

# 経済実験におけるスパイト行動

大 和 毅 彦<sup>1)2)</sup>

## 1. はじめに

新古典派経済理論の特徴の一つとして、以下のことがあげられよう。

- 1) 各経済主体は自己の利得のみに関心があり、何らかの制約条件の下で自己の利得を最大化する行動をとるものと想定する。
- 2) 経済活動は、時および場所に関わらず、1)の自己利得最大化モデルから導かれる帰結で近似的に説明できる。

前者は、「ホモ・エコノミクス(homo economicus)」の仮定と呼ばれ、アダムスミスが考え出した人間の概念と言われている。後者は暗黙のうちに仮定されてきたと思われるが、以下ではこの仮定を「普遍原理」と呼ぶことにしよう。

これに対して、哲学者デイヴィッド・ヒュームは『人性論 第2篇「情緒について」(1739)』の中で、以下のように述べている。

“Now as we seldom judge of objects from their intrinsic value, but form our notions of them from a comparison with other objects; it follows, that according as we observe a greater or less share of happiness or misery in others, we must make an estimate of our own, and feel a consequent pain or pleasure. *The misery of another gives us a more lively idea of our happiness, and his happiness of our misery. The former, therefore, produces delight; and the latter uneasiness.*”

(in “A Treatise of Human Nature: Being an Attempt to Introduce the Experimental

- 
- 1) 東京工業大学大学院 社会理工学研究科 価値システム専攻・社会学専攻。
  - 2) 大学時代より長年に渡り、ご指導ご鞭撻下さいました酒井泰弘先生に心より感謝申し上げます。本研究の一部は、文部科学省科学研究費補助金(課題番号15310023)の補助を受けました。

Method of Reasoning into Moral Subjects, Book II Of the Passions”)

他人との比較で自己の満足度を考えることが人間本来の特性であり、「他人の不幸は我々に自己の幸福（の事実）以上に生氣ある概念を与えるし、他人の幸福は我々の不幸と同様な概念を与える、従って前者は歓喜を産むし、後者は不快を産む」とヒュームは主張する。もしそうならば、人間は他人の幸福を減ずるため、相手の足を引っ張る「スパイト（いじわる）行動 (spiteful behavior)」をすると考えるのが自然であろう。しかし、そのようなスパイトをする人間がいる社会の方が、アダムスミスが想定したような純粹に自己の利得のみに関心のある人々からなる社会より、悪い悲惨な帰結が生まれるのではないだろうか？

しかし、全く逆のことが起こりうるものが、近年の実験研究で明らかになってきている。本稿では、公共財供給の参加ゲーム実験と最後通牒ゲーム実験を例としてとりあげて、前者では「スパイト行動が協力の源泉」、後者では「スパイト行動が公平性を生み出す源」となりうることを見る。これらの結果は、ホモ・エコノミクスを仮定した利己的な主体のみからなるモデルでは説明できない重要な経済現象であり、普遍原理の妥当性に関して大きな疑問を投げかけていると言えよう。

本稿は以下のように構成される。まず、2節では、経済学において実験研究の果たす役割について概説する。3節では、公共財供給におけるただ乗り問題を解決するメカニズムについて、これまでの研究と問題点を指摘する。4節では、公共財供給の参加ゲームに関する理論を説明し、5節では、そのゲームの実験結果について述べる。6節では、最後通牒ゲームの実験結果を紹介する。7節では、なぜ人々はスパイト行動をとるのかについて、進化学的な観点から考察する。最後の節で、今後の研究の方向性について述べる。

## 2. 経済学における実験の役割

現実の人間の経済活動や経済の動向を説明しようとする経済理論の分析方法は、基本的に自然科学のそれと同じである。つまり、まず現実をよく観察し理論モ

デルを構築し、次にその理論モデルが再び現実を説明できるか、その妥当性を検証して、さらに、必要ならば理論モデルを改訂していく。しかしながら、この「理論の構築→妥当性の検証→理論の改訂」という論理実証主義のサイクルにおいて、自然科学と経済学で過去大きく異なっていた点は、理論の妥当性の検証が自然科学は実験によって行われるのに対して、長年経済学のような社会科学において実験はできないと考えられてきたことである。

経済学では、代わりに現実経済で調査・収集された「経済統計データ」に基づいて、理論の検証が行われてきた。しかしながら、経済統計データをそのまま用いて、理論モデルの厳密な検証を行なうことには問題がある。理論モデルの想定とは異なる環境でデータが得られ、モデルで想定していない外的要因が影響している可能性が大きいからである。さらに、理論的に性能がよいと思われるが、歴史上存在しなかった新たな経済制度（例えば、温室効果ガス排出権取引制度など）を導入する場合、データそのものがないために、本当に理論通りうまく制度が機能するか否かを事前にチェックすることができない。

これらの問題点を解消しようと、経済学においても1980年代以降実験研究が発展してきた。経済学において実験手法を導入する目的の一つは、理論モデルの想定する環境を実験室の中で構築して、理論の検証を行なうことである。はっきりと定義された環境下で人間行動を観察し、意思決定の仕組みを左右する基本原理を明らかにするのである。これは、物理学における素粒子実験、原子核実験などに対応したものと言えよう。

経済学における実験研究のもう一つの重要な役割は、既存と異なる新しい制度の検証である。実験室で採用候補となる制度の性能を確認すれば、制度設計の失敗による損失を未然に防ぐことが可能になる。また、実験結果から、制度の不備を修正すると同時に、その制度に対する理論的な裏付けの問題点も発見<sup>3)</sup>できる。これは、航空工学や土木工学などにおける風洞実験に対応したものであ

3) 理論的に性能がよい制度を検証するというのであれば、「理論の検証」の一種であると言うこともできる。しかしながら、理論の検証では、人間行動の基本原理の解明 ↗

る。制度の性能を実験で検証し、さらに制度の理論を改訂していくというこのアプローチは、「制度設計工学」、「マーケット・デザイン」などと呼ばれている。例えば、地球温暖化ガス排出権取引市場、航空機発着権の割当制度、大気中の硫化物の排出許可証市場、全米医学研修生と病院のマッチング・プログラム、電波周波数帯のライセンスの割当制度、電気の取引市場、入札の談合防止制度などの制度設計に実験研究が生かされている。

もちろん、実験室で得られたデータは、現実の経済における生のデータではない。この点から、現実経済の調査で得られた経済統計データは重要である。論理実証主義における理論の検証において、経済統計データによる研究と実験データによる研究は、代替財ではなく補完財であると言うべきであろう。

経済学は実験ができない分野であるという考えが不幸にも支配的であった。しかし2002年度ノーベル経済学賞が実験経済学者バーノン・スミスに贈られたことからわかるように、上で述べた実験研究の重要性は認知されてきている。

1980年代以降の実験結果を大雑把にまとめると、まず、市場に関する実験に関しては、経済理論の仮説が実験室でもサポートされている。例えば、ダブルオークションなどの市場取引実験では、取引数量と取引価格の平均値は、需要曲線と供給曲線の交わる均衡数量と均衡価格とほぼ等しいことが観察されている。このことは、市場全体の需要曲線と供給曲線がどういう形をしているのか、実験参加者は誰も知らず、買い手は自分の留保価格（財に対して支払ってもよい最高価格）、売り手は自分の費用（財の仕入れ値）についてしか分からない状況で成立する。一人一人は自分のことしか知らず、自分の利得だけを考えて行動しても、取引数量や価格は効率的な均衡状態になる。まさに、アダムスミスの「神の見えざる手」が実験で体験できるのである<sup>4)</sup>。この市場実験に関する結果は、さまざまな国、多種多様な被験者に関しても成立する頑健なものである。

4) が強調されているのに対して、制度の検証では、制度が実際に機能するかどうかに重点がおかれている。さらに、理論的には複雑すぎて分析不能な制度に関しても、性能の検証を実験で行うことは可能である。

4) 市場取引実験の方法と実験結果については、例えば、大和（2001）を参照せよ。

このように、市場に関する理論については、普遍原理が成立することが実験でも確認されている。

これに対して、外部性があり市場が失敗するような環境での実験や、ゲーム理論が想定するような参加者数が少ないケースでの実験では、理論をサポートしているケースとそうでないケースの両方が観察されている。しかしながら、実験結果が理論予測とは違う場合でも、それらの相違の仕方が実験によって大きく異なるわけではなく、ある一定の傾向が観察されている。そこで、実験結果を分析し、理論が見落としていた点を考察することにより新たな視点、洞察を得ることが可能になる。

以下では、理論予測と実験結果が一致しない例として、公共財供給の参加ゲーム実験と最後通牒ゲーム実験をとりあげて考察する。

### 3. 公共財供給のただ乗り問題は解決可能か？

Samuelson (1954) は、どの経済主体も同時に消費できる（純粹）公共財が存在する経済では、社会的に望ましいパレート効率な配分を実現することは困難であると主張した。利己的な経済主体は、他の主体による公共財供給の費用負担にただ乗り（free ride）しようとするために、結果として、公共財は効率的な水準に比べて過小供給されてしまうからである。このただ乗り問題を解決する問題は、長年オープン・クエッションであったが、Groves-Ledyard (1977) が初めて、公共財のある社会において望ましい配分を達成する仕組み・メカニズムのデザインに成功した。彼らのメカニズムにおけるナッシュ均衡配分は常にパレート効率になるのである。彼ら以降、パレート効率性に加えて、個人合理性などの望ましい条件を満たす配分を遂行するメカニズムが構築されていった。よって、少なくとも理論的には、ただ乗り問題は解決できたように思われる。

しかしながら、これまでのメカニズム・デザインでは、社会の構成員全員がメカニズムに参加することが強要されてきた。これは、公共財の重要な特質である「非排除性（non-excludability）」が無視されていることを意味する。もし参加するか否かを定める自由が認められていれば、メカニズムに参加しないこ

とによって、参加した人々が生産した公共財からの便益を非排除性から享受できはすである。この「不参加によるただ乗りの問題」は、特に、国際公共財を供給するメカニズムである国際条約において重要である。例えば、地球温暖化を引き起こす二酸化炭素などの「温室効果ガス」を削減する目的で作成されたメカニズムである「京都議定書」は、2005年2月に発効した。しかしながら、世界最大の温室効果ガス排出国である米国のブッシュ政権は、2001年に議定書からの離脱を表明し、批准する意思がない。批准（参加）しなければ議定書に従う必要がなく、他の批准（参加）国による「温室効果ガスの削減」にただ乗りできる。このため、議定書の効力が疑問視されている。

もう一つの例としては、化学兵器禁止条約（chemical weapons convention, CWC）があげられる。化学兵器の禁止・処分に関して合意ができるまで24年間を要したが、1993年1月に各国によるCWCの署名が始まり、1997年4月には65カ国が批准（参加）し、CWCは発行した。2005年11月現在、175カ国が批准（参加）している。しかし、化学兵器開発疑惑国であるシリア、レバノン、北朝鮮などは署名も批准（参加）していない。また、イスラエルは1993年に署名したものの、未だ批准（参加）していない。したがって、CWCの効力も制約されていると言えよう。

これらの事実を踏まえて、Saijo-Yamato（1997, 1999）は、過去のメカニズム・デザイン理論とは異なり、各主体が参加するかどうかを選べる状況を考察した。彼らは、公共財の存在する経済において、全ての主体が参加するようなインセンティブを持つメカニズムをデザインすることはできないという不可能性定理を導き出した。この不可能性定理により、我々は公共財供給のためのメカニズムをデザインすることをあきらめなければならないのであろうか。この疑問に一つの回答を与えてくれるのが、Cason-Saijo-Yamato（2002）とCason-Saijo-Yamato-Yokotani（2004）の被験者を用いた参加ゲームの実験研究である。次節以降では、彼らの実験で検証した理論、実験の設定、および実験結果を説明する。

#### 4. 参加ゲームの理論

二人の主体、一つの私的財 $x$ 、一つの公共財 $y$ から成る経済を分析する。以下では、公共財供給のメカニズムの一つとして、公共財を作るために各主体が自分の私的財を自発的に支払うあるいは投資する「自発的支払メカニズム (voluntary contribution mechanism)」を考えよう。各主体 $i$ の私的財の消費量を $x_i$ 、公共財の生産量を $y$ と表そう。各主体は、私的財の初期保有 $w_x$ の中から公共財の生産のためにいくら投資するか、 $s_i$ 、を自発的に決める。二人が投資した私的財の量の合計がそのまま公共財の水準となる線形の生産関数を仮定する、つまり、 $y=s_1+s_2+w_y$ である。ただし、ここで $w_y$ は公共財の初期保有量である。各主体は、私的財と公共財に関して同じコブ=ダグラス型の利得関数

$$u_i(x_i, y) = a + (x_i^\alpha y^{1-\alpha})^\beta / b = a + ((w_x - s_i)^\alpha (s_1 + s_2 + w_y)^{1-\alpha})^\beta / b$$

を持っているものとする。以下では、 $a=500$ 、 $b=50$ 、 $\alpha=0.47$ 、 $\beta=4.45$ 、 $w_x=24$ 、 $w_y=4$ としよう。利得関数は二人の投資量 $s_1$ 、 $s_2$ の関数となり、各主体 $i$ は自分の利得 $u_i(s_1, s_2)$ を最大化するように投資量 $s_i$ を選ぶ。この自発的支払メカニズムのナッシュ均衡において、投資量は $\hat{s}_i=7.69$  ( $i=1, 2$ )、公共財供給量は $\hat{y}=18.38$ 、利得は $\hat{u}_i=7089$  ( $i=1, 2$ )となる。また、サミュエルソン条件より、パレート効率な公共財水準において、その供給量は27.04、二人の主体が同じ量の私的財を負担する場合の投資量は12.02と求まる。明らかに、ナッシュ均衡での公共財供給はパレート効率な水準より小さい。

ここまでの分析では、どの主体も自発的支払メカニズムに参加しないという選択肢が与えられておらず、メカニズムへの参加が強制されていた。しかし、Saijo-Yamato (1999) は、主体がメカニズムに参加するインセンティブを持たないような公共財供給メカニズムが数多く存在することを発見している。自発的支払メカニズムもそれらの一つである。いま、図1で表されているような二つのステージからなる参加ゲームを考えよう。まず、第1ステージでは、各主

体は、自発的支払メカニズムへ参加するか否かを同時に決定する。次に、第2ステージにおいて、相手が第1ステージで参加、不参加のどちらを選んだのかを知った後、「参加」を選択した主体は自分の投資数を決め、「不参加」を選択した主体は投資を全く行わない。もし二人とも参加すれば、二人は同時に投資数を決める。一人だけ参加した場合には、参加者のみが投資数を決定する。二人とも不参加ならば、公共財は全く生産されない。

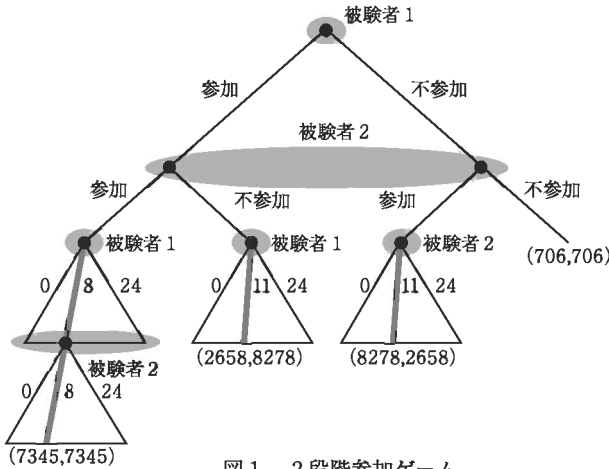


図1. 2段階参加ゲーム

ここで、「第1ステージで不参加を選択すること」は、「第1ステージで参加を選択した後、第2ステージで投資数ゼロを選ぶこと」と異なることに注意しよう。いま、あなたが参加を選択したとしよう。もし相手が不参加ならば投資数は必ずゼロであり、このことをあなたは投資数を決めるときにわかっている。これに対して、もし相手が参加を選んだならば、相手の投資数がいくらになるのかを知ることなく、あなたは自分の投資数を決めなければならない。たとえ、結果として相手の投資数がゼロだったとしても、これを事前に知ることはできない。



あなたの投資数

	0	1	2	3	4	5	6	7	8	9	10	11	12	
対戦相手の投資数	0	706	871	1072	1297	1536	1775	2003	2210	2386	2523	2615	2658	2648
	1	905	1127	1379	1647	1919	2183	2427	2641	2816	2944	3019	3039	3001
	2	1186	1465	1764	2072	2374	2658	2913	3129	3297	3411	3465	3456	3385
	3	1554	1888	2232	2575	2902	3202	3463	3675	3831	3925	3952	3911	3801
	4	2017	2401	2787	3160	3508	3817	4078	4281	4420	4488	4483	4403	4250
	5	2578	3010	3432	3831	4193	4507	4762	4950	5064	5101	5057	4934	4733
	6	3244	3718	4171	4590	4960	5272	5515	5681	5766	5765	5677	5504	5249
	7	4018	4529	5008	5440	5812	6115	6339	6478	6526	6481	6343	6114	5800
	8	4904	5447	5944	6383	6751	7038	7237	7340	7345	7250	7056	6765	6385
	9	5907	6475	6984	7422	7779	8043	8209	8271	8225	8073	7816	7458	7007
	10	7031	7616	8130	8561	8897	9132	9257	9270	9168	8951	8624	8193	7664
	11	8278	8873	9384	9800	10109	10306	10384	10339	10173	9886	9482	8970	8359
	12	9653	10250	10750	11142	11416	11567	11589	11480	11242	10877	10390	9791	9090

表1. 実験で配布した利得表の一部分（実験で配布した表にはグレイの部分はない）

実験では、被験者が選ぶ投資数を整数値のみに制約し、自分の投資数と相手の投資数によって自分の利得がどう決まるかを示す利得表を被験者に配った。私的財の初期保有量が24単位なので、ゼロ単位の投資数も含めると、利得表は25行25列となる。表1は、被験者に配布した利得表の一部を示している。いま、第1ステージで二人とも参加したとしよう。表1より、相手の投資数が8の時、自分の利得を最大にする投資数は8である。相手も同じことが成立する。よって、この部分ゲームでのナッシュ均衡は、各被験者が8を投資することであり、両者の利得は7345となる。次に、自分は参加したが、相手が参加しなかったとしよう。この時、相手の投資数はゼロなので、自分の利得を最大にする投資数11を選べば、自分の利得は2658、相手の利得は8278となる。二人とも不参加の場合には、両者とも706の利得になる。これらの第2ステージでの均衡結果を、参加・不参加の利得表にまとめたものが表2である。

		被験者2	
		$p_2$ 参加	$1-p_2$ 不参加
被験者1	$p_1$ 参加	7345 / 7345	2658 / 8278
	不参加 $1-p_1$	8278 / 2658	706 / 706

表2. 第1ステージでの参加の意思決定:部分ゲーム完全均衡

表2のゲームは「タカ・ハトゲーム」の一種である<sup>5)</sup>。従来、公共財供給の問題は「囚人のジレンマゲーム」として表現できると言われてきたが、メカニズムへの参加問題を考慮に入れると、より適切な表現方法はタカ・ハトゲームであることがわかる。このゲームでは、一人だけが参加する純戦略でのナッシュ均衡が二つ存在する。また、両者が参加率 $p_i$ を68%とする混合戦略でのナッシュ均衡も存在する。これら3つの均衡のうち混合戦略のみが、「進化的に安定な戦略（evolutionarily stable strategy, ESS）」となる<sup>6)</sup>。

このように、自発的支払メカニズムにおいて、二人とも参加することは均衡となりえず、相手が参加するならば参加しない方がよい。自発的支払メカニズムは単純で自然なメカニズムであるが、それすらも自発的な参加を誘発できないのである。前述したように、自発的支払メカニズムはパレート効率な公共財水準を達成できない。よって、たとえメカニズムにパレート効率性を達成することを要求しなくとも、各主体に自発的参加のインセンティブを持たせることは困難であることになる。

## 5. 参加ゲーム実験

実験では、20人の被験者が集められ、2人一組のペアを10組作った。各被験者はこのペアとなる相手と前節の参加ゲームをプレイする。実験は15回繰り返され、ペアは各回で変わり、同じ相手とは2度と対戦しない。各被験者は毎回異なる相手とプレイしていることは知っているが、実際に誰と対戦しているかはわからない。また、各被験者に同じ利得表を配り、相手も同じ利得表を持つ

5) 「参加」を「ハト」戦略、「不参加」を「タカ」戦略としよう。相手がハト戦略（参加）を選んできた時には、タカ戦略（不参加）の方がハト戦略（参加）より利得が高い。他方、相手がタカ戦略（不参加）の場合には、タカ戦略（参加）の方がハト戦略（不参加）よりも利得が高い。さらに、二人ともハト戦略（参加）を選んだ方が、タカ戦略（不参加）を選ぶ場合よりも、利得は大きい。

6) 2人対称ゲームにおいて、自分が戦略 $p$ を選択し、相手が戦略 $q$ を選んだ場合における期待利得を $\pi(p, q)$ とする。任意の他の戦略 $p'$ について、

1)  $\pi(p^*, p^*) \geq \pi(p, p^*)$  がかつ

2)  $\pi(p^*, p^*) = \pi(p, p^*)$  ならば  $\pi(p^*, p') > \pi(p, p')$

が成立するとき、戦略 $p^*$ はESSである（Maynard Smith-Price (1973)）。

ていることを周知した。さらに、被験者間の話し合いは一切認めなかった。

筑波大学と南カルフォルニア大学（USC）の学生を被験者として用いた参加ゲームの実験結果を見ていこう。図2は、実験における参加率（参加者の人数を全被験者数で割った割合）が各回でどのように推移したかを示している。もし進化的に安定な戦略による理論予測が正しいならば、実験で被験者の参加率は68%になるはずである。しかし、筑波大学での実験結果はこの予測とは大きく違っていた。実験の1回目では、40%に過ぎなかった参加率は、ラウンドが進むにつれて上昇していき、最後の方の回では85%から95%にもなっている。表2の参加ゲームでは、「二人とも参加する」のは部分ゲーム完全均衡ではない。ここで二人とも参加することを「協力」と呼ぶとすると、理論予測と異なり、協力が創発したのである。これに対して、USCでの実験においては、筑波の実験結果とは異なり、すべてのラウンドについて参加率は68%前後で推移しており、ほぼ理論予測通りの結果が得られている。

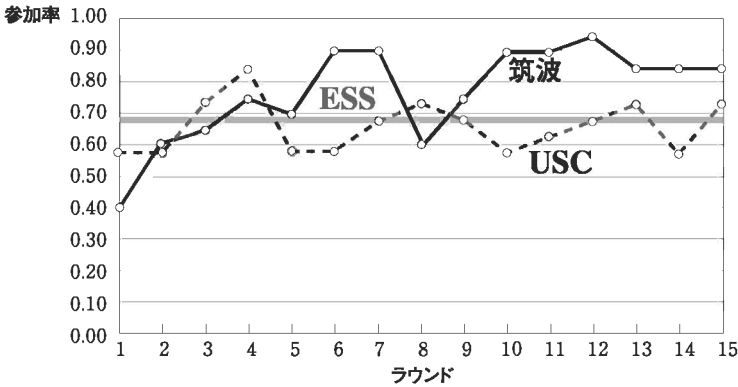


図2. 実験における参加率の推移

日本の実験では、なぜ協力が創発したのであろうか。相手が不参加の時、被験者が自分の利得を最大にする「自己利得最大化行動」をとるならば、投資数は「11」となるはずである。しかし、筑波の実験では、最も多く観察された投資数は「11」ではなく、「7」であった。これは、被験者が自分の利得だけで

はなく、相手がどれくらいの利得を得ているかを気にし、他人との比較で自己の満足度を考えて行動したからである。もし11を選べば、自分の利得は2658に対して、相手の利得は8278と自分よりかなり大きい（表1の1列目と1行目を参照）。他方、11の代わりに7を選ぶと、自分の利得は2658から2210と少し下がるものの、相手の利得は8278から4018へと大きく下げることができる。このように、自己利得最大化行動をとった時の利得に比較すると、自分の利得は少し下がるものの、相手の利得も大きく下がる「スパイト行動」がとられたのである。

参加しなかった被験者は、相手が参加し、自己利得最大化行動をとることによって得られる大きな利得8278を得ることを最初は目論んでいた。しかし、参加を選んだ相手は、自己の利得を犠牲にしてまでも、不参加者が大きな利得を得ることを防ぐスパイト行動をとるために、この甘い期待は実現しない。不参加によるただ乗りによって、大きな利得を得ることができないことを次第に多くの被験者は理解し、後半では多くの被験者が参加するようになった。つまり、「スパイト行動が協力の源泉」になったのである<sup>7)</sup>。

		2			
		参加	不参加	参加	不参加
1	参加	6494	6494	2349	5315
	不参加	5315	2349	706	706

筑波

		2			
		参加	不参加	参加	不参加
1	参加	7167	7167	2400	7279
	不参加	7279	2400	706	706

USC

表3. 5ラウンドまでのデータ平均値で作成した利得表

表3は、1ラウンドから5ラウンドまでで被験者が得た利得データの平均値をもとにして、参加・不参加の利得表を作成したものである。筑波での実験に関する利得表では、各被験者にとって、相手が参加するかしないかに関係なく、自分は参加した方がよい。スパイト行動によって、二人とも参加するのは均衡ではないタカ・ハトゲームが、二人とも参加するのが支配戦略となるゲームへと変容している。

7) スパイト行動は生物学の分野でも注目されている。Iwasa-Nakamaru-Levin (1998) は、大腸菌を出すバクテリアのスパイト的行動について吟味している。

これに対して、USCの実験に関する利得表は、数値は異なるものの、理論予測である表2のタカ・ハトゲームと基本的な利得構造は変わっていない。これは、USCの実験では相手が参加しない時でも、多くの被験者が自己の利得を最大にする投資数「11」を選択したからである。最初の5ラウンドまでで、一人だけ参加した時、投資数11が選ばれた割合はUSCでは88% = (14/16)であったが、筑波では37% (=7/19)にすぎない。他方、不参加者の利得を下げるスパイト行動（10以下の投資数）が選択された割合はUSCでは12% = (2/16)だが、筑波では63% (=12/19)もある。<sup>8)</sup> 米国の実験で多くの被験者がほぼ理論通りに行動したのに対して、日本の実験で被験者は理論の想定外の行動をとったのである。<sup>9)</sup>

## 6. 最後通牒ゲーム実験

ある一定の金額、例えば10万円を主体Aと主体Bに分配する方法として以下のやり方を考えよう。まず、Aが10万円の分配案を決めてBに伝える。次に、もしBがAの提案を受諾すれば、その通りに分配額が決まる。もしBが拒否すれば、二人の分配額は共に0円となる。これは「最後通牒ゲーム(ultimatum game)」と呼ばれている。各人が自分の受け取る額のみに関心を持ち、自己利得最大化行動をとるホモ・エコノミクスを仮定すると、Bは自分が1円以上もらえる提案ならば、拒否して0円となるよりはよいので、必ず受け入れるはずである。Aはこれを考慮に入れ、Bが受け入れてくれる提案のうち、自分の金額が最も大きくなる分配を提示する。つまり、「Bに1円、残りを全額自分がもらう提案をAが行い、Bがそれを受諾する」ことが、このゲームの部分ゲーム完全均衡となる。

8) 自己利得最大化行動をとった時の利得に比較すると、自分の利得は下がり、相手の利得が大きくなる「利他的行動」（12以上の投資数）が選ばれることは、筑波でもUSCでもなかった。

9) 東京都立大学とパディュー大学でも同じ参加実験が行われている。日本の都立大と筑波大に関しては実験データに有意な差はなく、また、米国のUSCとパディュー大学についても実験データの有意な差はなかった。しかしながら、日本と米国の間では、筑波とUSCの差と同様の違いが観察された。詳しくは、Cason-Saijo-Yamato (2002)を参照せよ。

しかし、この理論予測とは大きく異なる結果が実験では観察されている。実験では、分配する金額の40%~50%を相手に与える提案が全体の約2/3を占め、20%を下回る提案はほとんどない。また、自分の取り分が小さい不公平な提案ほど拒否する傾向が見られている。最後通牒ゲームについては、被験者の性別、年齢、学歴、計算能力、分配金額の総額と関係なく、上記のような実験結果が観察されている（Kagel-Roth(1995)、4章、Sigmund-Fehr-Nowak(2002)などを参照）。

ただし、これまでの実験の多くは西欧諸国、中国、日本など先進諸国で行われ、大学生が主な被験者であった。これに対して、Henrich et al (2004)は、さまざまに経済的・文化的条件の異なる小さな社会において被験者を集め、最後通牒ゲームの実験を行い、比較している。提案において相手の取り分が総額に占める割合は、典型的な西洋文明国では平均45%である。しかし、アマゾン川流域のマチグエンガ族でこの割合は平均22%とかなり低く、さらにほとんどの提案が拒否されることはなかった。この結果は、マチグエンガ族では、家族の単位を超えた社会的協力や交換がほとんど行われていないことが影響している。社会的制裁をあまり恐れず、他人の意見・評判に対して関心をほとんど示さないのである。部分ゲーム完全均衡により近い実験結果ではあるが、現代において非常に稀なタイプの社会の話である。

逆に、パプア・ニューギニアのアウト族では、相手の取り分が半分以上になる提案が多かった。さらに、自分の取り分が少ない提案だけではなく、取り分が多過ぎる提案も拒否する傾向もあった。一見すると、これは利他的で慎み深い行動のように思われる。しかしながら、実際は、贈り物をすることによって社会的地位を得ようとするアウト族での文化的伝統が結果に影響している。大きな贈り物をすれば社会的に優位な立場に立てる。また、贈り物を拒否すれば、相手に服従することを拒否したことになるのである。このように文化的な多様性は見られたものの、自己利得を最大化する行動をとるという予想とは実験結果は異なっているケースが多い。

では、なぜ先進諸国の実験では、公平な配分が多く提案されたのであろうか？考えられる理由としては、1) 公平な配分そのものが望ましいと提案者が考えたのか、もしくは、2) 不公平な配分の提案を行うと、拒否され利得がゼロになってしまう危険を提案者が回避しようとしたかであろう。このどちらであるかは、最後通牒ゲームの実験結果では明らかではない。そこで、一人が分配額の提案を行い、もう一人がその提案を必ず受託せねばならず拒否できない「独裁者ゲーム (dictator game)」に関する実験が行われた。独裁者ゲームの実験では、相手の取り分が分配総額に占める割合は、ほとんどの提案で30%以下となり、最後通牒ゲームの実験結果に比べてかなり低くなっている。相手に拒否されることがなければ公平な提案をしないので、公平な提案そのものが望ましいと考える人は少ないようである。

被験者は、自分の取り分が小さい不公平な分配案を提示されたならば、それを拒否し、自分の利得がたとえゼロになっても、相手が大きな金額を得ることを阻止する「スパイト行動」をとる。これゆえ、提案が拒否される危険を避けるために、公平な利得配分が提案されたと考えられる。いわば、「スパイト行動が公平性の源泉」となったのである。

## 7. なぜスパイトするのか？

不公平な提案の拒否という最後通牒ゲームにおけるスパイト行動も、相手が参加しないならば投資を少なくするという参加ゲームにおけるスパイト行動も、自己利得最大化行動をとった時の利得に比べて、自分の利得は下がるものの、相手の利得も下げる行動という点で、全く同じ種類の行動である。公平な分配、あるいは、投資への参加という社会規範 (social norm) を維持するために、自己犠牲を払っても、規範に従わない人を処罰しているのである。

しかしながら、実験は相手が誰かはわからない状況で行われている。最後通牒ゲーム実験では、ゲームは1回限りで2度と繰り返されない。参加ゲーム実験では、ゲーム1回ごとに相手が変わり、同じ相手とは2度と対戦しないように設定されている。よって、たとえ現在の相手が規範に従わなかったとしても、

スパイト行動によって処罰をしても意味がないはずである。それなのに、なぜスパイトするのであろうか？

この疑問に対する一つの答えを与えてくれるのが、進化モデルによる分析である。最後通牒ゲームの実験結果について、Nowak-Page-Sigmund (2000) は以下のような進化学的な説明を行っている。人類は何万年もの間、少人数集団の中で生活してきた。少人数集団の中では、お互いのことをよく知っていて、自分の判断がすぐにみんなに伝わってしまう。よって、一度少ない取り分を受け入れると、みんなが低い額を提案するようになってしまう。他方、少ない取り分を怒って拒否すれば、自分には高い額が提案されるようになり、将来の交渉に有利になる評判を得られる。少ない取り分を感情的に拒否した方が、自然淘汰の上で有利となるのである。人類の進化の過程で、交渉が1回限りで終わるケースはほとんどなく、我々は慣れていない。したがって、1回限りでの交渉でも、繰り返し交渉が続く場合でも、同じような行動をとってきたと思われる。自尊心が傷つけられたので、少ない取り分を拒絶したとも言えるが、進化学的には、自尊心を守ることは、将来に有利な評判を得るための仕組みと考えることができる。

一方、参加ゲームの実験結果をうまく説明するモデルは、まだ考案されていない。しかしながら、以下のGintis (2000, 2003) による進化学的な考え方が一つの手がかりを与えてくれる。彼の進化モデルでは、戦争、病気、飢饉といった社会の滅亡を招く危機に直面した時、社会規範から逸脱した人に対してスパイト行動をとり、罰を与える人々がたくさんいる社会の方が生き残りやすい、つまり、スパイト行動が適応度を高めるのである。<sup>10)</sup> 危機的状况では、同じ社会にいる人でも将来再び付き合うとは考えにくいので、利己的な人は相手に協力をしない。しかし、利己的な人を処罰する制度があれば、勝手な振る舞いを抑制でき、みんなが生き残りやすくなるのである。ただし、スパイト行動により罰を与える人自身がこの仕組みに気づいているわけではなく、単にスパイトする

10) ギンティスはスパイト行動を「strong reciprocity (強い互恵性)」と呼んでいる。



のが気持ちよいと感じているにすぎない。自らの適応度を高めるため遺伝子に組み込まれているのである。

ただし、この進化学的な考え方では、日本と米国の実験結果の相違、なぜ日本の方がスパイト行動をとる人が多いのかをうまく説明できない。歴史的に日本の方が米国より多くの危機に直面してきたので、スパイトする人の割合が多いとは言えないであろう。参加ゲーム実験結果について、日米の差の説明もできるような理論モデルを構築することは残された重要な課題の一つである。

## 8. おわりに

イギリス経験論を代表する哲学者であるデイヴィッド・ヒュームは、人間の本性を探究する方法として、ニュートン流の自然科学のやり方を用いることを主張した。つまり、経験（実験）に基づき原理（理論）を作成し、経験（実験）による検証を通じて原理（理論）をより確かなものへと鍛え上げてゆくという方法をとるべきである。この経験主義の立場に立つならば、人間の経済行動を探求する経済学においても、経済実験の結果を無視して理論構築はできないと言えよう。

実際、実験結果を説明できるように、いままでの経済理論では無視されてきた側面を取り入れて、従来の理論を拡張した新たな理論が生まれてきている。前節で述べた最後通牒ゲーム実験結果の進化学をベースにした説明は、その一例である。さらに、fMRIを使って、最後通牒ゲームをプレイする被験者の脳を調べる研究もなされている (Sanfey et al (2003))。経済合理性だけではなく、心理学、社会学、進化学、神経科学などにおける手法も取り入れて、行動経済学 (Behavioral Economics)、行動ゲーム理論 (Behavioral Game Theory)、神経経済学 (Neuroeconomics) などが誕生して来ている (Camerer (2003), Camerer-Loewenstein-Prelec (2004), Gintis (2004), Zak (2004)などを参照)。まさに、実験研究を通じた科学の統合化が始まっているのである。

参考文献

- Camerer, C. F., *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton University Press, 2003.
- Camerer, C. F., G. Loewenstein, and D. Prelec, "Neuroeconomics: How Neuroscience Can Inform Economics," *Journal of Economic Literature*, XLIII(2005), 9-64. Reprint in *Sistemi Intelligenti*, 2004, XVI(3), 337-48.
- Cason, T., T. Saijo, and T. Yamato, "Voluntary Participation and Spite in Public Good Provision Experiments: An International Comparison," *Experimental Economics*, 5 (2002), 133-153.
- Cason, T., T. Saijo, T. Yamato, and K. Yokotani, "Non-Excludable Public Good Experiments," *Games and Economic Behavior*, 49 (2004), 81-102.
- Gintis, H., *Game Theory Evolving*, Princeton University Press, 2000.
- Gintis, H., "Towards a Unity of the Human Behavioral Sciences," *Politics, Philosophy, and Economics* 3 (2004), 37-57.
- Groves, T., and J. Ledyard, "Optimal Allocation of Public Goods: A Solution to the 'Free Rider' Problem," *Econometrica* 45 (1977), 783-811.
- Henrich, J.P., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, eds., *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*, Oxford University Press, 2004.
- Iwasa, Y. M. Nakamaru, and S. A. Levin, "Allelopathy of Bacteria in a Lattice Population: Competition between Colicin-Sensitive and Colicin-Producing Strains," *Evolutionary Ecology* 12 (1998), 785-802.
- Kagel, J. H. and A. E. Roth, eds., *The Handbook of Experimental Economics*, Princeton University Press, 1995.
- Maynard Smith, J. and G. Price, "The Logic of Animal Conflicts," *Nature* 246 (1973), 15-18.
- Nowak, M. A., K. M. Page, and K. Sigmund, "Fairness Versus Reason in the Ultimatum Game," *Science* 289 (2000), 1773-1775.
- Saijo, T. and T. Yamato, "Fundamental Difficulties in the Provision of Public Goods: 'A Solution to the Free-Rider Problem' Twenty Years After," Osaka University, Institute of Social and Economic Research Discussion Paper No. 445, June 1997.
- Saijo, T. and T. Yamato, "A Voluntary Participation Game with a Non-Excludable Public Good," *Journal of Economic Theory*, 84 (1999), 227-242.
- Samuelson, P. A., "The Pure Theory of Public Expenditure," *Review of Economics and Stat-*

*istics* 36 (1954) 387-389.

Sanfey, A. G., Rilling, J. K., Aaronson, J. A., Nystrom, Leigh E., and Cohen, J. D., "The Neural Basis of Economic Decision Making: An fMRI Investigation of the Ultimatum Game," *Science* 300 (5626) (2003), 1755-58.

Sigmund, K., E. Fehr, and M. A. Nowak, "The Economics of Fair Play", *Scientific American*, January 2002.

Zak, P. J. "Neuroeconomics," *Philosophical Transactions of the Royal Society B*, 359 (2004), 1737-1748.

デイヴィッド・ヒューム『人性論，第2篇「情緒について」』岩波書店，1739年。

大和毅彦（2001）「実験経済学のすすめ」mimeo.