2016年1月2日

経済セミナー2016年3, 4月号

「オークションとマーケットデザイン」第**18**回

繰り返しゲームと感情

松島斉

東京大学経済学研究科教授

今回は、メカニズムデザインから離れたテーマだ。メカニズムデザインは、あらかじめルールをきちんとつくることによって、みんなにインセンティブを提供する。しかし、日常の多くの対立、ジレンマ、紛争の解決においては、ルールをいちいちデザインするのは現実的でない。

今回は、ルールを明文化せずとも、ジレンマが顕在化しないように、長期的な関係を利用して協調を引き出す、「ちょっと大人の対応」を考察したい。それは「繰り返しゲーム」と呼ばれる、重要なクラスのゲーム理論分析にあたる。

言いたいことは、長期的な関係における人々の振る舞いを理解するカギは「(寛容な) しっぺ返し戦略」であること、そして、プレーヤーの「感情」が行動に影響を与えていること。これである。

以下、Kayaba, Matsushima, and Toyama (2016)の実験結果を紹介することで、深遠なる「繰り返しゲーム哲学」に少しだけ迫ろう。

長い付き合いでジレンマ回避

利害の対立を示すゲームの代表格は「囚人のジレンマ」。図の例を考えよう。

	C	D
C	$Z-Y$　　$Z-Y$	$-Y$　　Z
D	Z　　$-Y$	0　　0

二人のプレーヤーがいて、各プレーヤーは「協力（C）」か「非協力（D）」を選択する。C を選択すると、コスト Y 円が自身にかかるが、相手には Z 円の便益を与える。便益 Z がコスト Y を上回ると仮定しよう（$Z > Y$）。これは大きな正の外部性だ。ならば、お互い協力 C を選択すれば、達成される配分は最適になる。

しかし、C を選択すれば本人にはコストのみが発生するだけだ。そのため、非協力 D を選択するのが優位戦略になり、これは唯一のナッシュ均衡戦略でもある。つまり、囚人のジレンマでは協調（C、C）は均衡にならないのだ。

しかし、長期的な関係となれば、話は別だ。例えば、各プレーヤーは、もし今日 C を選択するなら、相手は明日以降 C を選択してくれる（アメ）が、D を選択すると D で報復してくる（ムチ）と予想しよう。これなら、報復を恐れて、C を選択しようということになりそうだ。

トリガー戦略

トリガー戦略は、こんな「アメとムチ」を端的に示す好例だ。トリガー戦略においては、最初の期には協力 C を選択する。以降、お互いに（C，C）をプレイし続ける限りにおいて、C を選択し続ける。しかし、どちらかが一度でも D を選択すれば、次期以降ずっと D を選択する。

こんなトリガー戦略をふたりともプレイすれば、どちらも率先しては D を選択しないから、協力関係（C，C）がずっと維持されることになる。トリガー戦略のプロファイルは、将来利得をあまり割り引かず十分加味する、長期的視野のプレーヤーにとっては、ナッシュ均衡、より正確には、「部分ゲーム完全均衡（Subgame Perfect Equilibrium、SPE）になる。

今日 C を選択せずに D を選択するなら、費用 Y を節約できよう。しかし、明日以降は（C，C）でなく（D，D）を選択し続けることになる。つまり、毎期（C、C）がもたらす利得$Z-Y$を享受できたはずなのに、（D，D）がもたらす利得であるゼロ、つまり費用 Y は負担しないが相手から Z の施しもない､という状況にスイッチしてしまう。これは損だ。

もう少し正確に説明しよう。プレーヤーは将来利得を割引ファクター$\delta\in[0,1)$で割り引いて評価するとしよう。ならば、次期の損失は$\delta(Z-Y)$、その次の期の損失は$\delta^2(Z-Y)$、３期後の損失は$\delta^3(Z-Y)$、といった具合に将来損失を評価するので、合計

$$\delta(Z-Y)+\delta^2(Z-Y)+\delta^3(Z-Y)+\cdots$$
$$=(Z-Y)\sum_{\tau=1}^{\infty}\delta^{\tau}=\frac{\delta(Z-Y)}{1-\delta}$$

が将来損失にあたる。将来損失$\frac{\delta(Z-Y)}{1-\delta}$が利益 Y を上回るならば、協力的に C を選択し続けることが得になる。将来損失$\frac{\delta(Z-Y)}{1-\delta}$は、割引ファクター$\delta$が 1 に近いならば、十分大きな値になる。不等式

$$Y\le\frac{\delta(Z-Y)}{1-\delta}\text{、つまり}\ \delta\ge\frac{Y}{Z}$$

が成立していれば、トリガー戦略プロファイルは SPE になり、協力関係（C，C）は継続的に実現できる。

フォーク定理

ただし、ここで気を付けるべきは、トリガー戦略のみならず、（D，D）の繰り返しもまた SPE であることだ。実際、SPE はたくさん存在し、しかも各々性質の異なる配分を達成する。

トリガー戦略プロファイルでは、任意の時点からスタートする戦略プロファイルが、選択の履歴に応じて、トリガー戦略プロファイルになったり、（D，D）の繰り返しになったりしている。大事なことは、トリガー戦略プロファイルも、（D，D）の繰り返しも、ともに SPE であることがアメとムチの実行を保証してくれる点だ。

どの期からスタートしても、繰り返しゲームの SPE 全体の集合は同じである。まるで金太郎飴のようだ。普通の金太郎飴と違うのは、飴の断面にはいっぱい金太郎の顔があって、履歴に依存してどこかの顔につながっている。こうすることで、自由に繰り返しゲームの SPE を生成できる。繰り返しゲームの SPE のもつこの重要な性質は自己生成（Self-Generation）と呼ばれている（Abreu (1988)）。

Self-Generation によって、割引ファクターが 1 に近い場合には、SPE によって達成できる配分の範囲が非常に広いことが証明できる。このことは、囚人のジレンマに限らず、広範囲のゲームにおいて同じように成立しうる。この性質は、「フォーク定理」と呼ばれている。繰り返しゲームにおいては多種多様な配分が均衡として達成できる、というわけだ。この性質は、ずいぶん古くから推測されていたものの、満足のいく証明は Fudenberg and Maskin(1986)によってなされた。

しっぺ返し戦略（**TFT**）

トリガー戦略以外に、繰り返しゲームの戦略として非常にポピュラーなのは、「しっぺ返し戦略（Tit-for-Tat、TFT）」である。TFT では、前期に相手が C を選択したならば今期は C、前期に相手が D を選択したならば今期は D、といった具合に、前期の相手の選択と同じことをする。こうすることで、相手の選択の履歴と自分の履歴が似た状態に保たれる。だから、TFT をプレイすれば、相手がどのような戦略をとろうとも、相対的に大損する、あるいは大負けすることにはならない。そのため、TFT は、任意の戦略と総当たり戦をしてみると上位の成績を収めるらしい（Axcelrod (1984)）。TFT は、こんな必勝戦略の代名詞のようにいわれてきた。

しかしながら、TFT の戦略プロファイルはめったに SPE にならない。

相手が前期に C を選択したならば、あなたは今期 C を選択する、つまり費用 Y を負担しないといけない。そうしなければ次期に相手のしっぺがえしをくらって δZ 損する。よって、不等式

$$Y \le \delta Z \text{、つまり } \delta \ge \frac{Y}{Z}$$

が成立すれば、あなたは C をとるインセンティブをもつことになる。

逆に、相手が前期に D を選択したならば、D で報復しないといけないが、次期には相手にしっぺ返しを受けるので、逆の不等式

$$Y \geq \delta Z\text{、つまり}\ \delta \leq \frac{Y}{Z}$$

が成立しないと、あなたは D をとるインセンティブをもたない。TFT が SPE になるための必要十分条件は、C と D 両方のインセンティブをみたすことである。よって、等式

$$\delta = \frac{Y}{Z}$$

がピンポイントで成立しないといけない。

割引ファクター δ が $\frac{Y}{Z}$ より高いと、しっぺ返しの報復が強すぎて、D をとりたがらないし、$\frac{Y}{Z}$ より低いと報復が弱すぎて C をとりたがらない。割引ファクターが丁度 $\delta = \frac{Y}{Z}$ でない限り、TFT 戦略プロファイルは SPE にならない。

寛容なしっぺ返し（GTFT）

TFT の概念を少し拡張して、均衡分析にも都合がいいように、以下のような「寛容な」しっぺがえし（Generous　Tit-for-Tat, GTFT）」戦略を考えてみよう

GTFT では、相手が C ならば確率 $r(C) \in [0,1]$ で C をとる。相手が D ならば確率 $r(D) \in [0,1]$ で C をとる。これは、$r(C) = 1$、$r(D) = 0$ であれば、TFT そのものだ。GTFT は、C をランダムに選択することを認めることによって TFT を拡張した概念だ。

$r(C)$ が $r(D)$ より大きければ、相手には少なからず報復を与えることになる。D をとれば、費用 Y を節約できるが、次期に相手が C をとる確率は $r(C) - r(D)$ 分だけ下がる。つまり、期待値として

$$\delta\{r(C) - r(D)\}Z$$

円分を損してしまう。GTFT が SPE になるためには、プレーヤーにとって C と D が無差別にならないといけない。だから

$$\delta\{r(C) - r(D)\}Z = Y\ \text{つまり}\ r(C) - r(D) = \frac{Y}{\delta Z}$$

となるように確率差 $r(C) - r(D)$ が設定されれば、GTFT の戦略プロファイルは SPE になる。

報復の程度を示す確率差$r(C)-r(D)$を、強すぎず弱すぎず、いい塩梅に定めれば、そして、$r(C)=1$とおけば、TFT 戦略プロファイルはプレーヤーに（C，C）を恒常的に選択させることができる。$r(C)-r(D)$は 1 以下でないといけないから、$\frac{Y}{\delta Z}$も 1 以下、つまり

$$\delta \geq \frac{Y}{Z}$$

であれば、$r(C)=1$、$r(D)=1-\frac{Y}{\delta Z}$と設定された GTFT 戦略プロファイルは存在し、それは均衡になり、（C，C）を継続的に選択する協調関係が成立しうる。

不完全モニタリング

長期的に協調関係が維持されるためには、相手の選択が観察できるとする前提が不可欠だ（完全モニタリング）。全く観察できないとなれば、アメとムチは使えないので、こうはうまくはいかない。

重要なことは、相手の選択が直接観察できなくても、相手の選択に関連するなんらかのシグナルが観察できる場合（不完全モニタリング）には、GTFT は、不完全ながらも、ある程度は協調を達成できる点だ（Matsushima (2013)）。

プレーヤーは相手の選択を直接観察できない。このことは、現実の長期的な関係においてたびたび生じうる。ならば、直接観察はできないけれども、相手の選択に関係するシグナル、シグナル G（Good）あるいはシグナル B（Bad）のどちらかを観察できるとしよう。直接は無理だが間接的な観察ならできるというわけだ。ただしシグナルには、以下のように、ノイズがついているのだから、要注意。

相手が C を選択した場合、シグナル G を確率$p>1/2$で観察し、シグナル B を確率$1-p<1/2$で観察する。逆に、相手が D を選択した場合、シグナル B を確率$p>1/2$で観察し、シグナル G を確率$1-p<1/2$で観察する。$p>1/2$であることから、C を選択すれば G（Good）が観察されやすく、D を選択すれば B（Bad）が観察されやすい。

GTFT は、シグナル G を観察したならば確率 $r(G)\in[0,1]$ で C を選択する、シグナル B を観察したならば確率 $r(B)\in[0,1]$ で C をとる、と定義される。$r(G)$ は $r(B)$ より大きい。

D をとれば、費用 Y を節約できるが、次期に相手が C をとる確率、つまり Z 円もらえる確率が、$r(G)p+r(B)(1-p)$ から $r(G)(1-p)+r(B)p$ に、つまり確率差 $(2p-1)\{r(G)-r(B)\}$ だけ、下がってしまう。そのため、期待値として

$$\delta(2p-1)\{r(G)-r(B)\}Z$$

円分を損してしまう。C と D が無差別でないといけないから

$$\delta(2p-1)\{r(G)-r(B)\}Z=Y\text{、}$$

つまり、

$$r(G)-r(B)=\frac{Y}{\delta Z(2p-1)}$$

となるように、報復の程度を示す確率差 $r(G)-r(B)$ が定められれば、この GTFT は不完全モニタリングにおいて SPE になる。$r(G)-r(B)$ は 1 以下でないといけないから、$\dfrac{Y}{\delta Z(2p-1)}$ も 1 以下、つまり

$$Y<\delta Z(2p-1)$$

をみたすように Y が十分小さいならば、長期にわたって、C がある頻度で持続的に選択されることになる。

モニタリングの精度 p が 1 に近い、完全モニタリングに近い状況であれば、ほぼ確実に C が選択され続ける。しかし、不完全モニタリングである以上、C をとっても相手には B が観察され、次期に非協力的な D を選択される、という事態がおこることはさけられない。だから、一般に常に C が選択され続ける、というわけにはいかない。

モニタリングの精度と報復の強さとの関係

均衡における報復の強さをあらわす確率差 $r(G)-r(B)=\dfrac{Y}{\delta Z(2p-1)}$ は、モニタリングの精度 p の減少関数になっている。精度がよければ報復も弱くていい。弱い報復でも、精度がいいから十分効果的、というわけだ。

報復がもっと強いと、相手を報復するために D をとるインセンティブがなくなってしまう。報復すれば今度は相手に強く報復される、のを嫌うからである。逆に報復がもっと弱いと、今度は C をとるインセンティブがなくなってしまう。こうして、$r(G)-r(B)=\dfrac{Y}{\delta Z(2p-1)}$ は、丁度バランスのとれた報復の程度になり、C と D を無差別に保つ唯一無二の値になる。

このことは、精度がいい場合に効率性をより高く保つ秘訣にもなっている。均衡をプレイする合理的なプレーヤーは、精度が高ければ報復を低くして、効率性のロスをうまく節約している、というわけだ。

レヴューと長い報復

モニタリングが不完全であれば、GTFT の均衡では、非協力的な関係が定期的におこることをさけられない。ならば、GTFT よりももっと効率性を促進できる別の戦略的工夫はないだろうか。

任意の精度 $p\in(0,1)$ において、GTFT よりも協調を促進できる一案を紹介しよう。それは、相手のシグナルをしばらく観察してから、報復するか否かを決めるやり方だ。これは Reviewing とか Lenience とか呼ばれている。

十分長い一定期間、相手のシグナルを観察しつづける。相手がずっと C をプレイし続けていれば、この間ほぼ割合 p で G が観察されるはずだ。ならば、G が観察された割合が p よりもだいぶ下回れば報復し、そうでなければ C をとる、とすれば、不完全モニタリングにおいても、ほぼ効率性のロスなく報復ができるはずだ。つまりは、「大数の法則」を利用する、というわけ。

この際、GTFT のように一回限りの報復ではなく、報復期間も長くとって、報復をすると判断したからには厳しく報復する、といった追加措置も必要になる。相手がずっと D をとって得をしてきた分を相殺できるだけの十分な報復が必要だからだ。

公的モニタリングと私的モニタリング

さらに問題になるのは、もし相手の選択についてのシグナルだけでなく、自身の選択について相手が観察するシグナル、つまり、自分が相手にどう見られているか、に

ついても観察できる場合には、Reviewing におけるインセンティブの問題はかなりややこしいものになる。

自分が相手にどう見られているかがわかる、つまりシグナルが公的に観察されるケースは、「公的モニタリング」と呼ばれる。これに対し、相手にどう見られているのかわからない、シグナルが私的にのみ観察されるケースは、「私的モニタリング」と呼ばれる。

繰り返しゲーム理論史、特に「フォーク定理」の研究史においては、公的モニタリングにおいて、不完全にしかモニタリングできなくても、割引ファクターが 1 に近ければ、近似的にフォーク定理が成り立つことが、私的モニタリングに先んじて、証明された（Fudenberg, Revine, and Maskin (1994)）。シグナルが共通に観察できるため、観察されたシグナルに依存して次期以降の SPE がスイッチするように、Self-Generation をうまく使って、このフォーク定理が証明されている。

この証明には、メカニズムデザインとも共通するアプローチが使われている。一回限りのゲームにおいて、観察されたシグナルに応じて所得移転するルールを設計するメカニズムデザインのアプローチと、Self-Generation を使ったフォーク定理の証明の仕方は、とてもよく似ている。（Legros and Matsushima (1991), Matsushima (1989)）.

一方、私的モニタリングの場合にフォーク定理が成立することは、上述した Reviewing を使って証明された。Matsushima (2004)は、囚人のジレンマの例において、任意のモニタリング精度においてフォーク定理が成り立つことを証明したのだ。

公的モニタリングでは、Reviewing の途中で、自分の選択についてどのようにシグナルが相手に伝わっているかがわかる。よって、Reviewing 後に自分が報復を受けるかどうかが、だんだんわかってくる。そのため、C をさらにとり続けて相手に好印象をあたえるインセンティブが、だんだんなくなってしまう。このことは、公的モニタリングのケースに Reviewing を適用することをとても難しくしている。

しかし、Sugaya (2012)は、この困難を乗り越え、不完全モニタリングにおける非常に一般的なゲームについて、フォーク定理が成立することを証明してみせた。これは驚くべき快挙！

繰り返しゲームを実験する

フォーク定理は、ゲーム理論という応用数学を代表する、大定理だ。しかし、この定理には大きな欠点がある。たくさんの SPE の中でいったいどれが実現するのか？この問いに対しては、何も語ってくれないからだ。均衡条件だけではどの配分が達成されるか、全く予測できない。予測力に著しくかける定理なのだ。

繰り返しゲームにおいてどのようなプレイがなされるのか？どんな配分がなされるのか？理論だけではどうもらちが明かない。

ならば、もっと直接的に調べてみることにしよう。つまり、不完全私的モニタリングの繰り返しゲームを、実験室で、被験者にテストさせるのである。はたして、被験者はどのような振る舞いをするのか。

だがここで、さっそく問題発生だ。繰り返しゲームは無限回でないと意味がない。しかし、実験の時間には限りがある。

有限回の最後の期では、協調はとてもおこりにくい。ならば、最後から二番目の期でも、明日は非協調とわかっているから協調しない。といった具合に、後方帰納法的に考えていくと、最初から協調しないのが筋である。実際、過去の実験研究の結果をみても、有限回の繰り返しゲームを「繰り返す」と、被験者はだんだん協調しなくなる。

そこで、ちょっとした実験の工夫が必要になってくる。まず、実験は有限回だが、どの時点で終了するかは被験者に知らせない。そして、期ごとに終了する確率を固定して、この確率は被験者にあらかじめ教えておく。各期末では、その都度、終了するかしないかを決するルーレットを被験者に見せる。こうすれば、ああ、次期以降もまだ継続するんだなあと、「The shadow of the future」があることへの注意を被験者に喚起させることができる。（もっとも、これは「まやかし」なんだけどねえ。そして現実もまた、まやかしなんだ。）

このような工夫を実験に最初に取り入れたのは Dal Bo(2004)であり、これは比較的新しい試みだ。無限回の繰り返しゲームの実験は、この意味で、現時点でようやく端緒についたばかりだといえよう。

実験では GTFT がプレイされている

前述した不完全私的モニタリングの囚人のジレンマを実験すると、被験者ははたしてどのような行動をとるだろうか？

Kayaba, Matsushima, Toyama (2016)は、囚人のジレンマの利得の数値を固定して、割引ファクターを 1 に近く設定して、モニタリングの精度については、高いケース（$p = 0.9$）と低いケース（$p = 0.6$）の両方について、実験をおこなった。そして、実験結果から、被験者がどのような戦略に従って行動していたかを、以下のように推定した。

被験者には、ずっと D を取り続ける人、ずっと C を取り続ける人、ランダムに C と D を選択する人、TFT の人、Reviewing している人、ながく報復する人、などなど、いろんなタイプがいそうだ。そこで、「ありそうな」戦略をたくさんリストアップしておいて、それぞれの戦略が各被験者に利用されている頻度を、「Strategy Frequency Estimation Method」とよばれる最尤法（Dal Bo and Fréchette (2011)）によって推定してみた。すると、かなり多くの被験者（半数以上）が、精度の良し悪しに関係なく、GTFT に一番マッチしていることが分かった。

さらに重要な結果として、被験者がシグナル G を観察した場合とシグナル B を観察した場合とで、どの程度次期に C を選択する頻度が違うか、つまり報復の強さ、を調べてみると、精度の高い時には、GTFT の均衡の値よりも高く、精度の低い場合は GTFT の均衡の値よりも（かなり）低いことがわかった。

モニタリングの精度が高いと、概して C をとる頻度が高い。この意味では、被験者はシグナルを上手に活用しているといえよう。しかし、報復の強さは、精度が上がると、高くなっている。この意味では、シグナルの活用の仕方として、あまり利口とはいえまい。

被験者の選択の傾向をある程度予測できるのなら、精度が高いと報復を弱くして効率性のロスを下げるはず。なのに、実験結果はその真逆だった。精度が低い場合には、均衡より報復はかなり小さく、精度が高い場合には均衡よりも報復は大きかった、というわけだ。

やさしい感情といじわるな感情

この実験結果は、今までの理論的考察が意味することとずいぶん違っている。どうやら、今までとは違う発想から、長期的関係のことを考え直したほうがよさそうだ。例えば、プレーヤーは、金銭だけで動機付けられているのではなく、「感情」にも影響されていると考えてみてはどうだろう？

あなたは、相手が C を選択することを期待している。しかし、シグナル B をうけとった。あなたは不快に思い、D をとっていじわるすればすっきりする。これは、ずいぶん「いじわるな」感情だ。

逆に、あなたは、相手が C を選択することをあまり期待していない。しかし、シグナル G をうけとった。あなたはうれしく思い、C をとってやさしくおかえししたくなる。これは、「やさしい」感情だ。

いじわるな感情とやさしい感情のどちらが重要か。これは、協調しやすいか、つまり、モニタリングの精度が高いか、あるいはそうでないかに影響されそうだ。精度が低いと協調しにくい。だから、少しでも相手のいい情報が入れば、やさしくしてあげたくなる。逆に、精度が高いと協調しやすいので、いいシグナルがくるのはあたりまえ。だから今度は逆に、悪いシグナルに、やけに敏感になる。

このような観点から、均衡分析を再検討してみよう。

プレーヤーが任意の GTFT、$(r(G;p),r(B;p))$、をプレイするとしよう。実験結果に即して、$r(G;p)$、$r(B;p)$、$r(G;p)-r(B;p)$ はいずれも精度 p の増加関数としよう。特に、最後の増加関数の仮定は、実験結果とは整合するが、金銭だけで動議つけられた均衡理論には反する性質だ。

感情を考慮することによって、GTFT が均衡として記述される可能性を、以下のように考えよう。もし

$$r(G;p)-r(B;p)>\frac{Y}{\delta Z(2p-1)}\text{、}$$

つまり報復が理論値より強ければ、D をとって報復することを金銭的には躊躇することになる。ならば、B を観察した時に、いじわるな感情をしずめずに C をとってしまうと心理的コスト $c(B;p)\geq 0$ がかかるとしよう。すると、C と D が無差別になる条件は

$$\delta(2p-1)\{r(G;p)-r(B;p)\}Z-C(B;p)=Y$$

つまり

$$C(B;p)=\delta(2p-1)\{r(G;p)-r(B;p)\}Z-Y$$

になる。

モニタリングの精度が高いため報復 $r(G;p)-r(B;p)$ が強すぎる場合には、いじわるな感情をしずめる心理的コスト $C(B;p)$ は正になる。モニタリングの精度が上がると $r(G;p)-r(B;p)$ も上昇するから、心理的コスト $C(B;p)$ も上昇する。これは、精度が

上がって協調しやすい状況になると、相手の悪い情報に敏感になり、いじわるな感情が支配する関係になっていると解釈できる。

逆に、もし

$$r(G;p)-r(B;p)<\frac{Y}{\delta Z(2p-1)}、$$

つまり報復が理論値より弱ければ、C をとって協調することを金銭的には躊躇することになる。ならば、G を観察した時に、やさしい感情にさからって D をとってしまうと心理的コスト $c(G;p)\geq 0$ がかかるとしよう。すると、C と D が無差別になる条件は

$$\delta(2p-1)\{r(G)-r(B)\}Z=Y-C(G)$$

つまり

$$C(G;p)=Y-\delta(2p-1)\{r(G;p)-r(B;p)\}Z$$

になる。

モニタリングの精度が低いため報復 $r(G;p)-r(B;p)$ が弱すぎるならば、やさしい感情にさからう心理的コスト $C(G;p)$ は正になる。モニタリングの精度が上がると $r(G;p)-r(B;p)$ は上昇するので、心理的コスト $C(G;p)$ はダウンする。これは、精度が上がって協調しやすい状況になると、相手の良い情報に反応しなくなり、やさしい感情は影をひそめていくと解釈できる。どうやら私が思った通りだ。

モニタリングの精度が低い場合には、相手の良い点を積極的に評価しようとし、やさしい感情が長期的な関係を支配する。しかし精度が高くなって、協調がしやすくなってくると、相手のあらを探しては憤るようになり、いじわるな感情が支配するようになる。

ここには、金銭的には効率的な配分の達成と、冷たい社会との間の「不都合な」トレードオフが垣間見える。モニタリングしやすいために協力できる関係と、そうでもない関係とで、どちらがより「幸せな」社会といえるか。 経済学は、今のところ、この問いには答えてくれない。シグナルを刺激として受け、感情で反応するパターンを探求するアプローチ、つまり心理学的なアプローチとされるものからも、答えは期待できまい。

社会の規範が人々の行動を操っていることについて、もっと哲学しないといけない。この哲学の中に、幸せな社会をみいだすヒントがあるかもしれない。このヒントは、経済学をもっと価値ある方向に導くかもしれない。

村八分と居場所探し

今まで説明してきた繰り返しゲームは、常に同じプレーヤーがゲームを何度も繰り返している状況だ。今度は、大勢の経済主体が出会いと決別を繰り返している状況を考えてみよう。こんな、対戦相手が変わりうる状況でも、協調関係が成立しうることをうまく説明できるだろうか。

真っ先に思いつきそうなのが、「村八分（Ostracism）」だ。過去の選択の記録が、「評判」として、部外者にも広まっていく。そして、あなたは D を選択すると、そのことが他の全員に悪い評判として伝わり、次期以降誰からも相手にされなくなる。こんな村八分を恐れるので、あなたは C をとるインセンティブをもつことになる。

駆け出しの研究者だった 1980 年代、私は、このような、村八分を肯定する説明が大大嫌いだった。村八分は、限定された範囲の社会的関係においては、まあいいのかもしれない。しかし、複雑な現代社会の規範として肯定されるようなものではない。村八分は、差別や偏見の温床になるからだ。

村八分を実践する人たちの目は血走り、悪意に充ちる。こんな連中は地獄に落ちる。今もこの思いはあまりかわらない。

ならば、村八分とはことなる仕方で、インセンティブが提供される、別の見方はないものか？あれば、地獄に落ちる人も救済できそうだ。

では、私が駆け出しの頃考えていた、村八分にかわる協調の作法、「居場所探し」を、最後に紹介しよう（Matsushima (1990)）。私は今でもこのアイデアを、今後どう料理していこうかと、ひそかに温存している。

各プレーヤーには、「相性のいい人」が、社会に一定割合いる。そして、私にとって相性のいい人にとって、私は相性のいい人だ。しかし、誰が私にとって相性がいいかは、会ってみないとわからない。

この場合、相性のいい人にたまたま出会えたら、C をとり、お互いに C をとり続ける限りこのよい関係が維持される。しかし、いったん D を選択してしまえば、決別して、次期に別の人とマッチしなければならない。この時、また相性のいい人にマッチする、というわけには、必ずしもいかない。相性の悪い人にマッチすれば、お互いに D を選択して、また決別して、「居場所探し」をさらに続けないといけない。

こんな居場所探しは、時間がかかり、損である。ならば、よい居場所に巡り合えたならばよい関係を大事にしよう、ということになり、C を選択し続けて真面目につとめようとするだろう。

ここには、もはや村八分や評判などは機能していない。だから、社会に巣食うような差別や偏見も生まれまい。と、まあ、私は少しこう思うのだが、みなさんいかが？

参考文献

Abreu, D. (1988): On the theory of finitely repeated games with discounting, *Econometrica* 56, 383-396

Axelrod R. (1984): *The Evolution of Cooperation*, Basic Books, New York.

Dal Bo, P. (2005): “Cooperation under the Shadow of the Future,” *American Economic Review* 95, 1591-1604.

Dal Bó, P. and G. Fréchette (2011): "The evolution of cooperation in infinitely repeated games: Experimental evidence." *American Economic Review* 101, 411–429, 2011.

Fudenberg, D., D. Levine, and E. Maskin (1994): “The folk theorem with imperfect public information,” *Econometrica* 62, 997-1040.

Fudenberg, D. and E. Maskin (1986): “The folk theorem in repeated games with discounting and with incomplete information”, *Econometrica* 54, 533-554.

Kayaba, Y., H. Matsushima, and T. Toyama (2016): “Accuracy and Retaliation in Infinitely Repeated Games with Imperfect Private Monitoring: Experiments and Behavioral Theory,” University of Tokyo.
http://www.econexp.org/hitoshi/

Legros, P. and H. Matsushima (1991): “Efficiency in partnerships,” *Journal of Economic Theory* 55, 296-322.

Matsushima, H. (1989): “Efficiency in Repeated Games with Imperfect Monitoring,” *Journal of Economic Theory* 48, 428-442.

Matsushima, H. (1990): “Long-term partnership in a repeated prisoner's dilemma with random matching,” *Economics Letters* 34, 245-248.

Matsushima, H. (2004): "Repeated Games with Private Monitoring: Two Players," *Econometrica*, 72, 823-852.

Matsushima, H. (2013): “Interlinkage and Generous Tit-for-Tat Strategy”, *Japanese Economic Review* 65, 116-121.

Sugaya, T. (2012): *The Folk Theorem in Repeated Games with Private Monitoring*, Ph. D.

Thesis, Princeton University.