



Title	法政策と経済学 : 法政策の目的に関するゲーム理論的考察
Author(s)	町野, 和夫
Citation	新世代法政策学研究, 5, 287-314
Issue Date	2010-03
Doc URL	<a href="https://hdl.handle.net/2115/43728">https://hdl.handle.net/2115/43728</a>
Type	departmental bulletin paper
File Information	5_287-314.pdf



## 法政策と経済学

### －法政策の目的に関するゲーム理論的考察－

町野和夫

#### 1. はじめに：「多元分散型統御を目指す新世代法政策学」と ミクロ経済学・ゲーム理論

「多元分散型統御を目指す新世代法政策学」にミクロ経済学あるいはゲーム理論から何か貢献できるのかを考えるために、まず「多元分散型統御を目指す新世代法政策学」とは何かを経済学的視点で捉え直しておきたい。「法政策学」という言葉は経済学者には馴染みが薄いので、その定義を確認すると以下のようなものである。

- ① 一定の法目的の実現のために最も有効、能率的な法技術の体系を科学的に探究する学問。刑事政策学が最も代表的な学問分野とされる。
- ② 立法、政策決定に対する知識、技法を実践的立場から研究する学問。社会学、経営学等の意思決定理論を法的に再構成して用いる点に特色がある。

(有斐閣『法律用語辞典[第3版]』、2006)

このプロジェクトで使われる「法政策学」という言葉は主に①の意味で使われていると考えられるが、これは、ゲーム理論的には、社会的活動(個々の主体にとっては行動)というゲームの公式ルールである法律という制度の、設計方針・手法について考察する学問、と言い換えられるであろう。

社会的活動というゲームと言っても、社会には、立法、行政、司法、政治、経済(市場)、文化、社交、家庭生活等々、人間の活動の多様な側面に対応して多種多様な、かつ異なる次元のゲームがある。とくに立法、行政、司法は、それぞれが一つのゲームであるが、政府の一部門として、他

のほとんどすべてのゲームに影響を及ぼすことができ、なかでも立法は、行政や司法を含め、すべてのゲームのルールを決定できるゲームでもある。したがって、このプロジェクトで使われる「多元分散型統御」という言葉は、「一定の法目的実現のために」、立法が個々の分野のゲームのルールを設計するだけでなく、それらのゲームに影響を及ぼす行政や司法などのゲームのルールも含めた次元の異なる複数の多元分散的なゲームのルールの体系的な設計（統御）をすること、を意味すると考えられる。「一定の法目的」とは、たとえば知財法であれば科学技術の発展、環境法であれば環境の保全あるいは持続的発展、競争法であれば市場の競争環境の維持、といったものであろう。

所与の目的達成のための制度設計は、経済学ではメカニズム・デザインの問題であり、現在のマイクロ経済学、ゲーム理論にかなりの蓄積がある。多元分散的なゲームは、複雑なモデルになるかもしれないが、ゲームを定式化できる情報、（プレーヤの集合、戦略集合（即ち、それぞれのプレーヤの取り得る全ての選択肢）、利得、その他の「ゲームの木」を描くための情報）があれば、分析は可能である。しかし、本プロジェクトの問題意識に関するこれまでの議論の経緯を見ると<sup>1</sup>、まず上述の「一定の法目的実現のために」をどう考えるかに難しさが潜んでいる<sup>2</sup>。つまり「法目的」は、経済学的に明確な基準のある効率性の観点だけではなく公正や正義を含むと考えられるが、それらをゲーム・モデルにどう取り込めばよいのか、あるいは、そもそもゲームとして定式化できるのかどうかという問題である。公正や正義といった視点は、マイクロ経済学の理論モデルとしては定式化しにくい、法政策ではなくとも経済政策や制度を実際に考える際には避けて通れない観点である。そこで本稿では、まず次節で筆者の専門のマイクロ経済の分野で、公正や正義といった視点を含めた政策目的についての

<sup>1</sup> 「新世代法政策学研究」創刊号（2009年3月）の田村、藤谷、川濱、得津、の各論文、同第2号（2009年7月）の得津論文を参照。

<sup>2</sup> あくまでも理論上の難しさの比較であって、実際のメカニズム・デザインに際しては、無数にあり得る結果ごとの各プレーヤの期待利得、起こり得る全ての状況の確率分布の評価、そもそもどのようなゲームとして定式化するか、などについて合意することは容易ではない。

議論を整理し、第3節では、公正な結果が均衡となる（従来の経済理論モデルの均衡とは異なる）進化ゲーム・モデルの均衡の例を提示し、政策目的についての法学の考え方との共通点を議論する。それでも説明できない心理的要素について、第4節では人間の公平感についての人類学的根拠を与える研究例を、第5節ではさらに神経経済学的な裏付けを示唆する研究を紹介し、最後に第6節でそうした学際的研究成果を踏まえた上で、新しい法政策の目指すべき方向についての経済学的視点からの示唆を述べたい。

## 2. ミクロ経済学的経済政策：「市場の失敗」の補正と分配の平等

経済政策を、マクロ経済の安定や成長を目的とする財政政策や金融政策と、「市場の失敗（何を市場の失敗に含めるかは多少議論が分かれるが）」を補正する政策、分配の不平等を是正する政策に分けると、後の二つがマイクロ経済学的政策であり、そのうち「市場の失敗」の補正は効率性の観点から（総余剰やパレート効率という基準で）目的を明確に定義できるが、分配の不平等の是正の理論的に頑健な基準はない。実際の政策立案時には両者を明確に区分はできず、多くの場合、効率性と公平性の要素が混在するが、議論の明確化のために両者を分けて考えてみたい。

### 1) 「市場の失敗」を補正する政策

これが最も典型的なマイクロ経済理論で正当化できる政策（＝政府の介入）であり、その目的は、市場メカニズムが機能すれば実現したであろうパレート効率的な配分に近づく、という明確なものである。ここでは不完全競争、外部性・公共財、不完備情報の3タイプに分けて、タイプ別の政策のポイントを説明する。

#### A) 不完全競争

不完全競争の原因は以下の4つのうちの一つまたはこれらのうちの複数の組合せである。

- ① 財・サービスの異質性
- ② 需要者・供給者の数の少なさ
- ③ 不完備情報（これ自身「市場の失敗」の一つ）
- ④ 参入・退出の制限

不完全競争になると、プレーヤの戦略的行動によって非効率な結果がもたらされるので、競争法あるいは行政によって、完全競争であったら達成できたはずの総余剰を最大化する生産や価格の水準に近づけるための規制が正当化される。具体的には、価格や生産量そのものを直接規制したり、利得の上乗せというインセンティブによって望ましい生産量に間接的に誘導したり、といったメカニズム・デザインである。

### B) 外部性・公共財

あるプレーヤの経済活動が当該経済活動（ゲーム）の当事者（プレーヤ）ではないプレーヤの目的関数を変化させる場合である。当該ゲームのプレーヤたちは社会的影響を考慮しないため、フリーライダーや社会的ジレンマなどの問題、当該経済活動のアウトプットの過少供給（正の外部性の場合）や過剰供給（負の外部性の場合）という非効率が生じる。知財や環境など外部性のある経済活動に関連する領域で、法・行政による規制や、行政による公共財・サービスの供給などが正当化される。最近では行政だけでなくNPOなどによる公共財・サービス供給もある。コースの定理により、外部性を引き起こす側と影響を受ける側のどちらに権利を設定しても効率性は増加することが明らかにされているが、分配面では権利を得た側が多くの利得を得る。

### C) 不完備情報

情報格差によって、逆選択やモラルハザードなどの非効率的行動を招く。A)やB)と次元が異なる問題であり、多くの場合両者と併存する。情報格差を許認可や保証というメカニズムで公的機関が補うことによって、ある程度補正できることもある。また、オークション、インセンティブ報酬など民間によるメカニズム・デザインでの補正もあり得る。ルールによる規制と行政指導による規制のどちらがいいか、などの議論も含めてA)とC)の「市場の失敗」を補正する政策は基本的にはメカニズム・デザインの問題である。しかしB)のうち、公共財・サービスの公的供給の場合の負担をどうするか、あるいはコースの定理に基づいて権利の設定を行う場合には、効率性だけでは解が出ない<sup>3</sup>。

<sup>3</sup> 公共財の効率的な生産量決定のため真の限界効用を引き出す様々なメカニズムは考案されているが、実行可能性も考慮すると完全な解決方法はない。

なお、政府・行政の介入が正当化されるのは、政府に資源の効率的な配分を行う動機と能力があることが前提されているためであるが、その前提が満たされないこともある。これを政府の失敗と呼ぶが、政府に任せられた場合の主な問題点とその原因は、

- ① 非効率性（競争がないため効率化する誘因がない）
- ② レント・シーキングによる不公平（プレーヤの利害が異なる政治的決定には必然的に決定を自らに有利になるよう働きかけるインセンティブがある）
- ③ フリーライダー、モラルハザード、逆選択など（不完備情報の問題は完全には解決できない）

といったものである。①と③は効率性の問題、②は分配の公平性（と正義）の問題である。

「市場の失敗」にしても政府の失敗にしても、効率性を改善しようとするメカニズム・デザインの問題と分配の公平性の問題が混在するが、以上のように、それをある程度区分けできる。効率性の問題は、理論上はメカニズム・デザインの問題として解決されるとすると、残るは分配の公平性の問題である。

## 2) 分配の不平等に関する経済学の三つのアプローチ

市場メカニズムはパレート最適な資源配分を実現するが、民主的な社会でどのような分配が望ましいかについての社会的合意を保証するメカニズムは存在しない（アローの不可能性定理、後述）。社会的厚生関数は多数のパレート効率的な資源配分の結果の中から一つを選ぶメカニズム（社会的な順位付け）であるが、アローの定理は、社会の全員が最適だと思う何か一つの社会的厚生関数が存在することはない、と言い換えられる。したがって、分配の不平等を是正する政策は、「市場の失敗」を補正する政策が効率性という共有された一元的な基準で評価できると異なり、共有された基準はない。しかし、政策は現実決めなければならない。これに対する従来の経済学者の立場は大きく分けて次の三つである。

- ① 分配の問題を考えるのは政治の問題で、経済学の対象ではない。しかし分配の目標（あるいは基準）が与えられればそれを最も効率的に達成するメカニズムを考える。

- ② 政策決定に係わる政治家、官僚、利益団体、有権者もそれぞれで異なる目的の最大化を図って行動すると考え、政策決定過程をゲームとして分析する。（公共選択論や(新)政治経済学）
- ③ あくまでも社会の望ましさに対する何らかの基準を追求する。（厚生経済学や社会的選択論）

本稿では法政策を含む政策一般の目的に含まれる公正についてのゲーム理論的（とくに経済学的）説明を試みているので、①の立場ではない。ただし、①の立場を取ったとしても、分配に関する方針や、（上述の市場の失敗を補正する何らかのメカニズムによって）公共財・サービスの供給量が政治的に決まった後、いかにして歪みのない（＝効率的な資源配分にできるだけ近い）「公平な」課税をすべきか、といった分析をすることは財政学や公共経済学の重要なトピックである。

公共選択論に代表される②の立場では、個々のプレーヤの（必ずしも経済的だけではないが）自己利益を前提としており、外生的に「公正」あるいはそれに類似の価値観を持ったプレーヤを仮定する場合があっても、公正という概念自体を解明しようとはしない。しかし、法政策においても、それに関わる複数のプレーヤが戦略的に行動することを想定しなければならないので、公正の意味について考える場合にもこうした政策決定過程のゲーム理論的分析を理解しておく必要がある<sup>4</sup>。

上述のように、この分野では、広義ではあるが自己利益を前提としたプレーヤを前提として政治的意思決定をモデル化している。とくに公共選択論では、レント・シーキングによる政策的意思決定の歪曲という非効率や、政治家が選挙を意識して歳出に応じた負担増を有権者に求めない（従って赤字に歯止めがかからない）「ソフトな予算制約」という問題を重視し、これらの問題に対してレント・シーキングやソフトな予算制約を認めない

<sup>4</sup> 経済学でも、創成期には、アダム・スミスなどの古典派の学者は現実の政策についての政治経済学的議論を行っていたが、その後数理的な理論の精緻化と共に、経済学と政治学の分化、専門化が進んでいき、マルクス経済学を除けば漸く20世紀半ばにブキャナンやタロックなどによって創設された公共選択論において、経済学でも政治的意思決定過程の分析が始まり、さらに近年では、ゲーム理論や経済動学を取り入れた（次に述べる）新しい政治経済学も発展しつつある。

一般ルールの立法化（立憲的改革）を提案する。なぜなら、特定の改革で既得権益を損なう層は、政治家への圧力のための時間的金銭的費用をかけられるので、数は少なくとも反対の声は大きい。多数の有権者にとって、個別の改革によって受ける利益は小さく、既得権益層の反対を押しきってまで、改革のために費用をかけ努力するインセンティブは生じないからである。また、立憲的改革は、総論的、抽象的であり、有権者のどの層にとっても、改革によって自分が利益を受けるか損失を被るか分らないという段階での提案なので、合意が得やすいとも考えられる。具体的な例としては、均衡財政の立法化、貨幣供給のルール化、公共部門の一律削減などがある。しかし、現実には総論の段階でも、ある程度の利害計算がなされるので、立憲的改革であっても既得権益層を説得して合意を得るのは困難である。

新しい政治経済学では、公共選択論に加えて、選挙制度や政党制度など政治制度の分析が中心の合理的選択論（政治学）、マクロ経済学の動学的モデル（例えば複数期間の効用の最適化）などの手法も使い、現実の経済政策の立案、決定、実施を、それぞれの段階で関わりのあるプレーヤたちのゲームとしてモデル化する。そのモデルの中で、プレーヤの戦略（行動）やゲームの均衡を考察し、ゲームのルールである制度や慣習がプレーヤの行動や均衡に与える影響や、特定の制度の下で、どのような政策が選ばれ易いか、そのときの社会的厚生はどうなるか、などを考察する。しかし新しい政治経済学でも、政治的なプレーヤの目的は消費者や企業などの市場のプレーヤと違って定量化しにくいし、ゲーム・モデルを使うときは、多様なルールや複雑な政治環境をかなり単純化しなければならない。政策決定に関わる個々のプレーヤのインセンティブは、定量化しにくいだけでなく従来の経済学で前提としてきた単純で利己的な合理的「経済人」という枠組みを超える。上述の公共選択論で提唱されている立憲的改革にしても、それがうまくいくには、短期的な経済的利益だけを求めない広い意味での利己的動機（長期的帰結を考慮する繰り返しゲームや、非対称情報下での評判の確立やシグナリング）や、リーダーシップ、倫理的行動を考えなければならないが、まだ例示的で直観的な

説明に止まっている<sup>5</sup>。こうした「経済人」以外の意思決定主体は、次節や第5節で説明するような行動経済学や進化ゲーム・モデルなど、最近になってマイクロ経済理論の中でも本格的に組み入れられている。しかし、②の立場にとどまる限り、個人的な行動の動機が経済学的なものであろうがなかろうが、それを前提として社会的な意思決定の仕組み・過程とその結果を分析するものであって、結果に対する価値判断についてはパレート効率性以上のものは行わないか、行ったとしてもそれは分析者の主観的な意見に過ぎない。

一方③の立場は、ベンサムなどの功利主義の影響を受けたピグーに始まる「旧」厚生経済学においては、個人間比較可能な効用という現在の経済学では受け入れられない前提が必要な、最大多数の幸福を目指すナイーブなものであった<sup>6</sup>。その後個人間比較が不可能で序数的な効用を基にした社会的な評価基準の構築を目指した「新」厚生経済学は、バークソンによって導入されサミュエルソンによって精緻化された社会的厚生関数に辿りついた<sup>7</sup>。しかし、この関数が体現する価値観や形成プロセスは外生的に与えられるため、社会的価値判断に関する上述の立場としては①に退いたことになる。これに対して Arrow (1951) は、社会的厚生関数<sup>8</sup>の形成ルールが満たすべき最低限の望ましい条件を設定し、それを満たす社会的厚生関数が存在するかという公理的方法をとることで、社会的厚生関数の内容を研究対象とする③の立場に踏みとどまり、(i) 定義域の広範性、(ii) パレート効率性、(iii) 情報的効率性（選択に関係しない他の選択肢からは影響されない）、(iv) 非独裁制、という四つの最低条件を満たすような

<sup>5</sup> この点に関して、公共選択論の創始者の一人であるブキャナンは、合理的リーダーシップや経済的起源に基づく倫理という概念を提案している。合理的リーダーシップや経済的起源に基づく倫理については町野(2008)及びその参考文献を参照。

<sup>6</sup> 旧厚生経済学、新厚生経済学、アローの一般不可能性定理については例えば鈴木(2006；鈴木・長岡、花崎編第1章)を参照。

<sup>7</sup> もう一つの重要な流れとして補償原理の導入によるアプローチもあるが、社会的評価ルールとしては理論的欠陥があることはよく知られている。例えば、Suzumura (2002) 参照。

<sup>8</sup> アローの「社会的厚生関数」はバークソン＝サミュエルソンのそれよりは広義である。詳しくは Suzumura (前掲注7) 参照。

社会的評価形成のルールは存在しない、という「アローの一般不可能性定理」を証明し、これが社会的選択論の出発点となった。アローの4条件は非常に弱い基準ではあるが、パレート効率性という個人の効用（とその全体としての厚生）という経済学的評価尺度を前提としており、社会的な望ましさを探究する③の立場としては限定的とも言える。ある意味でこの限定性を問い直したのが Sen (1970) の「パレート派リベラルの不可能性定理」である。彼はパレート効率性という緩やかな厚生主義的要請でさえ、最小限の個人的自由の社会的尊重（最低限のリベラルさ）と矛盾する場合があることを示した。これをきっかけに Sen (1985) の福祉の経済学の「潜在能力」を初め厚生以外の様々な社会的選択の基準についての考え方が提案されてきた。政治・法哲学における Rawls (1971) の正義論の「社会的基本財」、Dworkin (2000) の平等論の「資源」、などもそのような非厚生主義的基準とみなすことができる。

なお、アローが分析から（意識的に）除外していた各メンバーの戦略的行動の問題（即ち、社会的評価形成のルールが決まれば、結論を自分に有利な方向に導くため自分の選好を偽る誘因が生じるという問題）に対する研究が、その後メカニズム・デザインの理論としてハーヴィッツやマスキンらによって進められていく<sup>9</sup>。メカニズム・デザインの理論自体は、社会的厚生関数に相当するような目的やその形成ルールが与えられた上で、それを実現するインセンティブを与えるメカニズムをデザインする、あるいはその存在可能性を議論する理論であるため①の立場に立っていると見えるが、この分野で使われる理論モデルは、その後センなどによって展開される現在の社会的選択論の有益な分析手法にもなった。それ以上に重要なのは、メカニズムのデザインによって遂行可能な均衡とその条件が明

<sup>9</sup> メカニズム・デザイン理論の他の源流はハイエクとラングなどの社会主義経済計画論争（あるいは計画経済論争）と「市場の失敗」が起こる環境での効率的メカニズムの追及である。前者は社会主義経済で市場メカニズムが達成するような効率的な資源配分は計算可能かという問題をめぐる社会主義経済計画論争であり、論争自体は社会的選択論と同様否定的な結論に収束する中で、メカニズム・デザインという有益な理論を生み出す一つのきっかけとなった。詳しくは、Suzumura (前掲注7)、西條・大和 (2006；鈴木・長岡、花崎編第2章) を参照。

らかになることによって、どのような条件が揃えば何を社会的な目的と心得るかというフィードバックが可能になったことである。このことによって、アローやセンの否定的な定理とは異なり、より実践的な政策的含意を導出できることもある。

以上、望ましい分配のあり方という観点から、政策目的の判断・評価基準に対する経済学の立場を三つに分けて整理したが、次節では、この経済学の議論と比較しながら、法政策の目的やあり方を考えてみたい。その上で両者に共通の分析枠組みを提供するモデルとして、進化ゲームを応用した規範形成モデルを紹介する。簡単に概略を説明すると、限定的ではあるが合理性を持った人間の試行錯誤的選択の累積の結果として規範が形成される、という立場である。これを、近年のミクロ経済学やゲーム理論の応用としての比較制度分析の立場から説明する。さらに第4節以降では、この進化ゲーム・モデルでも説明できない規範形成の心理学的要素の科学的説明として、まず第4節で人間の平等という価値観の進化的起源を説明する一つの生物進化論の学説を紹介し、第5節で、そのような説の行動経済学や神経経済学の最近の研究成果に基づく説明を試みる。

### 3. 法の評価基準と法の形成の進化ゲーム的説明

現代の「法と経済学」あるいは「法の経済分析」の代表的立場の一つとして、政策担当者選ばれた社会的厚生関数を基に厚生という基準のみで法ルールを評価すべきであるという主張（Kaplow and Shavell (2002)）がある。これは、法学者の考える一般的な評価基準、例えば Dworkin (1986) の考える法体系の純一性（integrity: 過去の政治的決定や、それらが前提とする正義や公正の諸原理から導かれる規準による体系としての整合性。新しい状況への共同体としての整合的な変更も可能。）とは対照的である。前者は単一の「客観的」基準で法ルールを評価するのに対して、後者の法的判断は総合的である。法の形成に関しても、前者であれば、効率的な無数の配分のうち、何らかの方法で決定された社会的厚生関数を最大化するような配分を実現するように法を設計することになり、後者であれば、過去の法体系との整合性を保ちつつ、その前提である共同体の諸原理に基づき、新しい状況にも対応して法を作っていくことになる。こうした両者の

違いの解釈として、常木 (2008) は、Kaplow and Shavell の主張は、Hart (1961) の「外在的視点」を体現した法外在的な規範と位置付け、Dworkin の主張は、Hart の「内在的視点」を強調したものと位置づける<sup>10</sup>。

Kaplow and Shavell の「外在的視点」にしても Dworkin の「内在的視点」にしても、ゲーム理論的に言えば、ゲームのルールである法を評価したり、法を意識的に設計したりする際の基準である。どちらの基準も、個人的利得の他に、社会的な結果に対する評価やプロセスについての評価も黙示的に含んでいるので、個人的利得以外の評価に関して社会のすべてのメンバーが納得できる基準は作れないであろう。前者では、社会的厚生関数の選択を委ねる政策担当者を決めるにあたって、誰が相応しいかを社会全体で合意することは難しい。後者では、まず“integrity”とは何かについて法学者の解釈は異なるであろうし、それについて抽象的なレベルで合意できたとしても、議論の分かれる具体的事例への応用に際しての解釈・判断は分かれるであろう。

したがって、Kaplow and Shavell の議論は、社会にとっての望ましきの基準を追求している点では、前節で述べた分配問題に対する第3の立場（厚生経済学や社会的選択論）にも見えるが、社会的厚生関数の選択を政策担当者に委ね、しかも担当者の決め方は、何らかの社会的厚生関数を体現しているという暗黙の条件を示しているだけで、結局は政治的に決まると推測されるので、第1の立場の一種であり、第2の立場での分析も可能であるが、第3の立場ではない<sup>11</sup>。Dworkin (1986) の“integrity”の議論も、既存の法体系に内在する共同体の正義や公正の諸原理を前提として、（法が道徳的にどうあるべきかではなく）法とはどのようなものである（はず）か、という問題意識に立っており、（そもそも経済学者ではないので上述

<sup>10</sup> 外在的視点と内在的視点は、それぞれ external point of view, internal point of view, に対応し Hart (1961) の邦訳では「外的視点」「内的視点」(2009版98頁)であるが、ここでは常木 (2008) の表現に従った。

<sup>11</sup> しかし、一旦社会的厚生関数が選択されたら、その中に公正に関するその社会の価値観は組み込まれており、その後で下手に場当たり的に公正や正義を考慮した分配のルールを追加適用することは、社会全体にとって望ましくないという Kaplow and Shavell の議論（証明：Kaplow and Shavell, 2001）の重要性は損なわれていない。

の3つの立場に分類するのは無理があるが) 第3の立場ではない。第2の立場での分析は可能であるが、第1の立場とも異なる。

経済学者でこれに近い立場を探すとすると、むしろ「過去から蓄積された経験に依拠して法の整合性が保障され、それが司法による法の適用を通じて法体系へ組み込まれると考えるので、法が立法に先立って存在すると考える点では Dworkin と共通する (常木(2008))」 Hayek<sup>12</sup>であろう。Hayek は、ルール (制度) を過去から蓄積された経験に依拠して自生的に秩序が形成される均衡であるとし、意識的に設計するべきでないという立場である<sup>13</sup>。したがって、「外在的視点」からの、(政府による) 意図的な制度設計に対しては否定的にならざるを得ない。実は Dworkin の言う integrity のある法体系も、慣例主義のように、過去の先例がない場合には従うべき指針がなくなるものでもなく、道具主義的なプラグマティズムのように過去の経緯を (戦略的な必要性がある場合を除いて) 全く無視するわけでもなく、法体系の整合性を保ちつつ徐々に新しい状況に対応して進化していく体系である。したがって、Hayek の自生的秩序形成の原理を解明する作業と、個別的ルールを integrity のある体系に相互に関連付ける黙示的な原理を解明する作業は、それほど違わないのではないだろうか。そこで本節の残りの部分では、Hayek の自生的秩序形成のモデル化と言える進化ゲーム・モデルの応用例の一つを見ていくことで、両者に共通する原理を探す端緒としたい。

自生的秩序は、慣習や規範という非公式制度とも言え、ゲーム理論的表現をすると進化ゲームの均衡である。進化ゲームの均衡は限定合理的なプレーヤたちがゲームを繰り返して学習を重ねて行くことで最終的に到達した均衡である。それは多くの場合、囚人のジレンマや共有地の悲劇が起こりうる状況で、それを回避するような非利己的で社会的に効率的な均衡と同じものになる。こうした規範形成過程の説明にもなるような進化ゲームの理論モデルの説明はゲーム理論の教科書に譲るが<sup>14</sup>、その成果を一言でいうと、限定合理的行動が非利己的均衡としての慣習や規範を形成する

<sup>12</sup> Hayek (1973) 邦訳 (1987) 「第4章 法概念の変遷」参照。

<sup>13</sup> 例えば同上「第1章 理性と進化」参照。

<sup>14</sup> 例えば大浦 (2008) 参照。

確率論的可能性を証明し、いくつかの形成過程のタイプを理論的に示したことである。

進化ゲームによって規範形成を説明するために使えそうな理論モデルの研究は数多くあるが、応用としては個別の制度や問題の単純化したモデルの分析に終わるものがほとんどである。ここでは体系的な応用を試みた数少ない研究として以下で説明する比較制度分析による (制度としての) 規範形成の説明を検討してみる。

規範を一つの制度として見ると、制度の変化について、例えば政治学者の Knight (1992) は、分配上不利なプレーヤが制度改革へのインセンティブを持つことを制度の変化の主要な原因として挙げ、逆に、あるプレーヤが均衡から逸脱した戦略 (行動) をとると、他のプレーヤたちは期待していた値以下の利得しか得られず失望する (従ってそれを防ぎたいと思う) ことを制度の維持の力として挙げている。叙事的説明としては理解できるが、モデル化しようとする、プレーヤたちの利得が徐々に累積して変化することによって、あるいは彼らの効用関数自体が変化することによって、繰り返されるステージ・ゲームそれぞれが異なるものになってしまうため、分析が困難になる。従来の慣習・制度の進化ゲーム・モデルの枠組を拡張して、こうした問題の一つの解消方法を提示したと言えるのが、比較制度分析である。代表的研究者である Aoki (2001) の主観ゲーム (subjective game) のモデルは、上記の問題に対して、戦略集合や利得や戦略選択ルールの固定性を仮定しないことで解決を試みた。従来のゲームでは所与であった、プレーヤの戦略集合、利得、効用関数を、主観ゲームでは、無数の戦略集合の内、ある認識された (activated) 部分集合がその時点での戦略集合であり、利得も不完備情報の下での利得予想と見なしたのである。このように、戦略集合の固定性を仮定しないことで、環境変化に伴って戦略集合の見直しを行うことが、モデルの内部で可能になった。

主観ゲームの解釈は以下のようなものである。大きな技術革新、国際的な競争の激化、長期的不況、政治構造の変化、戦争、関連する他の制度の大きな変化、などの急激な環境変化の際には、それらが均衡利得に与える累積的な影響のため、個々のプレーヤにとっては現実と予想のギャップが生じやすい。そのギャップがある大きさ以上になると、どのプレーヤもゲームの戦略や構造を見直す。その結果、従来は戦略集合として認識されて

いかなかった戦略が選択肢として認識されるようになり、その中あるいは従来最適ではなかった戦略の中から新しい最適戦略が見つげ出される。戦略集合の見直しや新たな最適戦略の採用をするプレーヤの数が一定の割合以上に増えると、均衡も変化し新たな制度が形成されることになる。環境変化とそれに対する新たな均衡変化が連続的だとみなされれば、それは新制度ではなく制度変化と呼んでもいい。

急激な環境変化が生じた場合、従来の戦略が不均衡になることは多くのプレーヤに明らかであるから、環境変化によって生じた特定の問題に対処するため、多くのプレーヤが同時期に集中して新たな最適戦略の発見に努める。その過程では、競争というメカニズムによって、最適な戦略を発見した主観ゲームの均衡が残ることになる。このようにして到達した均衡を断続均衡（punctuated equilibrium）と呼ぶ<sup>15</sup>。断続均衡とは生物学的進化の考え方の一つで、徐々に進化するのではなく、長期間一つの均衡に留まったあと一気に進化を遂げて次の均衡に到達し、また長期間その均衡に留まる、というタイプの均衡である（Dennet, 1995）<sup>16</sup>。

Aoki モデルは、制度変化は当初から存在していた複数均衡の間の移行であるという従来の進化モデルの想定を否定し、それは新たなルールの下での均衡だという発想を打ち出したことと、ステージ・ゲームのルールの固定性を捨てることで、その過程を内生的にモデル化したところに大きな貢献がある。しかし、変化を望む動機をモデルの中にどう組み込むかとか、アクティブとされていなかった戦略をどう発見するかという点は、理論的に曖昧さが残る。これらの点については、Aoki は企業化精神やイノベーションを念頭においているようであり、公共的なゲームにおいては上述の

<sup>15</sup> Aoki (2001) の邦訳 (2001) では、断続平衡という用語を使っているが経済学での equilibrium の定訳は (物理学では平衡だが) 「均衡」なので、本稿では断続均衡とした。

<sup>16</sup> これは必ずしも戦略集合の非固定性を必要とするわけではない。Young (1998) の確率的進化ゲームでも、ほとんどの期間社会状況は均衡の近くにあり、異なる均衡間を移動する期間は短いという性質を punctuated equilibrium effect と呼んでいるが、Aoki モデルでの特徴は、戦略集合の非固定性、均衡変化の文脈依存性と戦略集合探索努力の集中であり、同じ用語を使っているもののニュアンスは異なる。

Knight (1992) の分配上の不均衡や、公共選択論の「合理的リーダーシップ」の議論が参考になりそうであるが、モデル化は今後の課題として残されている。

進化ゲームによる規範形成の理論モデルで捉えられない現実の曖昧さは、上のような非経済的な動機や気づきなどの心理学的要素である。これらは、従来の経済学のモデルでは考慮されなかったり、進化ゲームによる利他性、互酬性、集団志向などのように、効用関数 (の一部) としてモデル外から与えられたりしてきた。このような非利己的動機は「社会規範の内面化」という説明もできるが、例えば、法の経済分析の分野で進化ゲームの研究を進める藤田・松村 (2005) は、それは一般的な心理学的説明に逃げ込むものであり、法の経済分析としては排除するべきであるとしている。しかし、「社会規範の内面化」という一般的に受け入れられている心理学的説明こそ、その原理が (経済合理的に) 明らかにされれば自生的秩序形成の原理説明の大きな前進ではないだろうか。次の二つの節では、この点について人類学や神経科学 (或いは脳科学) の知見も借りて検討したい。

#### 4. 倫理規範の進化的起源

従来の進化ゲーム・モデルでは、利他主義、互酬性などの既に何らかの規範が内面化されたプレーヤの存在を仮定していたが、ここではそのような目的関数あるいは選好が生じてきた過程のモデル化を検討する。町野 (2003) では、社会のルールとしての倫理規範を大人が子供を養育する社会化過程と、ある行為が社会のルールであるべきだと提案者が社会 (共同体) に説得する過程に分け、倫理的規範の形成過程のモデル化を試みた。前者では、とくに大人のインセンティブを厳密に考えることで、従来の学習モデルとは異なるゲーム構造を提示した。後者では提案者のイニシアティブの重要性を明らかにし、従来の動学的進化ゲームに欠けていた倫理性発生のモデル化を試みた。

後者の倫理規範を提案する側の動機について考えると、利己的なものもあるが、倫理的なものもある。それが倫理的なものだとすると、それを教えた者の動機は何か、そしてそれを教えた者の動機は… と、因果関係に

関する議論は無限に遡る可能性がある<sup>17</sup>。従って、そうした無限の連鎖を断ち、倫理規範の合理的形成の論理を提供しなければならない。町野（2003）で注目したのは、何らかの自己犠牲的行動を他の共同体メンバーに説得する、提案者の存在である。説得された行為が共同体にとって有益であることが実証されれば、この行為が次第に「倫理的行為」として認識され、その認識が共同体に浸透していく。このように、ある種の倫理的行為が集団の中で一定の割合で存在すれば、それが慣習・規範化することは、これまでの進化ゲームの研究成果から明らかである。しかし町野（2009）では、さらに遡って、なぜそのような利他主義的、互酬的、あるいは集団主義的提案を行うような選好を持つかという問い、言い換えると、人が何故ある行為を「善い」ものとする（或いは、ある行為を「悪い」ものとする）のかという問いに対して、そうした善悪に関する選好を、人間の社会的動物としての歴史あるいは民族的な歴史や個人的経験が積み重ねられて形成されたものと考えてみた。社会的動物としての歴史とは、自然淘汰が結果として生物学的な意味で「利他的」行為へと向かわせるよう人間社会に組み込まれた規則であり、我々の感情は、社会的に役立つことで是認された行為が「正しく」、否認された行為は「悪い」という感覚によって裏打ちされていると考えられる<sup>18</sup>。

人間の社会的動物としての本能については、近年の類人猿の研究から、類人猿が利他的、互酬的、あるいは集団主義的行動を取るのには、親子間の愛着というような一般の動物としての本能からだけでなく、社会的動物として生存のために集団に所属することが不可欠なことから、社会化によって集団の様々な規則（食物の分配や交配に関する規則など）に従っている事が明らかにされてきた<sup>19</sup>。規則を守らないメンバーに対する攻撃・制裁行動が「感情」を引き金にして引き起こされることも観察されている。彼らは自分と他者を区別して、お互いの行動を予測し、ある種の契約・信頼

関係を築けるので、他者の行動に対する予測が裏切られたときに、裏切ったものへの「怒り」が生じ、それが攻撃へと繋がる。しかし一方で、集団を維持することの高い必要性から、個々の対立が集団を崩壊させないような仕組みも内蔵されている。対立する当事者以外が和解を取り持ちたり、和解の直後に集団の他のメンバーたちが「喜ぶ」ことによって、彼らが他者の気持ちを予測して「共感」したり「同情」したりする能力があることも観察されている。

しかし、この類人猿の行動を規制するルールは人間の倫理規範とはかなり距離がある。その違いを、人間の主要な倫理的規範である平等の観点から考察するために、町野（2009）では、動物行動学・人類学の平等主義の進化的起源についての一つの有力な見方である Boehm（2004）の議論を紹介し、それを踏まえた上で、倫理的規範が集団に植えつけられるメカニズムのゲーム・モデル化を試みた。

Boehm（2004）はチンパンジーと非定住型狩猟・採集民との比較によって、平等主義的行動における両者の共通点と相違点を明らかにし人間社会の倫理的な進化的起源について考察している。彼の研究は大きな獲物を、その猟に貢献したかどうかに関係なく共同体全員に平等に分配する慣習に関するものである。チンパンジーも集団で狩をし、その獲物を分けるが、分配の仕方は食糧の豊富さなど様々な状況によって異なる。Boehmによると採集民の大きな獲物には共有財産という概念が当てはまるが、チンパンジーが共同の狩で得た獲物はあくまでも私的所有物としての性格を持っている。

人間も類人猿も大きな獲物を共同体のメンバー全員で平等に分け、人間はそれが制度化されているところに、平等主義という人間の倫理的性質が進化的起源を持つことが推測される。Boehmによるとこの分配行動の生物学的理由は、家族単位ではたまにしか獲れない大きな獲物を、自分たちが獲った時に他のメンバーに分ける代わりに他のメンバーが獲った時には自分たちにも分けてもらうということで、家族の食糧（タンパク質）の摂取が著しく変動する可能性を小さくすること（変動削減（variance-reduction）理論と呼ばれる）である。これを慣習化・規範化できた集団が、（12.8万

<sup>17</sup> 尾崎（2009・209頁）の「正統な規範ないし価値自体を究極的に根拠づける規範ないし価値をめぐるアポリア」に対応する。

<sup>18</sup> 進化論や社会生物学の倫理学への影響の歴史については内井（1996）を参照。

<sup>19</sup> 類人猿の社会的行動についての動物行動学の研究については de Waal（1996）を参照。

年前から7.2万年前にかけての) 氷河期を生き抜いてきたと考えられる<sup>20</sup>。また、その過程で脳が(従って社会的な関係を築く能力も) 発達し、ボスの力をその他のメンバーが共同して抑え込むという平等主義を実現する政治的構造が可能になったとも推測されている。

Boehm (1999) は、こうした(平等主義という) 倫理観を持つ道徳的共同体が成立するためのいくつかの前提条件(進化の準備段階)を挙げている。

- ① 支配したい、支配されたくないという性格
- ② 武器の発見・発明(これによって身体の大きさや強さの重要性が低下した。犬歯の退化や体毛の減少、二足歩行などとも関連すると思われる、50万年(2万世代)前までに変化したと考えられる。)
- ③ 知性の発達(多数の提携を形成する政治的知性と、(資源に余裕があれば障害を受けた者も世話するが、限りがあれば見捨てる) 合理的なセーフティーネットを構築する保険数理的知性。これらは10万年前から20万年前には脳の拡大に伴って発達したと思われる。)
- ④ コミュニケーション能力(言語、情報網、については類人猿も一定の能力はあったが、彼らの行動の選好は道徳規範として明確にできるほどのものではなく、言語と言えものは約10万年前に人類によって完成したと考えられる。言語がなければ協力、寛大さ、利他性などの複雑な概念を繰り返し説いたり教えたりできない。)

彼によれば平等主義的な共同体を維持するには、それに反するメンバーが出てきてから対処するのでは遅く、独裁者が現れないような不断の警戒が必要であり、共同体のメンバーが「道徳」的で相互に監視しあうことが有効である。

進化生物学者や人類学者は、従来、自然淘汰は縁故者を助ける性質を残すように働くが、利他的な性質を残すようには働かないと考えてきた。個体が集団の生存競争のための犠牲になるという集団選択(群淘汰)の理論も現在は否定的な評価が一般的である<sup>21</sup>。集団間の利他的行為は血縁選択

(淘汰) や強制、長期的利益との交換(利己的互惠性)によってなされると考えられている<sup>22</sup>。

この集団選択に対する批判はBoehmの議論と矛盾するとは限らないが、彼は、平等主義的倫理観が定着してきた旧石器時代の後期の人類にとっては、集団内の個体同士の競争による淘汰の方が集団間の淘汰より重要だという通説は妥当しないと考え、その理由として次の4つをあげている。

- a. 進化生物学者などの用いる数学モデルでは、肉の平等な分配など、自然な支配欲(あるいは支配されたくないという欲求)と異なっている性質が、結果的に食糧や生殖機会の平均化をもたらす集団としての再生産能力を高めることをモデル化できない。
- b. 狩猟・採集民の共同体は、狩猟や防衛のための集団行動のため合意形成が必要になる。合意形成がされなければ集団は弱体化ないし離散し、メンバーの生存確率は低下するのでメンバー間の協調行動が増え、個人間の競争による淘汰は緩和される。
- c. こうした合意形成による集団行動は、移住場所の決定など自然淘汰における成功と失敗の決定的要因になり、集団間の違いが集団内の個人差より重要になる。
- d. フリーライダーをコントロールするために、ゴシップ、からかい、村八分など多様な制裁が使われる。子供の社会化、利他的行為の賞賛とフリーライダーへの非難、によって社会的に利他主義を強制するとも言える。

このようにBoehmの議論は従来の集団選択論(群淘汰)とは異なる論理で集団間の淘汰が平等主義という倫理観を社会に浸透させるというシナリオを提唱している。とは言っても、Boehmの人間社会の倫理性の進化的起源についての議論は現代の狩猟・採集民や類人猿の共同体での観察に基づく一つの仮説に過ぎず、倫理規範形成の経済合理的説明として受け入れられるには、さらに裏付けが必要である。そのような裏付けとしては、近年の神経経済学の成果が参考になる。

<sup>20</sup> 大きな獲物の平等的分配に関しては、狩猟のうまさを誇示することによって性的な魅力になるという評判あるいはシグナリングの効果を重視する研究もある。(Hawkes and Bird, 2002)

<sup>21</sup> 群淘汰の評価については例えば内田(2007)参照。

<sup>22</sup> 親が自分を犠牲にして子を救うような行為は、進化生物学などでは利他的とは呼ばず血縁選択(淘汰)と呼ぶ。遺伝的繋がりがあれば親子に限らない。

## 5. 行動経済学、神経経済学による倫理規範形成の経済合理的説明

経済学で進化ゲームが必要とされた主な理由は、ゲーム理論の歴史の中で、均衡概念を精緻化するためにプレーヤに要求される合理性がエスカレートしていったことへの反省と、実験経済学や現実の社会で見せる人々の行動が合理的プレーヤを仮定する理論モデルの予想する行動と異なっていたことである。そこで、限界合理的な個人を前提とした進化ゲームで、現実の世界の人間行動を説明しようとし、一部のプレーヤの目的関数として利他主義、互酬性プレーヤという既に何らかの規範が内面化されたプレーヤの存在を仮定すれば、進化的均衡として、現実の人々の利他的行動や互酬的行動を説明できるようになった。

次に利他主義、互酬性プレーヤの行動を正当化するに当たっては、通常は前節のような議論を持ち出さず、行動経済学とその生理学的裏付けを提供することを（そして理論的体系化を）目指す神経経済学の成果が用いられるのである<sup>23</sup>。

ただし行動経済学については、人間の意思決定に特徴的なバイアスに注目し、経済学の理論の予測正確性に貢献している点を評価する経済学者も、理論の一般化と扱いやすさを両立させることの難しさ、またその問題への取り組みが不十分であると指摘することが多い。たとえば Fudenberg (2006) は、そうした留保を付けた上で、しかし経済学に与えられた目的が外生的変数からできるだけ正確な予測をすることだとしても、脳の仕組みが分かれば、人間の意思決定についての予測の正確さが向上する可能性はあるという期待を示している。その例として、1 回限りの囚人のジレンマ型ゲームでの協力的行動が、プレーヤの繰り返しゲームの一部という誤った感覚から生じるのか、人間の協力的交流からの喜びを求めての行動なのかという疑問に対する答えを得られるかもしれないと述べている<sup>24</sup>。

<sup>23</sup> 本節の以下の議論は主に町野（2008）の一部を要約したものである。

<sup>24</sup> Fudenberg (2006)．ただし一部のニューロ・エコノミストが主張するように、後者の解釈が、ゲームの際のプレーヤの脳の（報酬を得たときの情動的活動を示す）線条体の活動によって証明されたと考えるのは時期尚早であると注意を促している。

もちろん、神経経済学が現段階で従来の経済学のモデルを全面的に置き換える代替モデルを持っているわけではないが、そのための枠組み作りの基盤となる考え方として、Camerer, et al (2005) による脳の機能（情報処理過程）の整理は前節の議論を支持する見方である。彼らは脳の機能を、コントロールされているか自動的か、理性的か情動的かという二つの基準軸の組合せで計 4 つの象限（タイプ）に分けて整理している。コントロールされている情報処理過程とは、論理や計算のように意識的に段階を追って考える種類の過程であり、なにか意識して考えなければならない状況に直面した場合にそれなりの努力をしながら情報を処理する過程である。自動的な情報処理過程とは、それとは逆に無意識に脳の複数の領域を使って努力せずに情報が処理される過程である。自動的な処理で解決できない場合にコントロールされる過程へ移行する。情動とは、怒り、恐れ、嫉妬といった感情だけではなく、空腹、喉の渇き、性的欲求、肉体的な痛みや不快感、薬物中毒のような行動の動機をもたらす要因であり、それを脳の中で処理する過程が情動的な情報処理過程である。理性的情報処理過程とは行動するかどうかではなく、真実か否かを判断するような情報処理過程である。

この枠組みに従うと、従来の経済学のモデルはコントロールされていて理性的という脳の機能の一つのタイプに対応しているだけで、他の 3 つのタイプの脳の機能に対応する意思決定をモデル化できていなかったことになる。しかしそれ以上に問題なのは、コントロールされているか自動的か、理性的か情動的という分類軸はいずれも二者択一の分類ではないという点である。前者については自動的な情報処理は常にされていて、特殊な場合にコントロールされた情報処理が行われる。従ってコントロールされていて理性的なはずの意思決定も自動的な情報処理という意識されていない脳の活動に影響されている。また後者についても、進化の過程で情動的な情報処理を行う脳の領域に理性的情報処理を行う脳の領域が追加されてきたと考えられており、理性的情報処理過程は情動的情報処理過程の

脳のそれぞれの部分は様々な状況で活動するし、報酬を感じる場合も単に嬉しさを感じるだけではない様々な情報処理が行われ、ゲームを理解する仕方も文化によって異なるからである。

影響を免れない。情動と理性の二分法は、デカルト（1649）の単純な行動（反射的行動）と、複雑な行動（随意的行動）という二分法以来自然科学と人文・社会科学に共通の人間の認識に関わる考え方であったが、Damasioの一連の研究（例えば1994）などから両者の一体性が明らかにされてきた。感情などの情動によって合理的思考がじゃまされることもあるが、感情の衰退・欠如も不合理な行動の重要な要因となる。

前節及び本節のこれまでの考察から、人間の利他性あるいは公平性は、人類の進化の過程で利他的あるいは公平に振る舞うことが個人及び共同体のサバイバルに必要であったことに起源があり、それが脳に無意識な反応として組み込まれ、不公平な行為にはネガティブな感情を抱くようになった、という仮説が立てられる。しかし、それだけでは政策目的としての公平性について考えるには材料不足である。人間の利他性や公平性についての指向は、個人によって、また同じ個人でも相手によって大きく異なるし、状況に応じて、また文化に応じてその強さも異なる。この点については、Adam Smithが道徳感情論（1759）で拠り所としていた概念であるsympathy（同感）についての神経経済学的議論が参考になるかもしれない。Rustichini（2005）によると、スミスにとって、sympathyは他人の（主に不幸に向けられた）気持ちのシミュレーション（他の人の行動を自分の行動として仮想的に追体験すること）である。スミスの考えるsympathyは他人に起こった出来事の観察から生じるのであって、その出来事によって起こるその人の感情表現の観察から生じるのではない。しかし、このsympathyが、逆に他人（あるいは見えない第三者である観察者）から認められ得ることを認識させ、認められたいという欲求を生み、道徳的行動をもたらすというのがスミスの仮説である。

Rustichini（2005）は、この仮説の検証がミラー・ニューロンの発見で新たな段階を迎えていると考える<sup>25</sup>。他人の動きを観察したときに反応するのは、ミラー・ニューロンがシミュレーションによって他の人の行動を「理

<sup>25</sup> ミラー・ニューロンは前運動皮質（運動前野）の中央部にあり、何らかの目的をもった一連の身体の動きに反応して活動する。自分がその動きを行ったときと他人のそうした動きを観察したときのどちらの場合にも活動するのでこの名前が付けられた。

解する」機能を担っているからだと考えられている。動きだけでなく嫌悪感（吐き気）のような情動でも実際に経験した場合と、他人がそのような情動を経験した時の表情を観察した場合の双方で、脳の前頭部の広い部分が共に活性化する。

スミスはsympathyを、観察された出来事から他人の心の状態をシミュレートすること、と考えていたが、実際にはミラー・ニューロンによる“sympathy”は行動だけでなく情動の表現からもシミュレートできると考えられているので、スミスの言うsympathyはむしろempathyと言うべきだと指摘する議論もある（Binmore, 1998）。前者は相手の感情を自分も感じるが、後者は、（ゲームの木で他人のノードでの意思決定を他人の期待利得を計算して予測できるように）相手の立場に自分をおいて想像するだけで、痛みや喜びを感じることはない<sup>26</sup>。しかし、スミスのsympathyは他人の痛みを（強さの違いはあるが）感じたり、他人が手や足を打たれる場面を見て反射的に自分の手や足を引っ込めたりするというものであり、この点では上述のミラー・ニューロンのシミュレーションの方に似ている。

ただし、sympathyにしてもミラー・ニューロンにしても観察対象からその心的状態に関する情報を引き出してくるというだけあり、これだけでは、動機、意図、将来の行動をすべて理解することはできない。上述のDamasioの研究などで明らかのように身体と脳は強く相互作用しているため、脳は身体が接している周りの環境とも強く相互作用する。他人の立場に自分を置いてみて考えることで、環境全体（社会の一般的な選好の分布など）の事前評価を再評価するというような複雑な思考過程も必要となるはずである。また、意思決定は脳のどこか1箇所情報が集まってなされるのではなく、いくつかの場所がタイミングを同期させることによってなされる。従って作動記憶も重要になる。いずれこうした、より複雑な思考過程全体もニューロサイエンスによって明らかになるかもしれない。

なお、以上のような他人の情動、あるいは他人のおかれた状況を考慮して意思決定する（行動する）人間の能力は、人間が社会的動物であることから進化的に発達してきたと考えられており、上述の神経科学の研究成果

<sup>26</sup> Binmoreのempathy（共感）という概念は、後述するようにHumeから得られているが本稿でのempathyの議論はすべてBinmoreの解釈に従っている。

の一つとして、自己犠牲を伴うような行動（社会的行動）を司る脳の部位は進化的に新しいことも分かっている。Damasioの研究例でも、進化的には新しい頭前皮質を損傷した患者が、日常生活は（損傷しなかった進化的に古い脳を使って）スムーズに行えるのに、自分の意思決定が周囲に与える影響を推測するための脳の（進化的に新しい）部位を失ったことによって、スムーズに社会生活を送れなくなったのである。

## 6. 結論に変えて：共通価値基盤の形成のために

本稿では、「多元分散型統御を目指す新世代法政策学」において、法政策学の目的を考える上で経済学の知見がどう活かせるかを考えてきた。もし分配について合意された政策目的が与えられれば、それをいかに効率的に達成するかは、市場メカニズムやメカニズム・デザインなどの、まさにミクロ経済学の中心的な研究テーマであり、経済学で解決できる分野であるが、分配についての（すなわち正義や公正という価値判断を含む）政策目的を評価しなければならない場合は、政治経済学的分析や社会選択論などの研究成果はあるものの、経済学の主要な研究分析であったとは言えない。しかし本稿では、最近の進化ゲーム、行動経済学、神経経済学や進化生物学などの研究成果を総合すると、規範形成の進化的形成過程について一つの経済合理的説明が可能になることを明らかにしてきた。即ち、法の体系がHayek的自生的秩序であるにしても、Dworkin的な integrity を保ちながら変化する体系であるにしても、それらを人類の進化過程や脳の仕組みの解明によって、（経済）合理的に説明できる可能性があるということである。

人類の進化過程や脳の仕組みをすべて解明することは現段階では夢物語であろうが、人間の倫理的規範の形成という点に絞ると、筆者がその説明として注目しているのは Binmore (1994, 1998) の共感均衡のアイデアである。これは Binmore が、リベタリアンでもなく社会主義者でもないリベラルの立場を擁護するために、ゲーム理論を使って構築した壮大な哲学的議論の中核に位置する概念である。Binmore は、前節で述べたように Adam Smith の sympathy ではなく、(Smith の友人でもある) David Hume の議論の重要な要素である empathy という概念をゲーム理論的に解釈し直

し、それを Rawls の議論に触発されて、共有された効用評価とでも言うべき共感均衡という概念を考案した。

共感均衡形成プロセスを直観的に説明すると、個々のプレーヤがそれぞれ主観的に他人と自分の効用を比較し、その主観的な他者との効用比較に関して徐々に異なる個人の間で合意がなされていく、というものである。最終的には社会的に合意されたその効用評価が、社会全体の共有認識となった時点で共感均衡となる。さらに original position (=この場合は Rawls とは異なりもっと現実的な status quo とする) を共感均衡に照らし合わせて、status quo と代替案である制度や政策のどちらがその社会にとって望ましいかどうかを判断することになる。

主観的な他者との効用比較に関して徐々に異なる個人の間で合意がなされていく過程は、Hayek 的自生的秩序の形成過程にも当てはまると解釈できるし、ある程度合意がなされて社会全体の共有認識となったものの中には Dworkin 的 integrity のコアとなる共有された道徳が含まれるとも解釈できる。Binmore の共感均衡のアイデアは、Rawls や Harsanyi (1977)<sup>27</sup> と違ってシステムの外部に絶対的な評価基準を置いておらず、常に存在するとは限らない。その社会に、ボトムアップで創り出され、かつ合意できるような共通価値基盤が存在しなければ、共感均衡の実現は困難であることを示唆している<sup>28</sup>。したがって、社会の「多元分散型統御を目指す」ためには、こうした社会の共通価値基盤が何であるかを把握するステップが必要であり、共通価値基盤あるいは共感均衡が存在しない場合は、それらの形成のために、法政策の立案過程を設計するという発想が必要かもしれない。

<sup>27</sup> 同書 Chapter 4 Morality and social welfare 参照。veil of ignorance の背後で全てのプレーヤは同じ選好をもつので共感均衡のような面倒なプロセスは必要ない。その上で理想的観察者を想定する。

<sup>28</sup> 山岸 (1999) の安心社会から安全社会へという主張や Kaplow and Shavell (2002) が直感的な社会規範に頼ることを否定したこととも通じる。前節で議論したように、これがお互いに感情移入までできるような小さな、関係の密な共同体であれば sympathy に基づく同感均衡が成立するかもしれない。

## 参考文献

- 1) Aoki, Masahiko (2001) “The Subjective Game Form and Institutional Evolution as Punctuated Equilibrium” in *Towards a Comparative Institutional Analysis*, MIT Press. (滝澤・谷口訳 (2001) 『比較制度分析に向けて』 NTT 出版)
- 2) Arrow, Kenneth J. (1951) *Social Choice and Individual Values*, Wiley. (長名寛明訳 (1977) 『社会的選択と個人的評価』 日本経済新聞社；原著第2版 (1963、追加注付き) の訳).
- 3) Boehm, Christopher (2004) “What Makes Humans Economically Distinctive? A Three-Species Evolutionary Comparison and Historical Analysis.” *Journal of Bio-economics*, 6 2, pp. 109-35.
- 4) Boehm, Christopher (1999) *Hierarchy in the Forest: the evolution of egalitarian behavior* Harvard University Press.
- 5) Binmore, Ken (1994, 98) *Game Theory and Social Contracts I, II*, The MIT Press.
- 6) Camerer, Colin, George Loewenstein, and Drazen Prelec. 2005. “Neuroeconomics: How Neuroscience Can Inform Economics.” *Journal of Economic Literature*, 43 1, pp. 9-64.
- 7) Damasio, Antonio R. 1994. *Descartes’ Error: emotion, reason, and the human brain*. New York: G.P. Putnam. (田中三彦訳 (2000) 『生存する脳：心と脳と身体 の神秘』 講談社.)
- 8) Dennett, Daniel C. (1995) *Darwin’s Dangerous Idea - Evolution and the Meanings of Life* - Penguin Books. (石川幹人 他訳 (2001) 『ダーウィンの危険な思想：生命の意味と進化』 青土社.)
- 9) de Waal, Frans (1996) *Good Natured - The Origins of Right and Wrong in Humans and Other Animals* -, Harvard University Press. (西田利貞・藤井留美訳 (1998) 『利己的なサル、他人を思いやるサル』 草思社)
- 10) Descartes, René and Stephen Voss (1989) *The Passions of the Soul*. Hackett Pub. Co. (原著はデカルト著フランス語 (1649) 谷川多佳子訳 (2008) 『情念論』 岩波文庫.)
- 11) Dworkin, Ronald (2000) *Sovereign Virtue: the theory and practice of equality*. Harvard University Press. (小林公 他訳 (2002) 『平等とは何か』 木鐸社).
- 12) Dworkin, Ronald (1986) *Law’s empire*. Belknap Press. (小林公訳 (1996) 『法の帝国』 未来社).
- 13) Fudenberg, Drew (2006) “Advancing beyond Advances in Behavioral Economics.” *Journal of Economic Literature*, 44 3, pp. 694.
- 14) 藤田友敬、松村敏弘 (2005) 「社会規範の法と経済——その理論的展望」 *ソフトウェア研究* 第 1 号
- 15) Harsanyi, John C. (1977) *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*, Cambridge University Press.
- 16) Hart, Herbert L. A. (1961) *The Concept of Law*, Oxford University Press. (矢崎光 窓 監訳 (1976) 『法の概念』 みすず書房.)
- 17) Hawkes, Kristen and Bird, Rebecca B. (2002) “Showing Off, Handicap Signaling, and the Evolution of Men’s Work” *Evolutionary Anthropology*, pp. 58-67.
- 18) Hayek, Friedrich A. von. (1973) *Rules and Order*. University of Chicago Press. (矢島・永吉 訳 (1987) 『法と立法と自由 I：ルールと秩序 (ハイエク全集 8)』 春秋社.)
- 19) Kaplow and Shavell (2002) *Fairness versus Welfare*, Harvard University Press.
- 20) Kaplow and Shavell (2001) “Any Non-welfarist Method of Policy Assessment Violates the Pareto Principle,” *Journal of Political Economy*, 109 pp. 281-286.
- 21) Knight, Jack (1992) *Institutions and Social Conflict*, Cambridge University Press.
- 22) 町野和夫 (2009) 「平等主義の進化的起源と規範形成のゲーム・モデル」 *経済学研究* (北海道大学) 58巻 4号199-215頁.
- 23) 町野和夫 (2008) 「ニューロサイエンスと経済学モデル」 *経済学研究* (北海道大学) 58巻 3号187-198頁
- 24) 町野和夫 (2008) 『ミクロ経済学・入門』 有斐閣、第15章.
- 25) 町野和夫 (2003) 「倫理的規範形成のゲーム論的分析」 *経済学研究* (北海道大学) 53巻 3号283-297頁.
- 26) Mueller, Dennis C. (2003) *Public Choice III*, Cambridge University Press.
- 27) 西條辰義・大和毅彦 (2006) 「自然なメカニズム・デザインをめざして」、鈴木・長岡・花崎編『経済制度の生成と設計』第2章、東京大学出版会.
- 28) 大浦宏邦 (2008) 『社会科学者のための進化ゲーム理論：基礎から応用まで』 勁草書房.
- 29) 尾崎一郎 (2009) 「トートロジーとしての法学？—法のインテグリティと多元分散型統御—」 *新世代法政策学研究* (北海道大学) 3号191-220頁.
- 30) Rawls, John (1971) *A Theory of Justice*, Harvard University Press. (鈴木興太郎 訳 (1988) 『福祉の経済学—財と潜在能力—』 岩波書店).
- 31) Rustichini, Aldo (2005) “Neuroeconomics: Present and Future.” *Games and Economic Behavior*, 52 2, pp. 201-12
- 32) Sen, Amartya (1985) *Commodities and Capabilities*. North-Holland (矢島鈞次 監訳 (1979) 『正義論』 紀伊国屋書店)
- 33) Sen, Amartya (1970) “The Impossibility of a Paretian Liberal.” *Journal of Political Economy*, 78 1, pp. 152-57.
- 34) Smith, Adam (1759) *The Theory of Moral Sentiments*. A. Millar; and A. Kincaid and J. Bell. (水田洋訳 (2003) 『道徳感情論 (上・下)』 岩波文庫.)
- 35) 鈴木興太郎 (2006) 「制度の理性的設計と社会的選択」、鈴木・長岡・花崎編『経

リレー連載：「多元分散型統御の基礎理論を目指して」

済制度の生成と設計』第1章、東京大学出版会。

- 36) Suzumura, Kotaro (2002) “Introduction” Arrow, K. J., Sen, A. K., and Suzumura, K ed. Handbook of Social Choice and Welfare, Vol. 1, Elsevier. (鈴木他監訳 (2002) 『社会的選択と厚生経済学ハンドブック』丸善.)
- 37) 常木 淳 (2008) 『法理学と経済学』勁草書房.
- 38) 内田亮子 (2007) 『人類はどのように進化したか：生物人類学の現在』勁草書房.
- 39) 内井惣七 (1996) 『進化論と倫理』世界思想社.
- 40) 山岸俊男 (1999) 『安心社会から信頼社会へー日本型システムの行方』中公新書.
- 41) Young, Peyton, H. (1993) “The Evolution of Convention,” *Econometrica*, Vol. 61, p57-84.