奥野正寛教授（経済学部）最終講義 講義録

現代社会と行動様式の多様化

── 自発的継続繰り返し囚人のジレンマの分析を通じて ──

武蔵野大学政治経済研究所 2016 年度第 4 回研究フォーラム
2017 年 2 月 23 日 (木) 16 時 30 分～ 18 時
有明キャンパス 1 号館 207 教室

司会（大阿久博教授）　それでは時間になりましたので、奥野先生の最終講義を始めたいと思います。本日は天候が悪くなるようなことも予想されておりましたが、良いお天気で最終講義を迎えることが出来て大変良かったと思っております。

本日は大変遠方からお越しになった方もおいでになったとお聞きしております。誠にありがとうございます。

私は本日司会を務めさせていただきます経済学部経済学科の大阿久と申します。よろしくお願いします。私は不本意にも奥野先生の講義というものを一度もお聞きしたことがございませんでしたので、実は今日は非常に楽しみにまいりました。

それではお手元の式次第に沿って進めさせていただきたいと思います。はじめに、中村政治経済研究所長よりご挨拶がございます。

中村孝文所長　皆様こんにちは。今紹介いただきました政経研究所所長の中村と言います。よろしくお願いいたします。今司会の大阿久先生よりお話がありましたが、本日は遠方よりお出で下さった方もいらっしゃると伺っております。本当にありがとうございます。またいつもより沢山の所員の方々にご出席いただきまして本当にありがとうございます。

奥野先生のご略歴を紹介させていただきたいと思います。奥野正寛先生は 1969 年に東京大学の経済学部を卒業されました。その後アメリカの

スタンフォード大学大学院に留学され、1974年にPh.Dを取得されました。Ph.D取得後はペンシルベニア大学の経済学部客員講師、それからイリノイ大学経済学部助教授、そしてアメリカで教鞭をとられておられます。1977年に日本に帰国されまして、横浜国立大学の経済学部助教授、それから東京大学経済学部助教授を経まして1989年に東京大学大学院経済学経済科、経済学部教授にご就任になられました。2010年に東京大学をご定年になられて退職をされ、東京大学名誉教授の称号を授与されております。その後流通経済大学経済学部の教授をされまして、2013年の4月から本学の経済学部の教授にご就任されておられます。本年の3月31日をもちまして定年退職というご予定になっておられます。退職後は本学の政治経済学研究所の客員教授にご就任いただくという運びになっております。

奥野先生は皆様もご承知の通り、日本を代表する理論経済学者でいらっしゃいます。特にゲーム理論におきまして世界の第一線で活躍する研究者でいらっしゃいます。日本のみならずアメリカをはじめ多くの国々に名前を知られていらっしゃいます。これまで日本経済学会、日本応用経済学会の会長を務められ、経済学者が作る国際的な経済学会であるEconometric Societyの終身特別会員にも選出されておられます。このように先生の業績は日本だけではなく世界に広く知られており、経済学の発展に多大な貢献をなさっておられ今日に至っておられます。

1988年には先生のご著書であられます『産業政策の経済分析』が日経・経済図書文化賞を受賞されました。それから政府のさまざまな委員会の委員もお務めになられていまして、政策において経済学を実践する立場でもいらっしゃいます。

本学着任後の4年間は公共経済学、ゲーム理論というような授業をご担当されて、学生のみならず大学院生の指導にもあたっていただきました。

また2013年は法学部と経済学部が出来た年ですが、この年に開かれた法学部・経済学部開設記念シンポジウム、それから昨年度の政治経済研究所の講演会ではそれぞれご登壇いただきまして、研究所の活動にも大変ご

尽力をいただいてきました。それからこの大学の大学院の政治経済学研究科博士後期課程の設置の時にも先生に多大のお力添えをいただいております。

この博士課程設置の際の先生のご親切なアドバイスに対し、この場を借りまして感謝を申し上げまして、先生のご紹介と開会の挨拶とさせていただきます。先生どうもありがとうございました。

司会　ありがとうございました。今の話を聞きまして奥野先生には大変お世話になったことをあらためて感じました。それではさっそく奥野先生の最終講義の方に移らせていただきたいと思います。

タイトルは『現代社会と行動様式の多様化──自発的継続繰り返し囚人のジレンマの分析を通じて──』です。よろしくお願いいたします。

奥野教授　中村先生、ご丁寧な紹介、過分なお言葉ありがとうございました。そしてこんなにたくさんの方に来ていただいて大変光栄でございます。では、私の最終講義を始めさせていただきます。

始めに

今回の講演のタイトルに、「現代社会と行動様式の多様化」と書きましたが、私自身多様性が大事だと思っていて、それを「自発的継続繰り返し囚人のジレンマ」というゲームを使って研究するという作業を、この10年間くらい続けてきました。

「囚人のジレンマ」という有名なゲームがあります。それから「繰り返し囚人のジレンマゲーム」というゲームが、ゲーム理論の世界では30年以上前に大きな脚光を浴びました。それらをさらに拡張して、自発的に継続するか自発的に解消するかをプレイヤー自身が自主的に選択できる「自発的継続繰り返し囚人のジレンマ」というゲームを考えまして、ここ10年間ほど主に慶応大学のグレーヴァ香子さんと研究してきました。駒澤大学の鈴木伸枝先生と、最近では信州大学の西村直子先生を加えて、2人な

いしは 4 人の共同研究をしています。そこで、この 10 年間分の研究の結果[1] をザッとお話させていただきたいと思います。

囚人のジレンマ

経済取引というのは実はほとんどの場合、ある種のモラルハザード、つまり機会主義的行動が起こる可能性を含んでいます。例えば労働契約を例に取ってみましょう。労働契約自体は、雇用者が労働機会を提供する。それに対して労働者側は労働サービスを提供する。それで、お互いが協力をすることで利益を得よう、取引機会を実現しようという約束です。ところが、雇用者側からすると労働者側に働かせた上で逆に賃金を払わずに済めば、もっと儲かるわけです。そうしたいというインセンティブ、つまり「賃金不払いのインセンティブ」を雇用者は持っています。それに対して従業員の方も、働かないでサボって賃金だけを手にしたい。働いて協力する約束なのに、働かないで賃金だけを手にしたらもっと良いという、「サボリのインセンティブ」を持っています。両者が協力する（Cooperate する）ことが望ましいのに、お互いにモラルハザードがあり、賃金未払いをしたい、サボりたいインセンティブを持っているわけです。

これは、実は労働契約だけでなく、信用契約とか共同作業などを含めて、経済的取引はほとんどみんなそういう構造を持っています。協力すると大きな成果を得られるが、自分が協力しない方が自分の利益になる。「協力しない」ことを「裏切る（Deviate する）」とこれから言います。裏切ることによって、相手は損するかもしれないが、自分はもっと大きな利益を得るというこの構造は、経済取引の持つ一般的特性なのです。信用取引で返済をなしで済ます、商取引で支払いをしないで済ます、品質向上や生産増大の努力を惜しんで自分の利益を得たい。経済的取引が普遍的に

1　なお、この間の研究に対しては、JSPS 科研費 JP24330064, JP21530171, JP19530158 の助成を受けました。

持っているそういう構造を持つゲームを、「囚人のジレンマ」と言います。

図１が、「戦略形」という表現方法で表わした「囚人のジレンマ」のゲームです。どうして囚人のジレンマと言うのかについては、司法取引を基にした、二人の容疑者の行動とその結末という面白い話があるのですが、これを説明し出すと時間がなくなります。関心のある方は、例えばゲーム理論の入門書である渡辺隆裕著『ゼミナールゲーム理論入門』などをお読みになって頂ければと思います。

	協力（C）	裏切り（D）
協力　（C）	c, c	ℓ, g
裏切り（D）	g, ℓ	d, d

$$g > c > d > \ell$$
$$2c > g + \ell$$

図１　囚人のジレンマ

さてこの図は、二人の経済取引関係者で行うゲームです。この二人を以下、「行プレイヤー」と「列プレイヤー」と呼びます。行プレイヤーは「協力の行を選ぶ」か「裏切りの行を選ぶ」かという選択肢を持っています。列プレイヤーも「協力の列を選ぶ」か「裏切りの列を選ぶ」かという選択肢を持っています。そういう選択肢のことを「戦略」と言います。行プレイヤーが協力（C）を選んで、列プレイヤーも協力（C）をすると、それぞれのプレイヤーがcという利得（利益）を得ます。これは協力（cooperation）の利得と言われています。お互いが裏切る（D）とdという利得が得られます。二つ並んでいる利得のうち、左側が行プレイヤーの利得、右側が列プレイヤーの利得を表しています。cの方がdよりも大きいという関係があるので、協力した時の利得の方が、お互いが裏切った利得より大きいことになります。

ただ相手の列プレイヤーが協力をしてくれるなら、行プレイヤーは協力

をして c という利得しか得られないより、裏切って g の利得を得た方がより大きな利益（gain）を得られます。だから「裏切りたい」というインセンティブを持つわけです。実はこれは相手が協力してくる場合だけではなく、相手が裏切るという場合も同じです。行プレイヤーは、自分は協力しているのに相手が裏切ると、左側の ℓ というより小さな利得しか（loss）得られません。それに対して、自分が裏切ると d だけの利得を得られます。d の方が ℓ よりも大きいので、相手が裏切るときも、自分も裏切りたいというインセンティブを持っています。

具体的な数字を挙げた例が、図 2 です。お互いが協力すればお互いが 3 の利益が得られて、お互いが協力しないで裏切った場合の 1 の利得よりは利益が多いが、相手が協力しているときに自分が裏切ると、4 という大きな利得が得られます。お互いが協力した場合の 3 に比べて、1 だけ多い利得（deviation gain）が得られるわけです。相手が裏切るとき、自分も裏切れば 1 の利得が得られるが、自分が協力すると 0 の利得しか得られない。協力するとかえって利得が減る（loss）わけです。ゲームがこういう構造を持っているのが、囚人のジレンマのゲームです。

	C	D
C	3, 3	0, 4
D	4, 0	1, 1

図 2　囚人のジレンマ・ゲーム（数値例）

このゲームを一回限りプレイすると、相手が協力を選んでも裏切りを選んでも、どちらの場合でも、自分は裏切りを選んだ方がより大きな利得が得られます。相手がどちらを選んでも自分は常に裏切りを選んだ方が得ですから、裏切るという戦略が協力という戦略を「支配している」という意味で、裏切りを「支配戦略」と呼びます。結果として、このゲームを一回だけプレイすると、両方のプレイヤーが支配戦略を選びますから、(D, D)

が「支配戦略均衡」という形で実現することになります。これが、ゲーム理論ができた当初から良く知られている、囚人のジレンマの構造です。

繰返し囚人のジレンマ

次に使うのは、「ナッシュ均衡」という概念です。ナッシュ均衡とは、その戦略のペアが実現するだろうとプレイヤー同士が予想しあっているとき、予想通りの戦略を実現するのがお互いに最適になっている状態です。支配戦略均衡は、それが存在すれば必ずナッシュ均衡にもなります。どんなゲームでも支配均衡が存在するわけではありませんが、ナッシュ均衡なら、どんなゲームにも必ず存在することが知られています。

さて、現実の経済関係の多くは、基本的に囚人のジレンマのゲームという性質を持っていました。しかし、一回限りの囚人のジレンマでは、お互いが裏切ることを選ぶことが支配戦略均衡であるために、協力は実現できないはずだというのが、理論の予測です。しかし、現実の世界では結構みんな協力をします。では、囚人のジレンマのゲームが通常の経済取引を表しているとすると、どうして支配戦略均衡ではない協力が実現するのだろうか、という疑問が浮かびます。逆に言えば、(C,C) を選ぶことが双方に有益ですから、それを実現する方法はないだろうか、ということが問題になります。これはゲーム理論では、ずっと昔から問題になってきました。

その解決策として、30 年くらい前に完成された考え方が「繰り返し囚人のジレンマ (Repeated Prisoner's Dilemma;RPD)」です。この考え方自体は、もっと前から知られていましたが、それがきちんとした理論論文として完成されたのが、30 年ぐらい前です。RPD の考え方は、次のようなものです。

現実の経済取引は、確かに経済的な取引機会なので、囚人のジレンマ (PD) 的なゲームの構造を持っていますが、現実の世界では、PD を一回限りプレイするのではなくて、毎日毎日、毎月毎月経済取引 (PD) を繰り返すという構造を持っています。そういう意味で、現実にプレイされている囚人のジレンマは「一回限りの囚人のジレンマ」ではなく、「繰り返し囚人

のジレンマ」、RPD なのだということに、人々はだんだん気が付いたわけです。

もう少し厳密にいうと、二人繰り返し囚人ジレンマのゲームを、ある特定の二人のプレイヤーが、毎期ごとに一回限りの PD ゲームを、繰り返しプレイするというゲームとして定義します。毎期プレイする PD ゲームのことを「段階ゲーム」という言い方をします。段階ゲームを毎期繰り返しプレイすることで、毎期毎期の「利得の流列」を得ます。例えば今日の段階ゲームでは（C,C）が実現し、自分は c を得ます。明日の段階ゲームでは（D,C）が実現し、自分は g を得ます。明後日の段階ゲームでは（C,D）が実現し、自分は ℓ を得ます、結果として（$c, g, \ell, \cdots$）という利得流列を得ます。というように、毎日毎日、いろいろな利得が実現していくわけです。その利得流列の現在割引価値を最大にするように、プレイヤーたちがこの時間を通じたゲームをプレイすると考えるのが、繰り返し PD ゲームです。

割引係数、明日の１円を今日の段階でいくらと評価するかという値を δ と定義します。割り引くのが利子率だと考えれば、利子率は $\frac{1}{\delta}-1$ だと考えることもできます。繰り返し PD ゲームでは、次の二つの条件が満たされれば協力が実現されます。協力が実現するとは、毎期毎期の段階ゲームで（C,C）が実現するということです。そのための第一の条件とは、適当な大きさの制裁を与える仕組みが存在する、ということです。要するに、（C,C）をずっと続けましょうという約束をしておいて、どちらかが裏切ると、それに対して裏切られた人は裏切った人に対して十分な大きさの制裁を与えることができる。その仕組みが、この繰り返し PD ゲームの中に内包されているというのが、第一の条件です。第二の条件とは、両プレイヤーが十分に将来を大事にすること、つまり、あまり近視眼的ではないということです。将来が大事だと考えるとは、δ という割引係数が十分１に近い、あるいは利子率を考えれば利子率が十分０に近い場合です。これら二つの条件が満たされる場合、繰り返し PD ゲームで（C,C）が実現されることがわかったわけです。まずはそれを簡単に説明しましょう。

繰り返しPDゲームの戦略

次に、繰り返しPDゲームの「戦略」を説明しましょう。繰り返しPDゲームのような、時間を通じた複雑なゲームにおける戦略とは、そのゲームをプレイする前に予め、どのようにこのゲームをプレイしようかと考えておく計画のことです。繰り返しPDゲームをプレイする典型的な戦略が「C−トリガー戦略」です。C−トリガー戦略では、第一期目の段階ゲームではCをプレイします。問題はその後で、T期目ならT期目にCをプレイするかDをプレイするかを、T期目までに何が起こったかに応じて予め決めておきます。C−トリガー戦略の場合、T期目より前には、自分も相手も共にずっとCをプレイし続けている場合には、自分はT期目にCをプレイします。しかし、T期目より前に、自分か相手が一度でもDをプレイした場合には、自分はT期目にDをプレイします。これがC−トリガー戦略です。どちらかが一度でもDをプレイした、つまり、「どちらかが一度でも裏切ったら、その後は自分はずっと裏切り続ける」という戦略です。そういう意味で誰かが裏切ると、それがトリガー（引き金）を引き、その後は自分はDをプレイし続ける。それが、先ほど言った「相手に対する制裁」の役割を果たすという戦略です。

もうひとつの繰り返しPDゲームの典型的な戦略が、グリムD−戦略です。これは単純な戦略で、「何が起ころうともすべての段階ゲームでDをプレイする」という戦略です。裏切り続けるわけです。二人のプレイヤーがお互いにグリムDを戦略として選び合う場合を、グリムDとグリムDの組み合わせといいますが、この戦略の組み合わせはナッシュ均衡になります。その場合、毎期、毎期、段階ゲームではお互いの支配戦略であるDをプレイするので、利得流列は（$d,d,d,\cdots$）になります。ある段階ゲームで自分がそれと違うことをしても、その結果得られるのは、ℓという、もっと低い利得です。しかもそれ以降の段階ゲームで、相手の行動が変わるわけではありません。それなら最低限dを確保し続けた方が良いとい

うインセンティブが生まれますから、グリム D の組み合わせはナッシュ均衡になります。得られる利得の現在現在価値は、毎期 d の利得流列の割引現在価値ですから、それを δ で割り引くと $d/(1-\delta)$ になります。

繰返し PD ゲームにおける協力の実現

繰返し PD ゲームで、もしプレイヤーが将来を十分大切にするならば、お互いが C−トリガー戦略を選択することがナッシュ均衡になります。以下、それを説明しましょう。

いま、お互いが C−トリガー戦略を選ぼうという約束をしている状況を考えてみましょう。まず、相手も自分も約束にしたがって C−トリガーをプレイする場合を考えてみます。この場合、1 期目は二人とも C を選ぶ訳ですから（C,C）という結果が実現し、二期目も二人とも C をプレイして（C,C）という結果が実現し、……という形で、毎期（C,C）が実現します。その結果、両者が得る利得流列は（$c,c,\cdots$）になり、その割引現在価値は $c/(1-\delta)$ になります。

他方、相手が C−トリガーを選択しているときに、自分が約束を破って、C−トリガー以外の戦略を選択したら、自分の利得はどうなるでしょうか。任意の第 T 期目でも良いのですが、基本的に同じことなので一番簡単な場合として、C−トリガーをプレイする代わりに、1 期目の段階ゲームで D をプレイする場合を考えてみましょう。1 期目の段階ゲームで D をプレイする、要するに 1 期目に自分が裏切るわけですから、1 期目の段階ゲームの結果は（D,C）になり、1 期目には g の利得を得ます。ただ自分が裏切るので、相手のトリガーを引いてしまいます。その結果、相手は 2 期目以降のすべての段階ゲームで D をプレイし続けることになります。そうすると、自分にできる最善のことは自分もずっと D をプレイし続けることなので、2 期目以降に得られる利得の流列は（$d,d,d,\cdots$）になります。

つまり、1 期目に D をプレイした場合、自分にとって得られるベストの利得流列は、（$g,d,d,d,\cdots$）になります。従って、1 期目に g を得て、2 期目

以降は、来期のお金で評価して $d/(1-\delta)$ の割引現在価値が得られます。ただそれは、来期のお金での評価額なので、今期のお金で評価した割引現在価値にすれば、2 期目以降に得られるのは、それを δ 倍した $\delta d/(1-\delta)$ になります。結果として、約束を裏切るインセンティブが無いという条件、つまり約束通り C−トリガー戦略をプレイし続けたほうが良いという条件は、約束通りにプレイして $c/(1-\delta)$ を得た方が、約束を破って $g+\delta d/(1-\delta)$ を得るより望ましいという条件になります。式で書けば、

$$c/(1-\delta) \geq g+\delta d/(1-\delta)$$

になります。これを「インセンティブ・コンパティビリティ条件」略して IC 条件と呼びます。この式を整理すると、$\delta \geq \frac{g-c}{g-d}$ という条件が得られます。この式の右辺は、PD ゲームであるための条件から、必ずゼロと 1 の間の数になります。したがって、IC 条件とは、δ の値が十分に 1 に近く、$(g-c)/(g-d)$ を上回る、つまり、プレイヤーが将来を十分に大切にすることを意味しています。

以上をまとめると、お互い多 C−トリガー戦略をプレイするという状態は、次に述べる理由でナッシュ均衡になります。もし C−トリガー戦略にしたがって、お互いが C をプレイし続ければ、相手が自分を信頼して平均利得、つまり毎期段階ゲームから得られる利得は c になります。しかし、一度でも D をプレイして相手を裏切ると、相手は自分に対する信頼を失います。その結果、確かにその期は裏切ったことによって、c より大きい g という利益を得られますが、信頼を失ったことで相手から制裁されます。結果として翌期以降に得られる平均利得は c より低い d に低下してしまいます。つまり、D をプレイすると相手の信頼を失い、相手から制裁されて平均利得が低下するわけです。将来がそれなりに大事だと考えいる（δ の値が十分に 1 に近い）プレイヤーは、この制裁が十分に大きいと感じるので、相手を裏切らずに、C をプレイし続けることで制裁を避けたいと思うはずです。つまり、相手の信頼を失う恐れと、その結果起こる制

裁が強力なインセンティブを作りだす。それが起こるための条件は、将来を十分に大事にしていることなのだ、というわけです。

フォーク定理

以上は、お互いがC−トリガー戦略をプレイする状態が、繰り返しPDゲームのナッシュ均衡になるという話でした。しかし実は、この結論は、もっと一般的に成立する話であり、それは「フォーク定理」と呼ばれています。せっかくの機会ですから、簡単にフォーク定理を説明しておきましょう。

お互いがグリムD戦略をプレイする状況は、繰り返しPDゲームのナッシュ均衡でした。この場合、各段階ゲームで一回限りの囚人のジレンマゲームのナッシュ均衡をプレイするので、各プレイヤーは毎期毎期 d の利得を得続け、実現される平均利得の組は（d,d）になります。また、十分に δ の値が大きければ、お互いがC−トリガー戦略をプレイする場合も繰り返しPDゲームのナッシュ均衡になり、毎期毎期 c が実現する訳ですから、実現する平均利得の組は（c,c）になります。繰り返しPDゲームでは、実は他にもさまざまな戦略の組がナッシュ均衡になります。フォーク定理というのは、すべてのプレイヤーが将来が十分に大切だと考えるなら、どんな個人合理的な平均利得の組も、ナッシュ均衡として実現可能だ、という定理です。ここで個人合理的とは、次のようなことを意味しています。繰り返しPDゲームでは、相手がどんな戦略を選んでも、自分はDをプレイし続けることで、各プレイヤーは毎期 d の利得を確保できます。つまり、d とは、自分だけの独力で確保できる最低限の平均利得です。個人合理的な平均利得の組とは、両方のプレイヤーが、この独力で確保できる最低限以上の、実現可能な利得を得ている状態のことです。繰り返しPDゲームでは、どんな平均利得の組であっても、それが個人合理的で実現可能であり、しかも δ の値さえ十分に大きければ、その組を、繰り返しPDゲームのナッシュ均衡が実現する平均利得の組として実現できます。これが「フォーク定理」です。

繰返しPDゲームの限界と評判を通じた制裁

以上に述べたように、繰り返しPDゲームの社会では、もしプレイヤーたちがお互いを信頼するならば、その協力によって経済的利益が実現されます。お互いを信頼するのは、どちらかが相手を裏切って非協力（D）を選べば、そのプレイヤーに社会的制裁が加えられる、という制裁の仕組みがあるからです。これはプレイヤーが二人という1対1の状況だけじでなく、もっと大きな社会でも成立します。社会には、信頼を担保するさまざまな社会的な仕組みが存在するのです。

ただ、以上の論理を使うには、すでに述べた、(1）制裁する仕組みが存在する、(2）将来を十分に大事にする、ということに加えて、あと二つ重要な制約があります。まず、「繰り返しPDゲームの論理」が適用できるのは、二人より多くても良いが、少数の関係者間の固定的、永続的な関係であることです。少数の当事者の固定された関係でないと、この繰り返しPDゲームの論理は使えません。言い換えると、小さな共同体の中での慣習とか規範、典型的には「誰かが悪いことをすると村八分にされて内部制裁をされる」といった状況だったら、繰り返しPDゲームの論理を使えます。また、企業組織の中で命令違反が起きると、それに対する内部制裁が企業規則とか統制によって行なわれる。こういう場合も同じです。

ただ、古代から中世、近代社会から現代社会へと発展してくるにつれて、人々の移動費用が低下し、共同体や国家・社会の間の流動性が増大してきました。従って、パートナーの信頼を失っても、彼／彼女との関係を解消することで制裁を回避できる可能性が生まれてきました。相手を裏切っても、相手から逃げて別の相手と新たな関係を築けば、裏切った相手から制裁を受けることはない、というわけです。

こう考えれば、社会の流動性が高まれば、繰り返しゲームを考えても協力を実現できないかもしれません。ただ、次のような可能性も考える必要があります。現在の相手を裏切っても、彼／彼女との経済関係を解消して

逃げてしまったとき、次の経済機会はありますが、新たな経済取引の相手から信頼を勝ち得ることができるかというと、それは、必ずしも明らかではありませ。なぜなら、昔の相手との契約や信頼を破ったことがあるという評判が新しいパートナーに伝われば、新しいパートナーから信頼を勝ち取ることはできないからです。過去の行動履歴が新たなパートナーに伝わるなら、「評判」という仕組みを通じて、第三者を通じた制裁、裏切った当人の昔の相手からは制裁されないが、次の新しいパートナーという第三者を通じた制裁が与えられることになります。そういう意味で、流動化社会でも評判を考えれば、裏切った結果としての「悪い評判」が制裁の仕組みとして機能する可能性があるのです。

情報化社会と匿名性

ただ、現代社会は流動化しているだけではなく、IT 化や匿名化しています。現代社会では、ビッグデータの活用に典型的に表れているように、IT 社会化に伴って情報が洪水のように押し寄せていま。また、このような大量の情報の陰に隠れて、個人の行動はますます匿名化しています。たとえば、ある人の過去の行動履歴を把握しようとしても、よほど大きな事件を起こしでもしない限り、その人が国境を越えて当事国を脱出し、新しい人格で生きようとすれば、行動履歴を把握するコストはどんどん大きくなってきています。取引機会が世界に広がっていくことによって、世界の果ての聞いたこともない人から取引を持ちこまれた時に、その人の過去の行動履歴を把握することは困難です。そういう意味で、個々の法人、個人の評判は新しいパートナーに届きにくくなって、第三者による制裁が機能しにくくなっています。現在のパートナーシップを解消しても、次のパートナーが大きなコストをかけないかぎり、この悪い評判はなかなか新しいパートナーに伝わらないため、第三者による制裁は回避されてしまう可能性が次第に増えてきています。それを以下では、「匿名化社会」という言葉で表しましょう。匿名化社会では、信頼を作りだす社会的仕組みは、今

までの繰り返しPDのロジックでは説明できません。

では、匿名化社会における社会的仕組みは存在しないのか、するとしたらどんな仕組みかというのが、私たちがこの10年ほど研究している話です。それを考えるためには、実は「社会ゲーム」という視点から考える必要があります。社会ゲームは、東京大学の松井彰彦先生が言いだして、私たちも使っている概念です。社会ゲームとはどういう概念かと言いますと、国家とか大都市といった、非常に大きな社会を考えます。その大きな社会で、人々は次のような過程を経て、戦略的に行動します。まず、生まれて新しいパートナーと出会い、彼らとさまざまな付き合いをします。時には付き合いを継続しますが、ときには裏切られたりしてパートナーを解消します。パートナーを解消した場合は、別の新しいパートナーと出会い直して新しい付き合いを始めます。こういうことが、次々と起きる。その過程を全体として動学的に把握したものを、社会ゲームと呼びます。社会ゲームの一つの典型として、今日私がお話しようとしているのが、「自発的継続繰り返し囚人のジレンマ（VSRPD）」です。

通常の繰り返し囚人のジレンマではパートナーを変えることができないわけですが、自発的継続繰り返し囚人のジレンマ（VSRPD; Voluntarily Separable Repeated Prisoner's Dilemma）は、パートナーを変えられます。VSRPDは、私がグレーヴァさんと一緒に、2009年に『Review of Economic Studies』誌に掲載した論文で初めて体系的に定式化したゲームです。そういうロジックがあることは10年前くらいから知られていたのですが、それをきちんと体系的に分析しましょうということを始めたのが、この論文です。その後、鈴木さんなんかも含めて、『Games and Economic Behavior』誌に載せた論文とか、『Economic Theory』誌に載せた論文をはじめとして、ここ10年越しくらいでやっている研究プロジェクトの結果を、これから簡単にご紹介しようと思います。

自発的継続繰り返し囚人のジレンマ（VSRPD）

VSRPDとは、次のようなゲームです。社会では、昔からずっと、過去から現在、将来にわたって、図3のような状態がプレイされています。この大きな社会の中で、ある期に何が起こるかというと、その期の初めにパートナーがいない人は、ランダム・マッチング・プールというところに行って、そこで相手を見つける。結果として、二人のプレイヤーがパートナーシップを形成します。その後で二人は、段階ゲームとして通常の一回限りの「囚人のジレンマ（PD）」ゲームをプレイします。PDゲームの結果を見た上で、彼らは現行のパートナーシップを「継続（keep）する」か、「解消（end）する」かを決めます。もし二人が、ともに「継続（k）する」を選べば、次の期も同じパートナー同士でPDゲームをプレイします。繰り返し囚人のジレンマをプレイすることになるわけです。逆に、もしどちらかのプレイヤーがペアを「解消（e）する」を選べば、そのパートナーシップは解消します。ペアを解消したプレイヤーたちはランダム・マッチング・プールに戻って、新しいパートナーと無作為に出会って、新しいペアを形成する。その場合、過去の行動履歴は、匿名化社会だから新たな相手には引き継がれない。過去の行動履歴を知ることのコストがあまりに高いという状況を前提にします。結果として、新たなパートナーは相手の過去の行動履歴を知らないで、VSRPDをプレイすることになります。

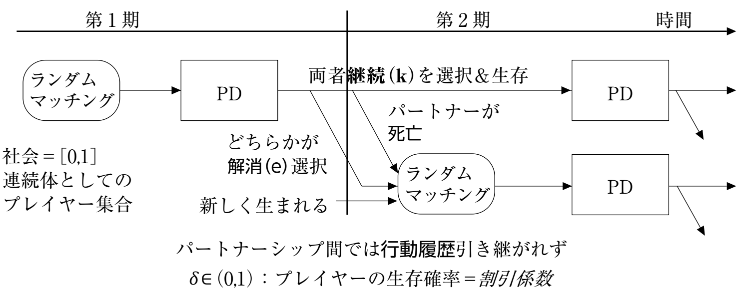

図3　VSRPDモデル

以上をビジュアルに考えて頂くと、図3のようなイメージになります。歴史はもっとずっと過去から続いていますが、例えば、今日が第1期で明日が第2期と、時間が左から右に流れていく状況を考えます。社会とは非常に大きな社会で、数学モデルなので少し極端に考えて、無限の数の人間がいる社会を考えます。連続体（continuum of players）といわれる、無限の数のプレイヤーの集合からなる社会を考えます。ある期の初めに、相手のパートナーがいる人もいない人もいるわけですが、パートナーがいない人はランダム・マッチング・プールに行きます。ここにも無限の数の人がいるわけですが、その人たちが無作為に二人のペアを作ります。新たに作られたペアも、過去から引き続いてペアを形成しているペアも、ペアを形成するパートナー同士で、一回限りの囚人のジレンマ（PD）ゲームをプレイします。プレイした結果を見て、その期の終わりに、各プレイヤーは、ペア継続（keep）と解消（end）のどちらを選ぶかを決めます。両者が継続（k）を選び、しかもどちらも死ななければ、次の期にも同じパートナー同士でPDゲームをプレイします。もしどちらかがその期のPDゲームの結果を見て解消（e）を選ぶと、そのパートナーシップは解消され、次の期には、生きている限り二人ともランダム・マッチング・プールに行きます。

両者が継続（k）を選んだ場合でも、自分や相手が死ぬ可能性があります。その確率は $1-\delta$ です。逆に δ は、各プレイヤーが次の期まで生き続ける確率で、割引係数にあたります。$1-\delta$ の確率で自分が死ぬ、あるいは相手が死ぬということが起こります。したがって、パートナー同士がペア継続（k）を選んだとしても、そのペアが次の期まで継続する確率は、δ^2 です。パートナーが死んだ人も、残された相手はランダム・マッチング・プールに戻ります。死んだ人と同じ数の新しく生まれて来る人がいるので、社会全体の人口は変わりません。マッチング・プールでは、前の期にペアを解消した人、解消された人、前の期の終わりにパートナーが死んだ人、新しく生まれてきた人、そういうさまざまな人々が出会う場です。こ

れらの人々が、ランダムに二人ずつのペアになってPDゲームを新たにプレイします。その結果を見て両者が継続か解消かの決断をする、……という形で、社会がずっと続いていきます……。というのが、このVSRPDのモデルです。パートナーシップ間では行動履歴は引き継がれません。こういう状況は、現実の経済取引でしばしば現れます。人々が経済取引を続けたり、何らかの理由で取引を止めて、取引の相手を変えるというケースもあります。研究者が共同論文を書いているとき、同じ人とずっと共同論文を書くことも、止めて別の研究者と共同論文を書くこともあります。結婚も、続けることもあるし、離婚したりパートナーと死別したりして、新しいパートナーと結婚するということもあります。このように、VSRPDのモデルは、現実にも当てはまる例がたくさん考えられるゲームです。

進化ゲーム

VSRPDのモデルの基本的に重要な点、普通の繰り返しPDとの違いは、プレイヤーが一方的にパートナーシップを解消できることです。他方、解消することは最大の制裁で、それ以上個人的な制裁はできません。何らかの社会的な仕組みがあれば別ですが、個人的にできる最大限の制裁は、自分とのパートナーシップを解消することでしかないという点です。

さて、VSRPDを分析するとき、限定合理性しか持たないプレイヤーからなる社会を考えます。「限定合理性」とはどういう意味かというと、普通のゲーム理論では、プレイヤーは「超合理的」な人たちで、ゲームの構造や相手の戦略を熟知しており、どんな複雑なゲームでも即座に最適な戦略を計算し実行できる存在である、と考えます。しかし、VSRPDを分析する際に使うのは、「進化ゲーム」という分析枠組みです。進化ゲームは、実は生物学で始まったゲームの考え方で、それがその後、経済学に導入されたという経緯を持つ分析枠組みです。そのため、進化ゲームでは、ゲームをプレイするプレイヤーは、通常のゲーム理論で仮定される超合理的な存在ではなく、できるだけ合理的にプレイしたいとは思っているが、限ら

れた能力しかないため最適な戦略を計算しようとしても、それができないという「限定合理的」な存在だと考えて、そこでの安定な均衡を考えます。社会という多数のプレイヤーが相互依存するゲームを分析対象とする社会ゲームにとって、プレイヤーの超合理性の仮定を避け、限定合理性だけを仮定することは極めて自然なことでしょう。

進化ゲームの分析枠組みを経済学に応用した著名な論文の一つが、東大の神取道宏さんが、論文執筆当時の彼の同僚だったMailathとRobと共同で書いた論文、通称KMR論文があります。この論文の仮定を例にとれば、進化ゲームで考えるゲームのプレイヤーの戦略選択は、次に述べる限定合理的な性質を持っています。

第一に、各プレイヤーは一生、生まれた時に決めた戦略にしがみつき、同じ戦略をプレイし続けるので、戦略を変更することがあるのは生まれた時だけです。進化ゲームの基となった生物学の世界では、生物の行動様式（戦略）は遺伝子によってコントロールされており、いったん遺伝が確定すると行動様式も確定し、一生変わらないというわけです。なおVSRPDにおける戦略とは、「毎回異なるプレイヤーとプレイする繰り返しPDをどうプレイするか」という意味の戦略です。いったん相手とパートナーになると、繰り返しPDのゲームをプレイするわけですが、その繰り返しPDの戦略です。この戦略は、いったん決めたら一生変えずに、相手が変わっても一生同じ戦略をプレイし続けます。

第二に、一部のプレイヤーを除けば、大部分のプレイヤーは生まれたときに、近視眼的に最適な戦略を採用します。言いかえると、生物学でいう適者生存ということを考えるわけです。現状の環境に適合する戦略、現状のマッチング・プールの戦略分布に対して、最も高い利得（環境適合度）を与えてくれる戦略が、そうでない戦略に比べてより多くの子孫を増やし、社会における人口の割合がだんだん増えてくる、という状況を考えます。

第三に、生まれ変わる時にだけ、プレイヤーは戦略を変えられるわけですが、一部のプレイヤーは現状の環境に適した、それに対して最適な戦略

を選ぶ代わりに、戦略選択を行う場合に、新たに実験的な戦略を選択したり、間違えて予定した戦略とは異なる戦略を採用したりします。結果として、彼らが選ぶ新しい戦略は、現状の環境とは全く無関係な戦略を、無作為にプレイすることになります。生物同士のゲームであれば、突然変異が起こるため、まったく新しい戦略が実験的にプレイされるのです。この三つの性質を考えた時に、ランダム・マッチング・プールの戦略分布は、最終的にどんな戦略分布に行き着くのか、どんな戦略分布が安定になるのか。これが、VSRPD を進化ゲームとしてとらえた時の考え方です。

VSRPD の信頼形成の仕組み：失業

では、具体的にどういう仕組みが VSRPD の世界で安定になるのでしょうか。そのような仕組の中で、おそらく最もよく知られている仕組みが、非自発的失業が存在するという仕組みです。この仕組みを提唱した論文として一番よく知られているのは、1983 年の『American Economic Review』誌に出た Shapiro & Stiglitz の論文です。

余談になりますが、ここで若干の自己宣伝をさせてください。Shapiro & Stiglitz の論文が出版される 3 年前くらいに、私はほとんど同じ仕組みを考えて、それを日本語と英語のディスカッション・ペーパーにして、Stiglitz を含めて世界中にばら撒いたのです。残念なことに、二つの世界的に著名な雑誌に投稿したのですが、両方から蹴られてしまいました。最初の雑誌のレフェリーの一人は非常に好意的だったのですが、もう一人がダメだと言うので、編集者が掲載を拒否しました。次はどうしようかと迷っていた時に、恩師の宇沢弘文先生から『現代経済』という雑誌に何か論文を出さないかというお話があったので、ちょうどよいかと考えて、1981 年に『現代経済』誌にその論文を出しました。当時は、まだ私も若かったので、一遍、ある論文を日本語で出したら、もう英語でも同じ論文は出版できないと思っており、英語ではこの論文は出版しませんでした。その後、1990 年に Stanford 大学のビジネススクールに客員教授として赴

任した時、Milgrom と Roberts という二人の友人から、“あの論文はどの雑誌に掲載したのだ”と聞かれました。“マイナーな雑誌でもよいから、出した雑誌を教えてくれれば、Shapiro & Stiglitz の先行論文だといって宣伝してやる”と言われました。それで、実はその論文は日本語で日本の雑誌に出してしまったと答えたら、日本語雑誌ではダメだから宣伝はできないといわれてしまい、せっかくの話もなしになってしまいました。私の若い時の一番の失敗です。

余談はこれくらいにしまして、失業が信頼を担保するという仕組みとは、どんな仕組みかを説明しましょう。さっきの VSRPD の世界で、パートナー同士が協力し、(C,C) を毎期実現させるとしましょう。パートナー同士の信頼がお互いを信頼して協力 (C) をプレイすれば、(C,C) が実現し、毎期、協力の利得 c が得られます。ただ、自分が裏切るとその期は g の利得を得られるかもしりませんが、相手の信頼を失って今期末にペアを解消されるでしょう。Shapiro & Stiglitz や私の論文が扱った雇用契約ならば、今期末にサボった（厳密には、これらの論文は確率的な監視をするので、サボったことが見つかった）労働者は解雇されるわけです。解雇された場合、すぐに雇用先、つまり新しいパートナーが見つかるなら、前のパートナーの信頼を失ったからといって、困ることはありません。

しかし失業があれば、（あるいは VSRPD のモデルに即していえば、ランダム・マッチング・プールで新しい相手をすぐには見つけられなければ）、新しいパートナーを見つけるまで失業状態に置かれるわけです。失業する、つまりパートナーがいないと、その期には d という利得さえ得られず、非常に大きな経済的損失を被ることになります。そういう意味で、失業状態に置かれるという利益機会の喪失が、裏切りに対する社会的制裁として機能することになり、協力をし続けようというインセンティブを生み出すことになります。つまり、パートナーがいない状態の利得が低く、人々が将来を十分に大事だと考えるなら、IC 条件が満たされ、現在のパートナーの信頼を保つために、協力して (C,C) を実現しようとする

インセンティブが生まれるのです。これが、失業が生み出す信用を担保する仕組みです。つまり、失業を避けるために、パートナーとの協力を選択するインセンティブが生まれる。別の言い方をすれば、Shapiro & Stiglitz の論文のタイトルのように、失業に陥る恐怖が、サボることを抑制するという「規律付け」の役割を果たす仕組みとして機能し、VSRPD 的な社会でも協力が実現されるのです。

VSRPD の信頼形成の仕組み：信頼構築期間

もう一つの仕組みとして、次に信頼構築期間という仕組みを説明しましょう。失業が存在しなければ、パートナーシップが解消されても、次の期には必ず新しいパートナーが見つけられます。その場合でも、相互の信頼を確立して協力を実現できる仕組みとして、1990 年代から知られている仕組みです。Greve さんと私の 2009 年の論文は、VSRPD という社会ゲームの全体像をきちんとした数学モデルに定式化して、この仕組みを分析したことに意義がありました。

この仕組みは、新しいパートナーとペアを作ると、お互いに信頼を確立して協力（C,C）をプレイし始める前に、まず 1 期以上の T 期間、（D,D）をプレイして、お互いが非協力をプレイし続けます。いわば、T 期間、お互いの信頼を構築するために、低い利得しか得られない（D,D）をプレイし続ける「信頼構築期間」を続けることが必要だ、という仕組みです。信頼構築期間が終わったら、T＋1 期目から、お互いは協力を始めて（C,C）をプレイし始めるという訳です。

信頼構築期間の仕組みの下で、どうして協力が実現するかを以下に説明しましょう。信頼構築期間を終えて信頼が確立すると、パートナーたちは協力を始め（C,C）をプレイし始めるわけですが、協力を始めた後に、もし自分が裏切ると、相手からペアを解消されてしまいます。その場合にはランダム・マッチング・プールに戻って、見つけた新しいパートナーとの間で信頼を確立するために、もう一度信頼構築をしなければなりません。

いったん信頼を確立すると、協力を続けて毎期 c の利得を獲得できるが、そこで裏切ると、またマッチング・プールに戻って新しいパートナーを探し、改めて信頼を確立するために、T 期間 c より低い d という利得に耐え続ける必要があるわけです。この、新しいパートナーと信頼を確立するために必要な経済的損失を考えれば、既存のパートナーとの協力の継続を選択したほうが良いでしょう。言い換えれば、既存のパートナーとの間ですでに信頼が確立されているなら、いわば「信頼資本」を所有していることになります。それが相手を裏切ることで失われてしまうので、そのぐらいだったら裏切らないようにしようというインセンティブが生まれるということになります。

この T−期間の信頼構築を経た上でしか協力を始めない戦略を、以下では c_Tー戦略と書くことにします。ここで T=1,2,3 とは、信頼を構築するために必要な期間の数です。信頼構築期間を経た後、C という協力を始めます。念のために述べておくと、T=1,2,3 という期間の数は、この社会ですべてのプレイヤーが共有している歴史的時間ではありません。パートナーシップを作って 1 期目、2 期目、3 期目という意味という意味での T です。ペアを作って T 期の間は D を選択して、どんな結果が実現しても継続 (k) を選択する。この間は信頼構築期間なので諦めて D をプレイする。その上で、T 期までパートナーシップが継続したなら、T+1 期からは「協力期間」に入って、C を選択し始める。その期の結果が（C,C）の場合、つまり、その期にお互いが C を選んだという場合にのみ、継続 (k) を選択する。協力期間に入ってからどちらかが D を選んだら解消 (e) を選んで、ペアを解消する。従って、お互いが c_Tー戦略を選んだ場合、T 期目までは（D,D）が実現して各期の利得は d であり、T+1 期目以降は（C,C）が実現し毎期の利得は c になります。ただし、T+1 期目以降は、その期の PD ゲームの結果が（C,C）の場合にのみ (k) を選択する。相手が裏切って（C,D）が実現したら、ペアはその期に解消する。これが c_Tー戦略です。

c_T－戦略と信頼資本

信頼構築戦略がどのような結果を生み出すかを、図４を使って説明しましょう。いま、すべてのプレイヤーが、３期信頼構築戦略である c_3－戦略を選んでいる社会を考えてみましょう。図の横軸には、ペアが形成されてから１期目、２期目、３期目と、パートナーシップが進んでゆく期間が取られており、縦軸には、各プレイヤーの平均利得が取られています。社会のすべてのプレイヤーが c_3－戦略を使っていれば、相手も c_3－戦略をプレイするので、各プレイヤーはペアができてから最初の３期間、（D,D）という結果に直面するわけです。従って、最初の１期目、２期目、３期目は、毎期 d の利得しか得られませんが、４期目からは協力が始まるので、毎期 c の利得を得られます。結果としてペアが続く限り得られる利得流列は、（d,d,d,c,c,⋯）となり、それを割引係数 δ で割り引いた平均期待利得は、最初の３期間に得る d と４期目以降に得る c の中間の水準、図の v（c_3）と記した水準になります。ペアが２期目に入ると、２期目以降に得られる平均期待利得は、１期目に得た d は過去のものになるので、将来得る利得流列は（d,d,c,c,⋯）となります。従って、すでに２期目に入った場合、そのペアが継続している限り得られる、と期待する平均期待利得は v（c_3）より大きくなります。ペアが３期目に入ると、もう１期だけ（D,D）をやって d の利得を我慢すれば、次の期からは c がずっと得られ続けるので、平均期待利得はさらに上昇します。ペアが４期目以降に入ると、それ以降、ペアが継続する限りずっと、毎期 c を得られ続けるわけですから、将来ペアが続く限り得られる平均期待利得も c になります。結果として、このペアが続く限り、t 期目以降にプレイヤーが得ると期待する将来平均期待利得は、図の下の破線で示した右上がりの曲線になります。

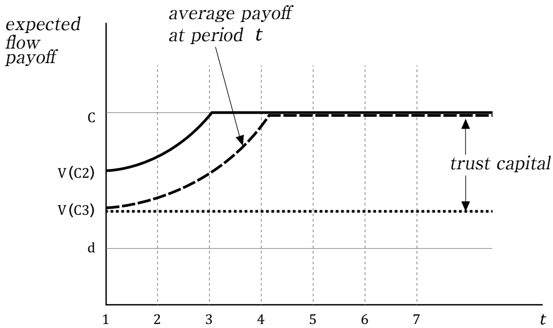

図 4　信頼資本と信頼構築均衡

信頼構築期間の仕組みは、次のように考えることができます。いま、信頼構築がすんだ 4 期目以降を考えると、このペアの相手がパートナーである限り、毎期 c の利得を得ることができます。これに対して、この相手を裏切ると、その期は g の利得が得られますが、ペアは解消され、翌期にはマッチング・プールに戻らざるを得ません。その結果、翌期以降に得られる平均期待利得は $\mathrm{v}(c_3)$ に低下してしまいます。つまり、信頼構築期間経過後に相手を裏切ると、毎期得られるはずの平均期待利得が、$c-\mathrm{v}(c_3)$ だけ低下してしまいます。この差額は、信頼構築期間の間、小さい利得で我慢することを通じて蓄積した「信頼資本」だ、と考えることができます。相手を裏切ると信頼資本を失ってしまい、それを再構築するためには、また信頼構築期間の間、小さい利得で我慢するという「投資」をしなければならない。もし、将来を大切にする人なら、相手を裏切ることで失う信頼資本を大事にして、今すぐ得られる $g-c$ の今期だけの利得増は追求すべきではない。そう考えるのではないでしょうか。これは、繰り返し PD ゲームのところで述べた IC 条件そのものです。つまり、信頼構築期間が十分に長ければ、信頼資本の大きさが大きくなり、IC 条件が満たされることになる、という訳です。IC 条件が満たされている場合、すべてのプ

レイヤーが c_T－戦略を選んでいる状態は、ナッシュ均衡になります。これを、以下では「T 期信頼構築均衡」と呼びましょう。

信頼構築均衡とその存在条件

上で説明した c_3－戦略の代わりに、社会のすべてのプレイヤーが 2 期信頼構築戦略である c_2－戦略を選んでいる場合の将来平均期待利得を表したのが、図 4 の上の実線です。2 期間で信頼構築がすむのなら、マッチの最初の期に予想する利得の流列は（d,d,c,c,…）ですから、それらを割り引いた平均期待利得は v(c_2) という v(c_3) より高い利得が得られますし、3 期目からは毎期 c が得られます。結果として、すべてのプレイヤーが c_2 戦略を使う場合の信頼資本の大きさは c−v(c_2) になり、すべてのプレイヤーが c_3 戦略を使う場合の信頼資本の大きさより、小さくなります。

以上を一般化したのが、図 5 です。横軸には、信頼構築期間の長さを、縦軸には割引係数の大きさをとってあります。社会のすべてのプレイヤーが、横軸にとられた期間 T の長さだけ信頼構築をするような戦略 c_T を採用しているとき、そのような状態が信頼構築均衡となるかどうかは、縦軸にとられた割引係数 δ の大きさに依存します。与えられた T に比べて δ が十分に大きく、（T, δ）の組み合わせが図の影を付けた部分にあるとき、T 期信頼構築均衡が社会全体の安定的な均衡になります。つまり、T に比べて δ が十分に大きければ、それは必ずナッシュ均衡になります。逆に、与えられた δ に対して一番短い信頼構築期間でナッシュ均衡になるような組み合わせが、図の太線の曲線です。T が十分に長ければ、社会ですべてのプレイヤーが c_T－戦略を使うというナッシュ均衡が存在します。ただ、T が短いほど、マッチング・プールで得られる平均利得は高くなりますから、信頼資本の大きさは小さくなり、同じ δ ではナッシュ均衡になり難くなるという性質があります。

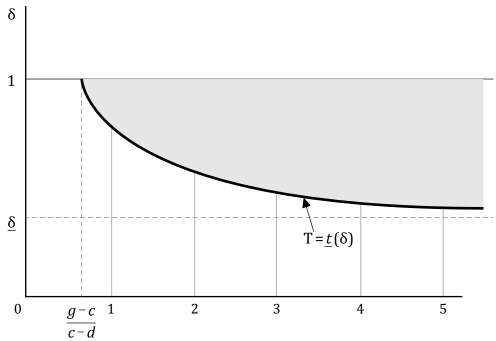

図 5　信頼構築均衡の性質

また、c_0－戦略だけからなる均衡は存在しないという事実も重要です。社会のすべての人が c_0－戦略をプレイしていれば、どんな相手もマッチした最初の期から協力（C）をしてきます。自分も c_0－戦略をとった場合、得られる利得は毎期 c です。しかし、自分はマッチの最初から非協力（D）をプレイすることで c の利得の代わりに $g\ (>c)$ の利得を得ることができます。いわば、相手をカモにするわけです。相手をカモにしたらすぐにペアを解消して、マッチング・プールで次の新しいカモを探して、また g の利得をとることができます。このマッチの最初に D をとり、何が起ころうとも期末にマッチ解消（e）を選ぶ戦略を、「ひき逃げ戦略」と呼び、以下では d_0－戦略と書きます。このように、社会のすべてのプレイヤーが c_0－戦略をとっている状態に対する最適反応戦略は、ひき逃げ戦略 d_0 であり、すべてのプレイヤーが c_0－戦略をとっている状態は、ナッシュ均衡にはなりえないのです。

なお、T 期信頼構築均衡、つまりすべてのプレイヤーが c_T－戦略を採用しているような状態が均衡として実現するためには、すべてのプレイヤーが、他のプレイヤーも信頼構築期間を T 期間続けることを確信していることが必要です。例えば、丁稚奉公を T 期間しないといけないというこ

とが慣習や規範になっているような伝統的社会でないと、この c_T－戦略均衡は成立しません。慣習や規範の役割が薄れた現代社会では、この均衡は実現し難いだろうと予想されます。

VSRPD の信頼形成の仕組み：非合理な人の存在

もう一つの知られている仕組みが、社会に非合理的な人が存在するために信頼が生まれるという仕組みです。普通の人は「合理的」な人で、ペアができた最初から相手を信頼して毎期 C をプレイして、その期の結果が（C,C）であるかぎり継続（k）を選び続けるという、c_0－戦略を採用しているとします。つまり信頼構築期間がなくて、最初の期から相手を信頼して C をプレイする。ただし、相手も協力を選んで囚人のジレンマの結果が（C,C）ならペアを継続するが、相手が裏切って今期の結果が（C,D）になると、自分はマッチを解消する、という戦略です。ただ、マッチング・プールには「非合理」な人が一定の割合、（1−α）の割合だけ存在します。非合理な人とは、理由はともあれ、とにかく絶対に D しかプレイしないという非合理な人です。非合理なので、D をプレイするロジックも持ちませんし、利得を最大化しようと考えているわけではありません。とにかく D しかプレイしない非合理な人です。

しかし、社会に非合理な人が一定割合いると、普通の人もマッチング・プールに行くと、非合理な人に 1−α の確率で出会います。非合理な人に出会った場合、普通の人が得る利得は ℓ になります。普通の人は普通の人とマッチすると、毎期ペアが継続して（C,C）が実現して、毎期 c の利得が得られます。しかし、自分が裏切って D を選ぶと、相手に現在のパートナーシップを解消されてしまいます。そうすると、ランダム・マッチング・プールに行きますが、そこには（1−α）の割合で非合理な人が存在する。非合理な人と出会うと、利得は ℓ と低くなり、信頼関係も確立できないので、次期には再度ランダム・マッチング・プールに戻って別のパートナーを探さなければならない。その期待損失が、せっかく普通の人と会っ

て（C,C）をプレイし続けているのだったら、それを継続しようというインセンティブを作り出します。

この非合理な人が存在する社会は、不完備情報ゲームで使われる「非合理タイプ（irrational type）」のプレイヤーという概念に対応しています。しかし、この種の文献では、どうして不合理な人が存在するのかという理由は、何も触れられません。ただ単に、非合理な人が存在することを前提として、分析が進められるのが普通です。

VSRPD の信頼形成の仕組み：多元的行動様式からなる均衡

これに対して我々が発見したもう一つの仕組みは、非合理的な人が存在する今の仕組みをうまく使って、以下のような仕組みを考えることです。社会の行動様式が多様であれば、結果として安定的な均衡を生み出す仕組みとして機能するということです。そのことを考えるために、先ほどの非合理的なプレイヤーが存在する社会をもう一度考えてみると、進化ゲームという視点からは何かおかしいことがわかります。つまり、非合理タイプと普通の合理的なタイプの人を比べると、通常、両者は異なる期待利得を得ているはずです。どちらかが、より高い利得を得ているはずです。その場合、より低い利得しか得られないタイプは、環境に適応していないということで淘汰されます。環境に対する適応ということを考えれば、均衡では、すべての存在している戦略は、同一水準の期待利得を獲得していないと、環境に適応していないと見なされるはずです。複数の戦略あるいは行動様式が同時に同一水準の期待利得を獲得しながら、生存し続けるという社会。そういう多元的な社会が、VSRPD には存在します。それが、これからお話をしようとすることです。失業が存在する場合や、社会に信頼構築期間が存在する場合は、社会の行動様式は一元的ですから、この仕組みの多元性は特徴的な性質です。

いま、マッチング・プールにいるプレイヤーのうち、α の割合が c_0 −戦略をプレイしているとします。念のために c_0 −戦略をもう一度説明して

おきましょう。パートナーシップが継続するかぎり毎期協力（C）を選んで、ゲームの結果が（C,C）ならばペアを継続（k）します。相手がDを選んで結果がその期の結果が（C,D）になれば、期末にペアを解消（e）します。一種のトリガー戦略です。残りの1−αの割合はd_0－戦略をプレイします。「ひき逃げ戦略」とも呼ばれるこの戦略は、パートナーシップができた最初の期に裏切ってDをプレイします。その期の末には、相手が協力しようがしまいがペアを解消（k）して、次の期はマッチング・プールに戻るという一種のグリムD－戦略です。さて、この二つの戦略が同時に共存している社会は、次に述べるような命題が成立します。

命題：δが十分に1に近ければ、ある$\bar{\alpha}$が存在して、$\bar{\alpha}$の割合の人たちがc_0－戦略をとり、残りの（$1-\bar{\alpha}$）の人たちはd_0－戦略をとるような戦略分布、つまり、c_0とd_0という二つの行動様式が同時に共存している社会状態が、局所的に安定なナッシュ均衡になる。

このような均衡を、以下では、c_0-d_0均衡と呼ぶことにしましょう。

c_0-d_0均衡：その直感的理解

この、二つの戦略が共存している状態が、局所的に安定なナッシュ均衡になるという命題を、第6図を使って、具体的かつ直感的に説明してみましょう。この図では、縦軸に平均期待利得を、横軸にはα、つまりマッチング・プールにおけるc_0－戦略の割合をとっています。図の左端はc_0－戦略がいない状態、つまり、社会がすべてd_0－戦略で構成されている状態を表しています。図の右端は$\alpha=1$ですから、社会全体がc_0－戦略をプレイしている状態を表しています。このとき、あるαが与えられ、マッチング・プールにおけるc_0の割合がα、d_0の割合が1−αである場合を考えましょう。以下、この状態を$\alpha c_0+(1-\alpha)d_0$と書きます。このとき、d_0－戦略の人たちが得る平均期待利得$v(d_0;\alpha c_0+(1-\alpha)d_0)$はどうなるで

しょうか。

d_0－戦略のプレイヤーは、α の確率で c_0－戦略と出会いますが、その場合、最初の期に自分は裏切ることで最大の利得 g を得ます。d_0－戦略はひき逃げ戦略ですから、このように、c_0－と出会うと相手を食い物にして、すぐペアを解消します。次の期には、またマッチング・プールに行って新しい獲物を探します。他方、残りの 1−α の確率で d_0－に出会いますが、その場合には、その期の利得は d しか得られません。すぐにペアを解消して、次の期にはプールに戻ることになります。結果として、d_0－戦略の人たちの平均期待利得は $\alpha g+(1-\alpha)d$ になり、そのグラフは、図の $v(d_0;\alpha c_0+(1-\alpha)d_0)$ と記した右上がりの直線になります。

次に、c_0－戦略の平均期待利得ですが、これは δ の大きさに依存します。c_0－戦略をプレイする人は、α の確率で c_0－戦略に出会いますが、その場合、両方が生きている限り、毎期 c の利得を取り続けられます。正確に言えば、期待値で言えば $1/((1-\delta^2)$ 期間、c を取り続けられます。他方、1−α の確率で d_0 と出会って、その時には１期だけ ℓ を得ます。結果として、獲得する平均期待利得 $v(c_0;\alpha c_0+(1-\alpha)d_0)$ は、$(\alpha\frac{c}{1-\delta^2}+(1-\alpha)\ell)/(\frac{\alpha}{1-\delta^2}+1-\alpha)$ という形になります。分子が同じパートナーが継続する間に得られる割引期待利得で、分母が期待パートナーシップ継続期間です。平均期待利得は、ペアが続いている間に得られる割引期待利得を、ペアの期待継続期間で割った値だ、というわけです。そのグラフは図に示したように、右上がりの上方に凸の形をした曲線になります。

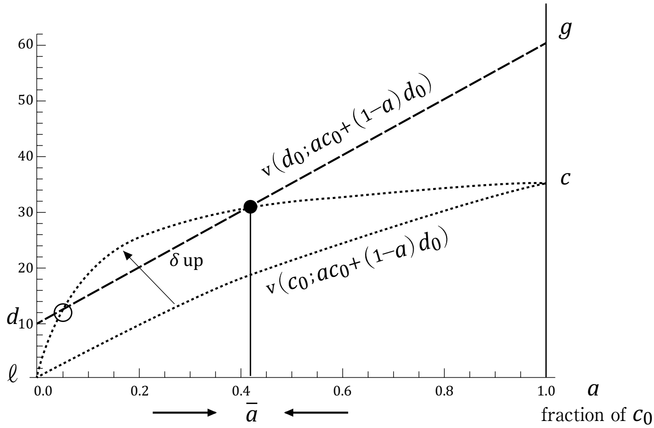

図６　命題１の直観的理解：利得の非線形性

大事なことは、が大きくなればなるほど、凸性が増すということです。もし α=0 で c_0－戦略が存在せず、相手がみんな d_0－戦略なら、平均期待利得は必ず ℓ になります。逆に α=1 で、相手がみんな c_0－戦略ならば、平均期待利得は必ず c になります。しかし、α がゼロと１の間の値にあると、δ が増えるにつれて、c_0 と出会ったときの期待継続期間が長くなり、それだけ得られる平均期待利得の値が大きくなります。結果として、δ が十分に大きくなると、二つの曲線が交わり始め、図の黒い丸い点のような交点が出現し始めます。マッチング・プールにおける c_0 の割合が $\bar{\alpha}$ であるこの点では、d_0 の平均期待利得も c_0 の平均期待利得も等しくなり、どちらの戦略を選んでも同一の平均期待利得が得られます。しかも、$\bar{\alpha}$ の近傍では、α が $\bar{\alpha}$ より増えると、d_0－の利得の方が c_0 の利得よりも大きくなるので、d_0 を選ぶ人の割合が増えて α が下がる。逆に、α が $\bar{\alpha}$ より減ると、c_0 の利得が d_0 の利得を上回るので α が増え、この状態は、動学的にも局所的に安定になります。

最後に、証明が技術的になってしまうので、ここでは説明しませんが、

この状態ではIC条件も満たされることが分かっています。つまり、c_0同士が出会って（C,C）という協力関係が始まると、裏切ってDをプレイするより、Cをプレイし続けて協力関係を維持し続けた方が利得が高い、つまり協力関係を維持しようというインセンティブが存在します。以上が、$\bar{\alpha}c_0 + (1-\bar{\alpha})d_0$分布が均衡になり、上記の命題が成立する直感的な理由です。

c_0－d_0均衡へのc_1－戦略の侵入の可能性

ところで、c_0－d_0均衡はナッシュ均衡ではあるのですが、進化的な意味で安定ではないことが知られています。言い換えると、c_0－d_0という戦略分布には、別の戦略が侵入できます。ここで「侵入」というのは、マッチング・プールの正の割合の人が同時にその戦略を採用した時、その戦略の平均期待利得が、既存戦略の平均期待利得を上回るため、その戦略が既存戦略を駆逐してしまう、という意味です。ここで、侵入できる戦略はc_1－戦略です。これは、信頼構築期間が1期間である信頼構築戦略です。つまり、1期目にはDをプレイして、相手が何をしてもペアを継続する。2期目から協力を始め、2期目以降はc_0のように行動します。c_1が、実はc_0－d_0均衡に侵入できるのです。

念のために、ナッシュ均衡と安定性の違いを説明しておきましょう。ナッシュ均衡とは、他の人たちがナッシュ均衡通りの戦略をプレイし続けるとき、自分だけがナッシュ均衡が指定する戦略を変更しても、その戦略変更によってその人の利得が増加しない状態のことです。これに対して、進化的に安定とは、他の人たちが現在プレイしている戦略を変えないとき、一人ではなく、二人以上の複数の人たちが、同時に既存戦略とは異なる同一の戦略を採用しても、それらの戦略変更をした人たちの利得が増加しない状態のことです。もし、ある状態が進化的に「不安定」なら、複数の人たちが、既存戦略とは異なる、別の戦略に戦略を変更することで、以前より得をすることになります。このように、ある戦略を複数の人たちが採用して、他の人より高い平均期待利得を得ることで、その戦略が既存戦

略を淘汰してその戦略をプレイする人口が増加してゆくこと、結果として既存戦略が淘汰されてゆくことを指して、その戦略が既存戦略分布に「侵入する」というのです。

VSRPD の場合、c_1－戦略がマッチング・プールに、わずかではあるが正の割合だけ参入すると考えます。マッチング・プールには無限の数のプレイヤーがいますから、参入する c_1－戦略をプレイする人は、無限の数のプレイヤーで、彼らが同時に c_1－戦略をプレイし始めることになります。このとき、c_1－戦略は既存戦略に出会うと、d_0－戦略と同じ利得を得ます。具体的に言うと、c_1－戦略は c_0 に出会うと 1 期目に D をプレイして g の利得を得ますが、相手から裏切ったなと思われて、ペアを解消されてしまいます。他方、d_0 と出会った場合、1 期目に D をプレイして d の利得を得て k を選択しますが、相手の d_0 からペアを解消されます。どちらの場合も、得られる期待平均利得は、自分が d_0 だった場合と同じです。

ただ、マッチング・プールには c_1－戦略が正の割合いるわけですから、正の確率で自分と同じ c_1－戦略と出会います。この時に何が起きるかが問題です。1 期目にはお互いに D をプレイして d の利得しか得られませんが、d_0 と出会った場合と違って、お互いにペア継続 (k) を選びますから、ペアは二期以降も継続できます。ところが、2 期目にはお互いが C をプレイしますから、2 期目以降は信頼が構築され、毎期ずっと c の利得が取れることになります。この利得（c）は、マッチング・プールに戻った場合に d_0 が得られる平均期待利得より大きいですから、c_1－戦略は明らかに d_0－戦略より高い利得が取れるわけです。つまり c_1－戦略が c_0－d_0 均衡に参入すると、既存戦略に比べてより高い平均期待利得を得ることができ、その人口比率は次第に増えていくことになります。そういう意味で、c_0－d_0 均衡は、進化的に不安定だ、というわけです。

c_1 同士が出会うと（D,D）をプレイした後、(k,k) が実現します。(k,k) が実現するということは、相手は d_0 ではないことを意味しています。お互いが、相手が d_0 ではなくて c_1 だと理解しますから、安心感が生

まれ、2 期目以降の継続戦略である c_0 を二人ともプレイし始めて、毎期（C,C）が実現する。これが c_1－戦略が c_0－d_0 均衡に侵入できた直感的な理由なのです。

c_0－d_0 均衡への c_1 戦略と d_1 戦略の同時参入

だとすると、実は次のような問題もあります。c_0－d_0 均衡に c_1－戦略だけでなく、d_1－戦略も同時に参入した場合には、事情は異なるのではないかということです。ここで、d_1－戦略とは、次のような戦略です。ペアを組んだ第 1 期には c_1 とまったく同じ行動をします。要するに D を選んで、何が起ころうと k を継続するということです。c_1－戦略は、（(D,D)，(k,k)）となった場合、継続戦略として c_0 をプレイするわけですが、d_1－戦略の場合には、継続戦略として d_0 をプレイします。つまり、（(D,D)，(k,k)）がおきて、c_1－戦略が相手と信頼構築ができたなと安心感をもったときに、そのような c_1 を逆に食い物にしようと D をプレイするというわけです。もし c_1 と同時に、d_1 も侵入すると、2 期目に c_1 は C をプレイし始めるわけですが、d_1 は同じ経路のあと 2 期目に裏切って D をプレイするので、自分はもっと高い g という利得を得ます。逆にその場合、相手の c_1 は利得が ℓ になってしまい、d_1 の割合が大きいいと、c_1 の平均期待利得の優位性が失われて、むしろ d_1 の平均期待利得が高まります。従って、d_1－戦略が大量に参入すれば、c_0－d_0 均衡に対する c_1－戦略の侵入に歯止めがかかることがわかります。

よく考えてみると、c_0－d_0 均衡に侵入するのが c_1－戦略だけだ、と考えるのは、進化ゲームで新しく参入する戦略は人々が新しい戦略を実験しているのだ、という考え方にそぐいません。侵入するのが c_1－戦略だけだと考えることは、人々が新たな実験を行う時に、みんなが連携（coordinate）して、一斉に c_1－戦略を実験することを意味するからです。人々が実験をするときに、このようなコーディネーションが起こるとは考えにくいでしょう。だとすると、例えば人々の一部は c_1－戦略を実験す

るが、残りの人々は d_1－戦略を実験するかもしれません。前者に比べて後者の割合が十分に大きければ、c_0－d_0 均衡は、進化的にも安定だということになるでしょう。現在、私は、Greve さんとの新しい論文で、このような形で進化ゲームの安定性の定義を拡張することを試みています。

c_0－d_0－c_1－d_1 からなる四戦略均衡

さて、以上の議論からわかることとして、次のようなことがあります。c_0－d_0 均衡に c_1 と d_1 が一緒に参入する場合、c_1 の相対的な割合が大きいと、c_1 の利得は既存戦略の平均期待利得より高くなりますが、d_1 の利得はもっと高くなります。c_1 の相対的な割合が小さいと、c_1 の利得も d_1 の利得も、既存戦略の平均期待利得より小さくなります。では、c_1 の割合が中間の適切な水準にあるとき、c_0, d_0, c_1, d_1 という四つの戦略が同時に存在しながら、それらの戦略がすべて同一の平均期待利得を獲得するというナッシュ均衡は存在しないでしょうか。実は、δ が十分に１に近く、$\bar{\alpha} c_1 + (1-\bar{\alpha}) d_0$ という戦略分布が c_0－d_0 均衡になる場合には、四戦略全体に占める c_0 の割合が $\bar{\alpha}$ で、しかも c_1－戦略が c_1 と d_1 の和に占める割合も $\bar{\alpha}$ になるような四戦略分布も、ナッシュ均衡になります。

図７は、この四戦略分布がナッシュ均衡になる直感的な理由を示しています。まず、c_0－戦略が得る利得流列を示しているのが、図７の上段左側部分です。c_0－戦略は、マッチング・プールで、自分と同じ c_0－戦略に出会う確率が $\bar{\alpha}$ ですが、c_0 －戦略と出会うと長期間にわたってずっと毎期 c の利得を取り続けます。次に、それ以外の d_0 や c_1 や d_1 といった戦略全体を d_{0+} と書きましょう。これらの人たちの人口割合は $1-\bar{\alpha}$ ですが、d_{0+} に属するどんな戦略が相手でも、相手は最初の期に裏切りますから、その期に c_0 は ℓ を得て、そのあとはマッチング・プールに戻ります。マッチング・プールから出発するとき、その後、毎期 c_0 が得る平均期待利得（c_0－d_0 均衡ですから、d_0 が得る平均期待利得も同じです）を v^* と書くことにすれば、次の期以降、c_0 は毎期 v^* の利得を得ることになります。

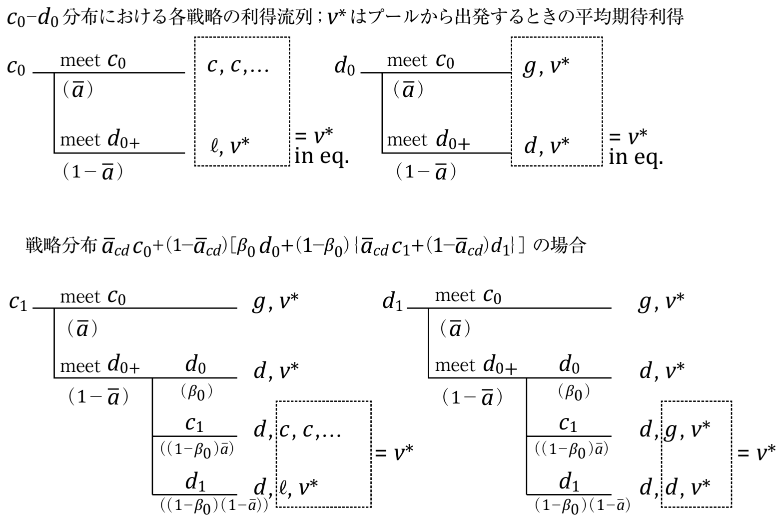

図７　四戦略分布がナッシュ均衡になる直観的理由

他方、d_0－戦略の得る利得流列を表しているのが、図７の上段右側部分です。d_0－戦略は、c_0－戦略と出会うと、今期相手は協力するが自分は裏切るので g を得ます。しかし、d_0 はひき逃げ戦略なので、１期でパートナーシップを解消して２期目にはマッチング・プールに戻るので、その後は毎期 v^* の利得を得ることになります。d_0 が d_{0+} と出会うと、今期はお互いが裏切って d を得ますが、d_0 はひき逃げ戦略ですから、期末にはペアが解消され、来期以降は毎期 v^* を得ることになります。

残りの二つの戦略の利得流列はどうなるかを示したのが、図７の下段です。まず、図７の下段左側が、c_1－戦略の利得流列を表しています。c_1－戦略は c_0 と出会うと、最初の期に D をプレイして利得 g が取れますが、即座に相手にペアを解消されます。結果として２期目以降はマッチング・プールに戻るので、毎期 v^* の平均期待利得を受け取ることになります。

他方、d_{0+} に出会う場合は、三つのケースに分けられます。まず、d_0 に出会うケースでは、最初の期に d の利得が得られますが、相手はひき逃げ戦略ですから期末にペアを解消されて、2 期以降の利得は毎期 v^* になります。c_1 に出会うケースでは、1 期目の結果は（D,D）で利得は d になりますが、1 期で信頼構築が終わって 2 期目から協力が始まり、以降毎期 c の利得が取れます。最後のケースは d_1 に出会うケースです。この場合、1 期目は（D,D）で利得は d ですが、2 期目は自分は協力を始めますが、相手はひき逃げを始めるので、自分は ℓ の利得しか取れません。相手がひき逃げなのでペアは期末に解消されて、3 期目以降はマッチング・プールに戻って毎期 v^* を得ることになります。

同様のことは、d_1－戦略についても考えられます。その結果は、図 7 の下段右側の図に示してあります。時間がないので詳しくは説明しませんが、図を見ながらよく考えていただければ、結果がこうなることは明らかでしょう。そうすると、図 7 の左上の図と左下の図の枠で囲った部分が一致すること、右上の図と右下の図の枠で囲った部分が一致することから、c_0－d_0 均衡の期待平均利得が v^* であるなら、この図の四戦略分布でも四戦略すべての平均期待利得が v^* になり、この戦略分布がナッシュ均衡になることが明らかです。

六戦略均衡と VSRPD の実験

さて我々は、この VSRPD を、理論的に研究するだけでなく、その実験研究を行っています。最近、経済学にも実験が導入されています。実は、日本で経済学の実験研究を行ったのは、たぶん私が初めてです。理由があって東大では認められず、実験自体は慶応大学でやったのですが、最近は東大でも認められるようになってきました。そこで、東大の本郷および駒場のキャンパスを使い、東京大学の学生や大学院生を被験者に、VSRPD のゲームをプレイさせるという実験を行いました。あまり細かい具体的な内容をお話しても意味がないと思うので、一番興味深い結果が

何だったかということをお話ししましょう。我々が事前に考えていた c_0, d_0, c_1, d_1 という四つの戦略に加えて、実験をしてみると、多くの被験者が「Forgiving 戦略」をプレイすることがわかりました。Forgiving 戦略とはどんなものかというと、1 期目に自分が C をプレイしたのに相手が D をプレイした、つまり「相手に裏切られた」としても、期末にペアを継続（k）して、2 期目にまた C をプレイする。つまり、「裏切った相手を許す」という行動が多数みられたことです。逆に、そういう戦略を真似しよう、1 期目に（C,D）が起こっても、期末に継続を選択するが、2 期目には C ではなく D をプレイする、という行動を選ぶ被験者も、結構多いということがわかりました。

このような実験結果を見て、理論的にこれらの行動を説明できないかと考えた結果、実はもっと別の多様な行動様式（戦略）が共存する、「六戦略均衡」と呼んでいる均衡が存在することを発見しました。いま、この理論的研究を、グレーヴァさんと私の共著論文として、まとめている最中ですが、時間がないので、それを説明する図 8 を示しておきます。ここで、Cc_0 というのが Forgiving 戦略で、1 期目は C をプレイし、結果が（C,C）ならペアを継続し、次の期以降は c_0 と同じようにプレイします。1 期目の結果が（C,D）でもペアを継続し、2 期目以降の継続戦略を c_0 にします。（つまり、2 期目には C をプレイし、（C,C）が起きた場合のみ継続を選んで C をプレイ、2 期目以降どこかで（C,D）が起きればペアを解消というわけです。）Cd_0 は、Cc_0 と同様、1 期目は C をプレイし、結果が（C,C）なら c_0 と同様に行動します。1 期目の結果が（C,D）の場合、Cc_0 と同様に継続を選びます。しかし、2 期目にはあたかもひき逃げ戦略のように行動し、継続戦略は d_0 になります。（つまり、2 期目には D を選び、相手が何をしようとペアを解消します。）Dc_0 は今まで c_1 と呼んでいた戦略、Cd_0 は今まで d_1 と呼んでいた戦略です。図 7 を説明したのとまったく同じ理由で、図 8 を見れば、この六戦略均衡が均衡になることがわかります。

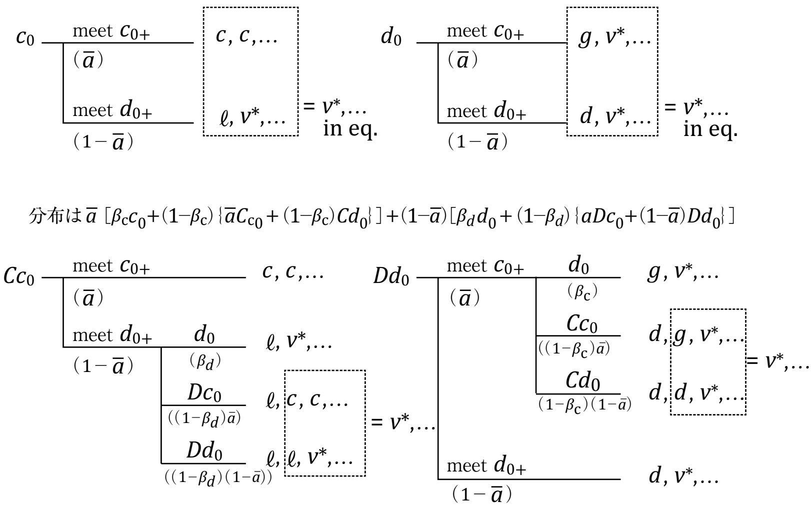

図 8　六戦略分布が均衡になる直観的理由

実験研究の結果

理論予測と実験結果を簡潔に説明しておきましょう。$\bar{\alpha}$ の値、つまり c_0 ー d_0 均衡で、1期目にCをプレイする戦略が全体に占める割合は、理論予想では0.368ですが、実験の結果は0.344です。値が近いという印象はありますが、厳密に統計検定をすると、等しいという仮説は棄却されしまいます。もう一つ重要な発見は、1期目に（C,D）が実現した後に、期末にkという「ペアを解消しよう」という選択を行う被験者たちの存在です。結果として、((C,D),（k,k)）が実現して、ペアが2期目まで継続した場合に、被験者がCをプレイする割合が、異常に高いことがわかりました。すでに述べた六戦略均衡のうち、Forgiving戦略をプレイする人たちの割合が、理論予想に比べて格段に大きい、と言い換えてもよいかもしれません。六戦略均衡の場合、その割合の理論値はやはり0.368ですが、実験してみるとその値は0.6を大きく超えます。その結果をいろいろ議論

した結果、このような行動を被験者が取るのは、k という選択を、被験者たちはペアの相手に対して、「次の期には C をとろう」というシグナルとして使っているのではないか、と推測しています。k がシグナルとして機能しているのなら、それはどのような理由なのか、理論的な理由は何か、統計的に検証できないか、といった作業を、現在やっています。

最後に

最後のまとめに入りましょう。グローバル化し匿名化が進んだ現代社会において、見知らぬ人たちが信頼を構築して協力を始めるためには、失業とか信頼構築期間とか、いろんな仕組みが知られていいます。しかし、これらはすべて、すべての人が同一の行動をとるという一元的な仕組みです。しかし現実を見る限り、人々が多様な行動様式をとるという、多様性に基づいた仕組みの方が、どうも現代社会に合っているのではないか、というのが我々の問題意識であり、また、我々の研究結果です。どうしてかというと、例えば T 期信頼構築均衡では、ちょうど T 期間の間、すべての人が信頼構築期間を行うという共通認識を、社会のすべてのメンバーが共有していなければなりません。丁稚奉公を 20 年するといった社会慣習や社会規範がある社会だったら、確かに信頼構築型の社会ができるでしょう。しかし、グローバル化して異なる社会慣習が混じり合う現代社会では、それは難しいのではないでしょうか。丁稚奉公の仕組みに基づいて社会慣習や社会規範が生まれることを、coordination device といいますが、それは昔の小さな共同体では可能かもしれないが、グローバル化して匿名化が進んでいる現代社会では、社会慣習や社会規範が coordination device として機能するのは困難ではないでしょうか。むしろ現代社会では、多様な行動様式が共存することによって、裏切るインセンティブを抑えるという社会的な仕組みの方が、coordination device が不要だということもあり、信頼や協力を生み出しやすいのではないか、というのが今日のお話の要点の一つです。

では、多様化する社会でも、どうして見知らぬ他人を信用できるのでしょうか。その直感はどう説明したらよいでしょうか。それは実は、人を信頼できない人がいるからこそ、逆に協力をしてくれる人を信頼できるのだ、ということにほかなりません。信頼できない人が存在するからこそ、信頼できる人に対する信頼の礎が生まれてくる。そういう意味で信頼できない人がいるということ自体が、社会にとってはある種の必要悪なのです。多様な価値観が存在することは、むしろ、自分と似た価値観を持つ人たちが、お互いの社会関係が壊れることに恐怖を与えるために、価値観の共通した人たちがより社会的に強い紐帯を持とうとするというインセンティブをもたらし、結果として社会の強さをもたらしているという逆説的な見方もできる。このように、現代社会を社会ゲームとしてとらえると、真の意味で、多様性が重要だという結論が得られます。詳しくお話しする時間がありませんでしたが、先ほどお話しした「裏切りを許す人」、「裏切りを許す人をカモにする人」、それに加えて様々な行動様式を持つ人が社会には同時に存在する均衡があります。結果として、実は協力に行き着くまでに、どんな行動パターンを考えても、そういった行動パターンを経ることで初めて協力に行き着く。どんな行動パターンを経ても協力に行き着くことが可能だ。そんな、多様性を内包する社会の均衡が存在していることがわかってきました。それが、今、まとめている最中の論文です。

定刻なので私の報告はこれで終わりにしたいと思いますが、もし皆様がよろしければ５分ほど時間をとって、質問にお答えしたいと思います。もし、何かあればですが。

ディスカッション

植田健一教授（東京大学）　二点ほど、質問があります。一つは比較的テクニカルな質問ですが、多様性均衡、たとえば c_0-d_0 均衡というのは、一人の人が、$\bar{\alpha}$ の確率で c_0 をとって、$1-\bar{\alpha}$ の確率で d_0 をとるという混合戦略とは考えられないのでしょうか。もしそうだとすると、多様性というの

は、人々のなかでそれぞれの個人が c_0 と d_0 という異なる戦略を確率的に取っている社会になります。その場合、多様性とは個人のなかでの多様性かもしれないと思います。そういう理解でも良いでしょうか、という点が第一の質問です。

もう一つはより大きな質問です。先生のお話で、私も学生だった頃を覚えているのですが、日本経済では、アメリカの場合と違って、一見合理的でないが、良く考えてみると慣習とかそういうものを使って説明できる事象がありました。私自身も、あまりフォーマルな契約とかを考えないで、どんな社会が出来るのかということに興味があり、分析もしています。一方で、現実問題としてどんどん契約社会化が進んで、結婚にさえ契約が入り込んでいるという状況です。契約を通じてしっかりとペナルティーをかけることができるようになって、法整備を通じて社会的な制裁ができるように、日本の社会がなってきたという側面もあると思います。そうした背景を考えた時に、それでも契約によらずに経済問題を考えた方がいいのか、それともやはり契約社会化を進めて、最適な契約を考えていく方向で政策を考えた方がいいのか、そのあたりをお聞きしたいと思います。

奥野教授　前者の混合戦略ですが、もちろん多様性均衡を混合戦略均衡として考えることもできます。ただ混合戦略には、直感的に理解しにくいという問題があります。ゼロサム・ゲームである、ペナルティーキック合戦とかじゃんけんなどをプレイする場合の混合戦略にはそれなりの意味がありますが、普通の社会、普通のゲームで混合戦略を無理に考えるのは、本当はあまり適切ではないという強固な議論があります。混合戦略という考え方は、人間の直感に反しているのではないか、というわけです。実はゲーム理論研究者の中でも、ノーベル賞をとったハルサニなどはそういう意見です。混合戦略均衡の場合、そのゲームの利得構造を私的情報と考えたうえで、利得の値を少しだけ動かすと、混合戦略均衡が純粋戦略均衡に変わります。それを purification と呼びます。purification が言わんとするところは、混合戦略均衡が存在するのは極めて例外的な場合で、ほとんど

の場合、純粋戦略できちんと考え均衡が計算できる、という考え方です。私は、purification の考え方が好きです。特に VSRPD は、人間が大きな現代社会をプレイする社会ゲームです。このような社会で、人間は自分の行動を、コイン投げとかサイコロを振って無作為に決めるのではなく、同じ文脈である限り、身に付いた一つの行動様式をプレイし続けるのではないか、と思います。そうならば、それはやはり純粋戦略で考えるべきではないでしょうか。ただ VSRPD を研究する人たちの中にも、混合戦略で考える人たちもいます。

それから二つ目の質問ですが、昔の日本は同質的な社会で、その意味で coordination device が機能しやすい社会でした。今でも日本にはそういう側面があるので、日本社会は慣習や規範である程度説明できるかもしれません。逆に、欧米は昔から契約社会という側面が強いという違いがあるかもしれません。ただ問題は、契約にはコストがかかるという点です。したがって、大きな stake がかかった経済関係であれば、もちろんきちんとした契約をするでしょう。しかし、小さい、例えばネットショッピングだとかヤフオクであるといった、stake が小さい経済関係では、契約ということまでやらないし、評判も機能しにくいわけです。さらに特殊詐欺をはじめ、詐欺のような stake はある程度大きいけれども、評判も機能しないという経済機会も増えてきている。だから、経済機会の一部は契約とか慣習とか規範を使って implement できるけれども、stake の小さい経済機会をはじめ、そうできない部分もいっぱいあるのではないか。その残っている部分をきちんと分析しようというのが、私の今日の問題です。両方ともが重要で、契約を使って implement することに反対しているわけではありません。

中泉拓也教授（関東学院大学）　信頼の話は非常に重要だと思うのですが、最近、フェイクニュースの問題が昨年の大統領選以降重要になって来ています。先生がご指摘のように、情報化によって情報の非対称性が増えるため、フェイクニュースも増えると考えられます。これについて、本講演の

趣旨である信頼形成のメカニズムからはどう考えればよいでしょうか。

奥野教授　基本的にフェイクニュースが存在することが、社会にある種の安定性を作り出しているということだと思います。「これはフェイクニュースかもしれない」という不安が、本当に信頼できるニュースソースを信頼する理由になっているという……。だからこそ、ある人を信頼し始めたならば、その人を信頼し続けようというインセンティブが生まれます。そういう人がせっかくいるのだから、裏切らないようにしようというインセンティブが生まれます。逆説的ですが、社会には変な人がいるということが、逆に真っ当な人が生きていける環境を作り出している、社会に安定性を作り出している、ということこそ、今日、私が強調したかったことです。

中泉拓也教授（関東学院大学）　まさに、フェイクニュースの問題が深刻になったことで、むしろニューヨークタイムズの有料会員が増えているということが、まさにご指摘のことを裏付けていると改めて認識いたしました。ありがとうございます。

司会　まだまだご質問があるかもしれませんが、次に懇親会が控えておりますので、その時にお話いただければと思います。奥野先生、今日は本当に貴重なご講演ありがとうございました。

奥野教授　こちらこそ、ご清聴ありがとうございました。

司会　それでは今までの多大なる御貢献に感謝を込めまして花束の贈呈をさせていただきたいと思います。以上を持ちまして奥野先生の最終講義を終了させていただきます。この後向かいのロハスカフェという所で6時15分から懇親会を行う予定です。20分くらいに伸ばすようにしますので、その頃まで来ていただきたいと思います。行く途中に図書館がございます。そこに奥野先生の一連のご著書が展示してありますので、是非ご覧いただきたいと思います。本日はどうもありがとうございました。