期間ごとにしか行動を調整しないことはしばしばある。再び労働市場の例で考えると，賃金改定や昇進は臨機応変に行われているわけではなく，年 1 回など，ある程度の期間を置いて行われるのが普通である。もちろん本業の仕事の合間に業務評価，労使交渉等をしなくてはならないわけであるから，任意のタイミングで調節するのは難しいのであるが，逆に言うと，このような離散的意思決定にはコストがあるということがわかったのである。なかなか就職のチャンスがないにもかかわらず，新規採用には低い利得しか与えないという均衡より，就職のチャンスをさらに少し下げるとしても，最初から協力関係に入れればそれに越したことはない場合がある。このように，離散的意思決定と相関戦略との比較が明らかになったのは，本モデルの構造によるところが大きく，新たな貢献と言える。（通常の離散時間繰り返しゲームにおいては，相関戦略をとろうが，純戦略による適切な行動の列をとろうが，ほぼ同じ利得を達成できるので，両者はほぼ同じものとなる。）

さらに，完備情報ゲームにおいて[11]，漸次協力均衡が確率的受け入れ均衡より効率的であるのはどのような場合かを考えてみる。例えば，δq が非常に大きい場合，将来が重要となるので，d があまり小さくなければ（13）の逆が成立することがある。このとき，中間的な p の値について，漸次協力均衡の方が確率的受け入れより効率的となる。これは，プレイヤーたちが将来を重視するので，当初の (D, D) をがまんできるということである。

この他には，学習がある場合も確率的に受け入れるより，誰でも受け入れてしまう方がよいことが考えられる。（ただし，これはモデルを拡張しなくては正確な議論ではない。）もし，関係が長くなるとゲームをよりよくプレイできるようになり，両者の利得が大きくなるというような構造になっていれば，早く関係を成立させた方がよい。スポーツや職人の世界のように，技術の学習が存在する場合，やはり誰でもまず参加させて，ただし当初は「修行」させていることが多い。

最後に，残された課題について述べる。本稿では，対称戦略による均衡の比較しか行わなかったが，非対称戦略による均衡で，対称戦略均衡より高い利得をもたらすものが存在することが Fujiwara-Greve and Okuno-Fujiwara (2008) で示されている。均衡の比較はまだまだ可能であろう。さらに，モデルが複雑であることから，自発的繰り返し囚人のジレンマにおける均衡利得の全体の集合につ

（11） 相手の利得関数がわからない場合，それを知るために漸次協力をすることには意義があるので，不完備情報ゲームにおいては漸次協力は重要である。例えば Ghosh and Ray (1996) を参照。

いてはまだ一般的な分析がなされていない。これらは重要な将来の課題である。

（経済学部教授）

参 考 文 献

Abreu, D., D. Pearce, and E. Stachetti (1986). "Optimal Cartel Equilibria with Imperfect Monitoring". *Journal of Economic Theory*, **39** pp.251–269.

Aumann, R. (1981). "Survey of Repeated Games". In *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern*, Mannheim. Bibliographisches Institut.

Carmichael, L. and B. MacLeod (1997). "Gift Giving and the Evolution of Cooperation". *International Economic Review*, **38** pp.485–509.

Datta, S. (1996). "Building Trust". Manuscript. London School of Economics.

Ellison, G. (1994). "Cooperation in the Prisoner's Dilemma with Anonymous Random Matching". *Review of Economic Studies*, **61** pp.567–588.

Fudenberg, D. and E. Maskin (1986). "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information". *Econometrica*, **54** pp.533–554.

Fujiwara-Greve, T. (2002). "On Voluntary and Repeatable Partnerships under No Information Flow". In *Proceedings of the 2002 North American Summer Meetings of the Econometric Society* (http://www.dklevine.com/proceedings/game-theory.htm).

Fujiwara-Greve, T. and M. Okuno-Fujiwara (2008). "Voluntarily Separable Repeated Prisoner's Dilemma". Forthcoming in *the Review of Economic Studies*.

Fujiwara-Greve, T., M. Okuno-Fujiwara, and N. Suzuki (2008). "Voluntarily Separable Repeated Prisoner's Dilemma with Reference Letters". Manuscript, Keio University, University of Tokyo, and Komazawa University. Available at http://web.econ.keio.ac.jp/staff/takakofg/papers.html.

Ghosh, P. and D. Ray (1996). "Cooperation in Community Interaction without Information Flows". *Review of Economic Studies*, **63** pp.491–519.

Kandori, M. (1992). "Social Norms and Community Enforcement". *Review of Economic Studies*, **59** pp.63–80.

Kandori, M. and H. Matsushima (1998). "Private Observation, Communication and Collusion". *Econometrica*, **66** pp.627–652.

Kranton, R. (1996). "The Formation of Cooperative Relationships". *Journal of Law, Economics & Organization*, **12** pp.214–233.

Kreps, D. and R. Wilson (1982). "Sequential Equilibria". *Econometrica*, **50** pp.863–894.

Matsushima, H. (1990). "Long-Term Partnership in a Repeated Prisoner's Dilemma with Random Matching". *Economics Letters*, **34** pp.245–248.

Okuno-Fujiwara, M. and A. Postelwaite (1995). "Social Norms and Random Matching Games". *Games and Economic Behavior*, **9** pp.79–109.

岡田章（1997）．『ゲーム理論』，有斐閣。