

Motivation Crowding Out と利他性

Motivation Crowding Out and Altruism

早稲田大学社会科学総合学術院教授
井上 正 INOUE, Tadashi

1. Motivation Crowding Out

標準的な経済理論によれば、ある活動に対してインセンティブとして金銭的報酬の提供を増やせば、パフォーマンスは向上すると考えられている。すなわち、活動に対する金銭的報酬とそのパフォーマンス・レベルの間には、単調な正の関係があると考えられている。しかしながら、多くの実証結果から明らかになっているように、実際には金銭的報酬がパフォーマンスに及ぼす効果はそう単純ではない。

例えば、人は金銭的報酬とは無関係な行動を規定するような社会規範に影響を受ける場合がある。献血は社会に対する貢献であり、可能な限り行うべきものと考えられている。しかしながら、献血者に金銭的な補償を行うことは、Titmuss (1970) も言うように、義務感を破壊し献血という行動を減少させる可能性がある。また、互惠性 (reciprocity) のような社会規範がある場合には、行動の対価として金銭が支払われると、その行動の動機としては、互惠性ではなく報酬が考えられるようになってしまい、互惠性が破壊され、行動が互惠性に基づいて行われなくなることがある。Fehr, Gächter and Kirchsteiger (1996) は、労使関係において、外発的なインセンティブの導入は、互惠性の機能不全を引き起こすため従業員の

パフォーマンスを低下させるとしている。

このように、パフォーマンスと報酬の関係は必ずしも単調なものではない。それゆえ、金銭的報酬とパフォーマンス、ひいてはそのパフォーマンスを作り出す努力との間には別の要因が影響していると考えられる。

ところで、報酬が行動や活動に与える影響については、従来心理学分野で研究されてきた。1940年から60年代に隆盛した行動主義理論の一つであるオペラント条件づけ (operant conditioning) では、行動は報酬によって動機づけられると説明され、もう一つの動因低減説 (drive reduction theory) では、行動は生理的動因や欲求を低減させるべく動機づけられるとされた。すなわち、行動主義によると、中立か、あるいは軽度の不快感を伴う活動に対して報酬が与えられると、最終的にはその活動に肯定的な価値を関連付けることになる。言い換えれば、報酬はその活動のパフォーマンスにプラスの影響を与えると考えた。しかし、この結論は、70年代の初めに登場してきた認知心理学派によって異を唱えられた。認知心理学派によると、ある活動を行う場合に、報酬とは無関係にその活動自体の動機 (motivation)¹⁾ があり、これを内発的動機 (intrinsic motivation)²⁾ と呼んだ。そして、内発的動機に基づいて行われている行動や活動に対し、外発的

な報酬（特に金銭的報酬）が与えられると、内発的な動機が低下し、活動そのものが縮小してしまうことがあるとした。言い換えれば、金銭的な報酬のような外発的な要因の介入によってパフォーマンスが低下する現象、すなわちクラウディング・アウト（crowding out）³⁾ が起こると主張した。そして、人間行動は、内発的動機づけ、および外発的動機づけの両方により影響を受けていると考え、外発的動機づけは、外部から活動を活性化されるものであるのに対し、内発的動機づけは、まさに遂行しようとしている活動に関連して生じるものであり、人はその活動を行うことが「楽しい、興味がある、好きである」ことから、活動自体からある種の満足を得られると考えている。

このように、人は内発的動機に基づく行動に従事するとき内的な満足を得るが、そのような行動に対し外発的な報酬が提供された場合、内発的動機に基づく努力が低減または排除されるというクラウディング・アウト効果をどのように説明するかは、人間の行動を理解するうえで重要な課題であると思われる。

また、以上では、内発的動機を中心にみてきたが、内発的動機だけでなく、他の人の利益向上を意図する向社会的行動（prosocial behavior）⁴⁾ においても、外発的な報酬と努力の間で、クラウディング・アウトが生じることがあることが分かっている

ここで、内発的動機と向社会的動機の関係について考えてみよう。内発的動機に基づく行動は、上記したように自分自身の効用を追求する個人主義的な行動に近いものであるのに対し、向社会的動機は、「他を重視する」という選好に従う、その人自身の目的関数と他の人の厚生との間のより複雑な相互作用である。もちろん、内発的に動機づけられた行動が、他の人に正の外部性を生み出し、それゆえ他の人の厚生に影響を与える可能性は常にあるが、それはその人自身の効用に見返りのない意図しない影響である。一方、向社会的動機とは、人が他の人の厚生を向上させることで自分の効用をも追求する意図的な行動である。それゆえ、向社会的動機を刺激することは、その人の

パフォーマンスや他の人や社会の厚生に直接的な影響を与える。このような向社会的動機は、内発的動機の特定の形態、すなわち、内発的動機は向社会的行動へと人を導く動機であると、しばしば想定されてきた⁵⁾。しかしながら、この場合の内発的動機は、単に自己の内面に沸き起こった興味・関心や意欲、といったものではなく、他の人や社会への貢献に価値を置き、他の人や社会の利益に貢献すること自体に興味、関心、喜びを感じるといったものである。他方、外発的なインセンティブが向社会的行動を促すこともある。すなわち、他の人や社会へ貢献することにより、自分自身も利益を得たり、何らかの目に見える見返りを得ることによって向社会的行動を行う場合である。Grant（2008）によれば、内発的動機づけのレベルに応じて、向社会的動機づけが異なる自律的な規制を受けるとし、人が他の人に利益をもたらす動機は、その人を「助けたい」と思っている場合と、「助けなければならない」と感じている場合があるとしている。そして、「助けたい」と思っている場合には、上記の意味での内発的動機が向社会的行動を促すが、「助けなければならない」と感じている場合は、それは純粋な意味での内発的動機というよりは、義務感や使命感⁶⁾ であり、外発的誘因の誘惑を受けやすくなると考えられる。

本論文は、金銭的報酬のような外発的動機づけにより向社会的行動がどのように変化するか、すなわちクラウディング・アウト効果が、どのように生じるのかに関し、個人のもつ利他性に注目し、フォーマルなモデルを用いて検証することを目的としている。

2. 関連する研究

本節では、内発的動機や向社会的行動におけるクラウディング・アウト効果に関連する研究を整理する。クラウディング・アウト効果を検証する研究や実証実験は多々ある。上述のように、心理学の分野では、内発的動機による活動が、外的な報酬の導入によってクラウディング・アウトがお

こるという現象についての研究は、70年代から行われてきたが、寄付行為やボランティア活動のような向社会的行動と外発的報酬との関係は扱われていなかった。しかしながら、2000年以降になると、向社会的活動に関する動機づけと、外発的な報酬との関係についての実証実験が進められるようになってきた。本節では、それらの研究や実証実験結果を概観しておこう。

まず、Titmuss (1970) は、献血者に金銭を支払うことは、献血する彼らの意欲に負の影響を及ぼすことを報告している。すなわち、金銭による補償が、市民としての義務感を台無しにしてしまう傾向があるとした⁷⁾。Deci (1971) は、自宅の芝刈りを自発的にやっている少年に対し、彼の父親がこの芝刈りに対し小遣いを提供した場合、このインセンティブに少年がどのように反応するかを観察し、報酬(小遣い)を得ると、少年が自発的に芝刈りをする意欲を失うことを報告している。Fabes et al. (1989) は、子どもが紙を色別に分ける手伝いをするという実験を行った。ある条件下で、実験者はこの手伝いに対し報酬を明示的に提供した。そして、報酬が提供されない後続のフェーズで、実験者は「後で戻ってくる」と言って部屋を出て行き、子どもたちを密かに観察した結果、報酬が与えられるのを見ていない子どもに比べて、報酬を提示された子どもたちは大人がいないところで紙を分別することができなかったことを発見している。Frey and Oberholzer-Gee (1997) は、スイスの小さな町で、核廃棄物をリサイクルする工場の設置受容度を分析している。そして、その貯蔵施設の受け入れ賛成率は、市民に補償を提示した後に、予想に反し51%から25%へと低下したことを発見している。Gneezy and Rustichini (2000a) は、イスラエルの幼稚園でフィールド調査を行い、両親が彼らの子供を迎えに来るのが遅くなる頻度を調査した。そして、遅刻に対し罰金が導入されると、罰金導入前に比べ両親は余計に遅刻して迎えに来るようになったという事実を発見している。Warneken and Tomasello (2008) の、子どもが落ちている物を拾って大人を助けるという実験では、殆どの子ど

もは、報酬が約束されていなくても、自発的にこのような援助行動を行う内発的動機を持っていることがわかっているにもかかわらず、何人かの子どもを無作為に選び、援助行動の見返りに物質的な報酬を提供すると、報酬を提供しない次の段階で、一度でも報酬を得たこれらの子どもは報酬を全く提示されなかった子どもたちに比べて手伝いをする可能性が低くなることを発見している。Deci, Koestner and Ryan (1999) は、多くの実証研究のレビューから、クラウディング・アウト効果が存在する確かな証拠を見つけている。

以上のような多くの実証研究によって示されたクラウディング・アウト効果を、どのように説明するかについて、以下ではいくつかの理論研究を紹介しながら議論を進めていくことにする。

パフォーマンスに連動したインセンティブ報酬が内発的動機や向社会的行動の抑制につながる理由やそのメカニズムについては、さまざまな理論づけがなされてきている。クラウディング・アウト効果を説明するための最も有名な理論は、Deci (1980) によって提唱された心理学分野の認知評価理論(Cognitive Evaluation Theory)であろう。この理論は、外発的要因に対する人の受け止め方(認知的評価)次第で、外発的要因は人の内発的動機づけを促進したり抑制したりするという理論である。すなわち、外発的要因は時として、人の内なる欲求を阻害し、やらされ感を強め、自律性や有能感を低下させるため内発的動機を低下させる一方、同じ外発的要因でも、認知的な評価によっては自律性や有能感を促進する場合には、内発的動機づけが高まることとしている。そして、なぜ、どのような条件で外発的動機づけが内発的動機づけの移行をもたらすのかを「因果性の所在(locus of causality)」⁸⁾という概念を用いて体系的に説明している。因果性の所在とは、ある行動に関する人の認知の帰属を表したものであり、内的な因果性の所在を持つとは、自分の活動を自分自身の自発性や承認に帰することを意味し、外的な因果性の所在を持つとは、自分の行動を外部の制約に帰することを意味している。そのゆえ、外的な因果性の所在は外発的動機につながるのに対

し、内的な因果性の所在は内発的動機につながるものであると考える。外的条件は、情動的なフィードバックを提供し、それによって人が学習し、自己決定を感じることができれば、内的な因果性の所在、ひいては内発的動機を支持するとされている。しかし、インセンティブ報酬が、自己をコントロールしていると認識され、自己決定の認識を低下させると、個人は「糸で操られている人形」のように感じ、外部からの強い因果性の所在を経験し、内発的動機づけが阻害され、動機づけは外発的なものが主になるように変化するとしている。すなわち、このような場合にクラウドディング・アウトが発生すると主張している。Frey and Benz (2004) は、この考えを用い、金銭的なインセンティブや罰則を重視した規制は、それが自分をコントロールするものと認識され、自発的な関与を認めない場合、内発的動機をクラウドディング・アウトするとしている。

外発的報酬が内発的動機をクラウドディング・アウトするのは、外発的報酬のもつ情動的な側面が原因だとしている研究もある。プリンシパルがエージェントを信頼していないことや、プリンシパルがある仕事を魅力的ではない、達成が難しい、危険であると考えていることをエージェントに伝えるためのシグナル (signal) として外発的報酬は提案されると考えるのである。そして、外発的報酬の支払い、従事する仕事の性質についてエージェントの信念に影響を与えるプリンシパルからのシグナルを作成するというのである。(Bénabou and Tirole (2003))。Bénabou and Tirole (2006) は、このようなインセンティブのシグナリング効果についての議論を拡張し、報酬は人の意図を伝えるものでもあるとしている。人は自分の好みや本当の「善良さ」について自分では明確にはわかっておらず、人は自分が向社会的な行動を行うと、自分は「善良」であるという利己的な仮定を確認することができる。しかし、これらの「善行」に罰金や報酬が与えられると、人は、自分の貢献が、当初想定していたような高い道徳的価値よりも、むしろ外発的なインセンティブによって動機づけられていると考えるようにな

り、報酬がない場合には、自発的にはそのような行為に貢献することをやめる傾向があるとしている。Bolle and Otto (2010) は、「価格」というものが持つシグナリング効果が持続的な効果をもたらすことを示している。彼らは、外発的報酬は活動に価値を与えるものであり、特にある活動に対する「公式の価格」がない状況では問題となりうると主張している。例えば、献血は多くの国で自発的に行われているが、これに報酬を導入することは経済的な評価の付与を意味し、その結果、しばしば献血量にマイナスの影響を与えるとしている。しかし、クラウドディング・アウトに対する、シグナリングに基づくこのような考え方は、上述の例で示したような子どもに対する実験に適用できる可能性は低い。なぜなら、子供が物を拾ったり、紙を分けることについて、それらの活動は、困難で面倒で、危険であるということを、報酬の支払いによって子供に明らかにすることがたとえ可能であったとしても、一旦報酬がなくなると、その報酬は再びシグナルとして機能するはずであるにもかかわらず、報酬がなくなった後に、子供たちはそのような活動にあまり従事しなくなるということが説明できない。

クラウドディング・アウトに関する他の研究は、人は戦略的に行動する可能性があるというものである。Frey (1994) は「選好における根本的な変更を仮定することなく、合理的な方法でクラウドディング・アウトを説明するのは困難または不可能」であると主張する。しかしながら、この説明の欠点は、なぜ選好が変化するのかについての理由に言及しておらず、またそれらが発生する条件にも言及していない点にある。そこで、選好の変更も生じず、シグナルとして行動を捉えない考え方として、活動の基本的な性質は影響されなくても、外発的報酬によって、人は「ペイオフ最大化」という思考が働きだす可能性があるという、人のもつ戦略的思考に着眼したものである。例えば、人助けのような活動に向社会的動機を持つ人を考えてみよう。他の考慮すべき事項が存在しない場合には、その人は、活動自体に従事するように動機づけられ、自主的にそれを行うだろう。し

かしながら、この活動が、一旦経済的取引の対象となることが分かると、その人は、その内発的な価値に加えて、その活動の交換価値（「価格」）を考慮するようになる可能性がある。そして、彼は可能ならば、この交換価値を増やしたいと考え、ある条件下では、この目標は彼が従事している活動や行動の程度（範囲）を意図的に減少させることで達成しようとするだろう。こうして、たとえ向社会的動機は、外的な報酬がない場合は変化しないとしても、外的な報酬が出現した場合には活動の実行に消極的になることは起こりうるだろう。Schnedler and Vanberg (2014) は、このような人の持つ戦略的な行動を用いて、向社会的動機づけにおいてクラウディング・アウトが起こる理由を説明している。このことは、外発的な報酬がその活動は取引可能なものであるという「市場フレーム」を形成することで、人を支配する考え方に影響を与えると考えることができ、それによって、意思決定を、内的志向なものから、市場志向へとシフトさせる可能性があるということが出来る。すなわち、その活動が経済的取引可能なものであると認識したときには、内発的動機や向社会的動機を持つ人でさえ、活動の「価格」（外的な報酬）を増加させるために、活動の「供給」を意図的に削減する可能性があるというのである。Bénabou and Tirole (2006) では、外発的報酬は当該活動がコストのかかるものであるということ伝えるシグナルの役割を果たすと考え、外発的報酬がその人の持つ内的な価値を破壊するというものであり、外発的報酬は活動に従事する人の元来持っていた動機を消し去るというものであった。しかし、Schnedler and Vanberg (2014) は、外発的報酬は、この元の動機に影響を与えないというものである。そのかわり、活動が外発的報酬の大きさに影響を与えるとき、戦略的行動を促す可能性があるというものである。

以上のように、内発的動機や向社会的行動におけるクラウディング・アウト効果を説明しようとする研究は、様々な分野で様々な観点から行われている。本論文では、人の向社会的行動において生じるクラウディング・アウト効果に関し、人の

持つ利他性という特性から分析する。利他性とは、他の人のための行動自体（そして、それだけ）が動機となっている場合の「純粋な利他主義」(pure altruism) と、それだけでなく（あるいは、それと同時に）自分の心に何らかの満足感 (warm glow) が得られることから利他的行動をとる場合もある。つまり、ここでは何らかの行動を取る場合、利他的であると同時に利己的な（自分本位の）動機の両方が混合している。このような利他性を「不純な利他主義」(impure altruism) と呼ぶことがある⁹⁾。そして、特に向社会的行動においては、利他性というものが人の行動を左右する重要なカギであるという観点から、利他性が「純粋」な場合（純粋な利他主義）と、純粋でない場合（不純な利他主義）のモデルを開発し、全く利他性を持たない、利己主義者との比較を通し、金銭的なインセンティブ報酬を与えることが果たす役割について次節では議論する。

3. モデル

人は自分自身の行動や意思決定に際し、自分のことだけを考えて行動しているわけではない。例えば、寄付や募金活動、被災地におけるボランティア活動、チームワークなど、人は必ずしも利己的行動をとっているわけではないだろう。そして、実証実験では多くの「利己的でない」人の行動が観察されてきた。

本節では、人の持つこのような利他性という特性をとり入れたプリンシパル・エージェント・モデルを分析することで、クラウディング・アウト効果がいかに生じるかを検証する。

次のような状況を考えよう。公共財や社会的利益に貢献するような、何らかの社会的活動への参加を考えているエージェントを想定する。ただし、他の人や社会に負の外部性を与えないことを条件とする。プリンシパルは、エージェントから社会的利益への貢献に値する努力 (e) を引き出すために、外発的なインセンティブ報酬 (r) を提供するとする。また、エージェントの努力に対する不効用は $e^2/2$ とする。そして、 S は、エージェン

トの1単位の努力から作り出される社会的な利益であるとする。この社会的利益は、受益者からなる社会ヘリターンをもたらし、ここでのプリンシパルは、そのような受益者に代わって行動する受託者とする。

すると、プリンシパルの目的関数 (W) は次のようになるであろう。

$$W = eS - C \quad (1)$$

ここで、 C はコストとする。ただし、簡単のためコストは、エージェントへのインセンティブ報酬支払いだけで、その他のコストはゼロとする。すなわち、ここでは、エージェントが向社会的行動をとることによってエージェント自身が何らかの効用を得たとしても、プリンシパルはその効用を直接評価しないことにする。これは、例えば、公共サービスの運営が受益者(納税者であったり、サービス料を支払ったりする場合)にもたらす利益のみを重視し、公共サービス提供者の効用は重視しない場合や、医師が治癒の喜びを経験したとしても、そのようなペイオフを医療制度の運営方法の決定に考慮しないような場合を想定する。

インセンティブ報酬の導入がエージェントの努力の減少をもたらすというクラウドディング・アウト効果を説明するために、エージェントの効用は、外発的源泉および内発的源泉の両方から成り立っているとすると、外発的動機付け要因および向社会的動機付け要因の両要因から構成される次のような効用関数を最大化するように、エージェントは努力 (e) を選択するであろう。

$$e(r+a) - \frac{e^2}{2} \quad (2)$$

ここで、 $r (\geq 0)$ はプリンシパルが提供するインセンティブ報酬、 $a (\geq 0)$ は向社会的行動からエージェント自身が得る効用とする。

さらに、エージェントの利他性を導入するため、Andreoni (1989, 1990) の議論に基づき、利他性をモデル化する。Andreoni の warm-glow¹⁰⁾ モデルでは、利他主義者を二つに区別し、人に何かを与えたいという欲求だけが動機となる「純粋な利他主義者」(「完全な利他主義者」とも呼ばれる)と、人に与えることが自分自身の喜び (warm

glow) ともなる、「不純な利他主義者」に分けている。

ここでの、不純な利他主義者とは、人は他人を助けるために「自分の役割を果たす」ことで喜びや満足感を体験すると考え、この満足感(つまり、warm glow)は、善行から得られる利己的な喜びを表している。それゆえ、このような warm glow の枠組みでは、利他的な動機とその行動をおこなう際に生じる利益(例えば、寄付に対する減税措置など)を目的として行動を行うという意味で利己的な動機の両方を同時に持ち合わせていることになる。こうして、純粋な利他主義者の行動は他の人が体験するペイオフによってのみ動機づけられるのに対し、不純な利他主義、すなわち warm glow によって動機づけられた利他主義者は、社会的利益と共に自分自身の貢献をも考慮する要素を持つと考えることができるであろう。

純粋な利他主義と warm glow によって動機づけられた不純な利他主義との違いをモデル内に明示的に導入するため、向社会的行動からエージェントが得る効用 (a) は $a = \lambda (S - r)$ で与えられるとする。ここで、 $\lambda (\geq 0)$ はエージェントの利他性の程度を表す。ただし、 λ のレベルは私的情報で、プリンシパルはいかなる費用でもこれを知り得ないと仮定する。また、インセンティブ報酬が多いと受益者の余剰が減るという事実を反映して、 S から r を差し引く ($S > r$ と仮定する)。こうして、社会的利益に貢献することをエージェントは目指しているが、それは利他性の強度にもよると考えるのである。

その結果、(2) 式は次のようになる。

$$e(r + \lambda(S - r)) - \frac{e^2}{2} \quad (3)$$

そうすると、エージェントの最適努力選択 (\hat{e}) は、(3) 式を努力 e に関して最大化することで得られ、次のようになる。

$$\hat{e} = r + \lambda(S - r) \quad (4)$$

ここで、二つの異なったシナリオを考える。第一は、利他性を持っているが、インセンティブ報酬がない ($r = 0$) 場合である。すなわち、エージェントは、インセンティブ報酬は提示されないが、

向社会的行動に動機づけられている場合である。このときには、エージェントは $\hat{e} = \lambda S$ の努力選択を行うであろう。第二は、インセンティブ報酬は提供 ($r > 0$) されるが、利他性を持っていない、いわゆる利己主義者の場合である。言い換えると、エージェントは努力のためのインセンティブ報酬を提示されるが、インセンティブ以上の努力をする向社会的動機がない場合である。この場合には、エージェントは $\hat{e} = r$ の努力選択を行うであろう。この第二ケースの場合、努力は、インセンティブ報酬が増すにつれて増加する。

実証結果によると、向社会的動機に動機づけられたエージェントが、インセンティブ報酬がない下で供給する努力は、しばしば、インセンティブ報酬により統制され向社会的動機がないエージェントによって供給される努力と等しいか、あるいは、それを越える。すなわち、 $\lambda S \geq r$ あるいは、 $\lambda \geq r/S$ である。言い換えれば、社会的利益の大きさに比してインセンティブ報酬が小さいなら、利他性が多少でもあれば、外発的動機づけが提示されないとき、向社会的動機に動機づけられたエージェントにより供給される努力は、インセンティブが提示され、しかし、向社会的動機づけが欠落するエージェントにより供給される努力以下になることはないということが分かる。

以上の前提の下で、エージェントが利己主義者、不純な利他主義者、純粋な利他主義者という三つの場合のモデルを提示する。

まず、利己主義者の効用 (U_e) であるが、彼は社会的利益への貢献には全く関心がなく、すなわち向社会的動機が全くなく、単に自身の活動から得られる効用にのみ関心を持つので、(3) 式において $\lambda = 0$ の場合に対応する。すなわち、この場合の効用は、

$$U_e = er - \frac{e^2}{2} \quad (5)$$

となる。

続いて、不純な利他主義者であるが、この場合には利他的行動を取ることにより自分自身も何らかの満足感 (warm glow) が得られる場合である。つまり、不純な利他主義者の場合には、利他的な

動機があると同時に利己的な (自分本位の) 動機の両方が混合していることになる。それゆえ、効用 (U_{im}) は、次のようになるであろう。

$$U_{im} = er + \lambda(eS - er) - \frac{e^2}{2} \quad (6)$$

最後に、純粋な利他主義者であるが、この場合には他の人に対して与えること自体 (そして、それのみ) が動機となっているので、彼の選好 (U_p) は次のように表すことができるであろう。

$$U_p = \lambda(eS - er) - \frac{e^2}{2} \quad (7)$$

以上で、利己主義者、不純な利他主義者、純粋な利他主義者の選好をモデル化することができたので、インセンティブ報酬の導入が、それぞれの場合のエージェントに対し、どのような影響を及ぼすかを見ることにする。

最初に、 $r = 0$ の場合、すなわちインセンティブ報酬が全く提示されないケースから考察する。エージェントが利己主義者の場合の効用は次のようになるであろう。

$$U_e(\hat{e} | r = 0, \lambda = 0) = 0 \quad (8)$$

すなわち、利己主義的エージェントは、 $\hat{e}_e^i = 0$ の努力選択を行う。

一方、不純な利他主義者、純粋な利他主義者のように向社会的行動をとることに動機づけられているなら、そのときエージェントの効用は、次のようになるであろう。まず、不純な利他主義者の効用 (U_{im}) は、(6) 式より

$$U_{im}(\hat{e} | r = 0, \lambda > 0) = (\lambda^2 S^2) / 2 \quad (9)$$

となる。一方、純粋な利他主義者の効用 (U_p) は (7) 式より、

$$U_p(\hat{e} | r = 0, \lambda > 0) = (\lambda^2 S^2) / 2 \quad (10)$$

となる。

インセンティブ報酬がない場合には、不純な利他主義者、純粋な利他主義者ともに $\hat{e}_{im}^i = \hat{e}_p^i = \lambda S$ という努力選択に対応していることがわかる。

次に、インセンティブ報酬が提示される場合、すなわち $r > 0$ のケースを考える。エージェントが利己主義者ならば、このときの効用は、

$$U_e(\hat{e} | r > 0, \lambda = 0) = \frac{r^2}{2} \quad (11)$$

これは、 $\hat{e}_e^2 = r$ の努力選択に対応している。

一方、エージェントが不純な利他主義者や純粋な利他主義者のように向社会的行動に動機づけられているなら、そのときの効用は、それぞれ次のようになる。

不純な利他主義者の効用は、

$$U_{im}(\hat{e} | r > 0, \lambda > 0) = \frac{1}{2}(\lambda^2(S-r)^2 + 2\lambda r(S-r) + r^2) \quad (12)$$

これは、 $\hat{e}_{im}^2 = r + \lambda(S-r)$ の努力選択に対応している。さらに、純粋な利他主義者の効用は、

$$U_p(\hat{e} | r > 0, \lambda > 0) = \frac{1}{2}(\lambda^2(S-r)^2) \quad (13)$$

これは、 $\hat{e}_p^2 = \lambda(S-r)$ の努力選択に対応している。

以上の分析から、エージェントが利己主義者、不純な利他主義者、純粋な利他主義者である場合において、インセンティブ報酬がない場合と、インセンティブ報酬が提供されたときの、最適努力選択と効用が明らかになった。

まず努力選択に関しては、

- 1) 純粋な利他主義者の場合には、常に動機のクラウディング・アウトが発生する。なぜなら、 $r = 0$ のときの純粋な利己主義者の努力選択は $\hat{e}_p^1 = \lambda S$ であるのに対し、 $r > 0$ すなわち、インセンティブが提示されると、努力選択は $\hat{e}_p^2 = \lambda(S-r)$ となり、明らかに努力量は減少する（すなわち、 $\hat{e}_p^1 > \hat{e}_p^2$ ）。それゆえ、純粋な利他主義者の場合には、クラウンディング・アウトが生じることが分かる。
- 2) 不純な利他主義者については、 $r = 0$ のときの努力選択は $\hat{e}_{im}^1 = \lambda S$ であり、 $r > 0$ のときの努力選択は、 $\hat{e}_{im}^2 = r + \lambda(S-r)$ であるので、利他性が大きい場合（すなわち、 $\lambda > 1$ ）にはクラウンディング・アウトが起こるが、利他性が小さい場合（ $0 \leq \lambda < 1$ ）にはインセンティブの提供により努力選択は増える。 $\lambda = 1$ のときは、変化しない。
- 3) 利己主義者の場合には、 $r = 0$ のときの努力 $\hat{e}_e^1 = 0$ から、 $r > 0$ のときの努力 $\hat{e}_e^2 = r$ となり、

インセンティブ報酬の大きさに応じて努力量も増えるということが分かる。

次に、3者の効用の変化に関しては、

- 4) 純粋な利他主義者に関しては、(10)式から $r = 0$ のとき $U_p = \lambda^2 S^2 / 2$ であり、 $r > 0$ のときには (13)式より $U_p = \lambda^2 (S-r)^2 / 2$ となる。その結果、インセンティブ報酬が提示されると効用も減少することが分かる。
- 5) 不純な利他主義者については、 $r = 0$ のとき (9)式より、 $U_{im} = \lambda^2 S^2 / 2$ であり、 $r > 0$ のときには (12)式より $U_{im} = (\lambda^2 (S-r)^2 + 2\lambda r(S-r) + r^2) / 2$ となる。その結果、利他性が小さい場合（ $0 \leq \lambda < 1$ ）には、効用は増加するが、利他性が大きい場合（ $\lambda > 1$ ）には減少する。 $\lambda = 1$ のときは、変化しない。
- 6) 利己主義者の場合には、 $r = 0$ のとき $U_e = 0$ であり、 $r > 0$ のときには (11)式より、 $U_e = r^2 / 2$ となり、インセンティブに応じて効用も増加することが分かる。

ここから明らかになったことは、外発的なインセンティブ報酬が提供されると、エージェントが純粋な利他主義者の場合には、努力量が減るというクラウディング・アウトが起こるのである。また、効用も減少するということである。不純な利他主義者については、努力選択に関しては、彼の利他性の程度（ λ ）に依存し、利他性が大きい場合（すなわち、 $\lambda > 1$ ）には、努力選択を減らすというクラウディング・アウトを引き起こすが、利他性の程度が小さい場合（ $0 < \lambda < 1$ ）には、逆にインセンティブ報酬に応じて努力選択を増やすということが明らかになった。また、効用については、利他性の程度が小さい場合（ $0 < \lambda < 1$ ）にのみ、効用は増加するということが明らかになった。

以上の分析結果によれば、前述したような献血行為に関して、献血者に対し金銭的な補償を行うことは、献血を純粋な向社会的行動として行っている人については献血を減少させるということを示唆している。すなわち、純粋に慈善行為や社会貢献の点から献血を行っている人に、インセン

ティブを提示することは、その行為を忌避させることを示している。しかし、慈善や社会貢献のためだけでなく、自分の心に何らかの満足感 (warm glow) が得られることから献血を行っているような人、すなわち不純な利他主義者に関しては、インセンティブ報酬の提供は献血の継続可能性があることを示唆している。

さらに分析を進め、利他性の程度 (λ) や社会的利益 (S) の変化に応じて、純粋な利他主義と不純な利他主義者の努力選択および効用がどのように変化するかを見てみると、表 1 のようにまとめることができる。利己主義者に関しては、努力量も効用もインセンティブ報酬にのみ影響を受けるので、この表からは除外されている。

次に、社会的利益の受託者としてのプリンシパルのペイオフは、3つのケース (利己主義者、不純な利他主義者、純粋な利他主義者) でどのようになるかを検討しよう。

プリンシパルの目的関数は、(1) 式で示したように $W = eS - C$ であり、 C はエージェントへの支払い (r) だけで、その他のコストはゼロとしたので、 $W = e(S - r)$ となる。

その結果、インセンティブ報酬が支払われる場合、3者の最適努力選択に対応するプリンシパルのペイオフは、次のようになるであろう。

(a) 利己主義者のケース：

最適努力選択は、 $\hat{e}_e^2 = r$ であったので、

$$W = r(S - r) \quad (14a)$$

(b) 不純な利他主義者のケース：

最適努力選択は、 $\hat{e}_{im}^2 = r + \lambda(S - r)$ であったので、

$$W = r(S - r) + \lambda(S - r)^2 \quad (14b)$$

(c) 純粋な利他主義者のケース：

最適努力選択は、 $\hat{e}_p^2 = \lambda(S - r)$ であったので、

$$W = \lambda(S - r)^2 \quad (14c)$$

(14a), (14b), (14c) の3式を比較すると、(14b) \geq (14a)、そして (14b) \geq (14c) は、すぐに分かる。

しかしながら、(14a) と (14c) の大小関係に関しては、インセンティブ報酬の大きさ (r)、利他性の程度 (λ)、社会的利益の大きさ (S) に依存する。

インセンティブ報酬の大きさ (r) に関しては、 $r \leq \frac{\lambda S}{(\lambda + 1)}$ ならば (14c) \geq (14a) となり、 $\frac{\lambda S}{(\lambda + 1)} \leq r \leq S$ ならば (14a) \leq (14c) となるが、 $\frac{\lambda S}{(\lambda + 1)}$ の値は、 λ が非常に大きい (すなわち、利他性の程度が非常に大きい) と S に近づくので、エージェントの利他性が大きい場合には、インセンティブ報酬の大きさの如何にかかわらず、エージェントが純粋な利他主義者の場合の方が、プリンシパルのペイオフを大きくする。

利他性の程度 (λ) については、 $\lambda \geq \frac{r}{S - r}$ ならば (14c) \geq (14a) が成り立つ。この場合、インセンティブ報酬が非常に小さければ、これは常に成り立つ。一方、インセンティブ報酬の大きさが社会的利益と同じほど大きくなると (すなわち、非常に大きなインセンティブ報酬が提示されると)、(14c) \geq (14a) が成立するためには、非常に大きな利他性が要求される。

また、社会的利益 (S) に関しては、 $S \leq r + (r/\lambda)$ ならば、(14a) $<$ (14c) が成り立つ。しかしながら、この不等式は $S - r$ が (r/λ) 以下でなければならぬことを示しているので、 λ が十分大きければ、 S と r がほとんど等しくなり、この不等式が成立することは稀である。すなわち、非常に

表 1 利他性及び社会的利益に関する努力量と効用

	純粋な利他主義者	不純な利他主義者
利他性の程度 (λ)	λ の増加にともない努力も効用も増加する。	λ の増加にともない努力も効用も増加する。
社会的利益 (S)	S の増加にともない努力も効用も増加する。	S の増加にともない努力は増加する。効用は $S \geq r(\lambda - 1) / \lambda$ なら増加する。

利他性が高い純粋な利他主義者は、利己主義者よりも社会的利益を高めることができるということがわかる。

以上より、社会的利益の受託者であるプリンシパルのペイオフを最も大きくするのは、不純な利他主義者である。そして、純粋な利他主義者と利己主義者に関しては、純粋な利他主義者の利他性が非常に大きいならば、純粋な利他主義者の方がプリンシパルのペイオフを大きくするが、一般的には、インセンティブ報酬の大きさ、利他性の程度、社会的利益の大きさに依存するということが明らかになった。

4. 結論

本論文では、向社会的行動に関してエージェントが、利己主義者、不純な利他主義者、純粋な利他主義者である場合の効用を特定化したモデルを導入し、インセンティブ報酬の提供によって向社会的行動がどのように影響を受けるかを検証した。

分析の結果、クラウドイング・アウトに影響を及ぼすのは、「利他性の程度」「インセンティブ報酬の大きさ」「社会的利益」の3つの要因であることがわかった。そして、社会的利益を追求することに動機づけられた純粋な利他主義者であるエージェントは、外発的なインセンティブ報酬を提示されると、向社会的動機についてクラウドイング・アウト効果を経験する可能性が高くなるということが明らかになった。例えば、他の条件が等しければ、社会的利益への忠誠心が高い人ほど、インセンティブ報酬の提供は動機のクラウドイング・アウトにつながる可能性が高くなる。一方で、エージェントが不純な利他主義者の場合には、必ずしも動機に関するクラウドイング・アウト効果を経験するとは限らない。これは、内在的な社会性や興味への動機が高い活動は、外発的な報酬によって必ずしもマイナスの影響を受けないといういくつかの経験的な証拠と一致している。したがって、クラウドイング・アウトが起こるかどうかは、エージェントが生来持っている向社会的動

機がどれだけ強いのか、そしてその動機が何に結びついているかに根本的には依存している。

本論文におけるモデルは、多くの経験的な証拠と一致しており、多くの検証可能な仮説を生み出すことができる。また、本論文で用いた効用はベンサム（Bentham）的功利主義での総合的な満足度としている。すなわち、快樂・苦痛を量的に勘定できるものであるとし、効用は比較可能であると仮定し、効用を総て足し合わせたものを最大化することを重視している。そして、インセンティブ報酬が導入されたときに、エージェントが向社会的行動を取ることで効用が高くなるのか、それとも効用が低くなるのかを、3つの場合（すなわち、エージェントが、純粋な利他主義者、不純な利他主義者、そして利己主義者の場合）にわけて検討した。しかし、動機づけに関するクラウドイング・アウト効果を、ベンサムの効用の観点からモデル化し説明するには、さらなる努力が必要であることは明らかである。例えば、本論文におけるモデルでは、罰金ではなく報酬が向社会的行動に与える影響を検証しているが、罰金の場合も同様の結果が得られるかどうかは残された課題である。

また、本論文におけるモデルは、エージェントの効用を加法的に分離可能なものとして扱っているという点で、このモデルが単純であることは認めざるを得ない。これは、外発的動機づけと内発的動機づけの関係が相互作用的であることを示唆するいくつかの経験的知見と一致しないかもしれない¹⁾。さらに、本論文のモデルは、すべての経験的証拠に完全に説明できるわけでもない。

本論文の利他性に基づくモデルで説明できる行動は、献血等の向社会的動機に関するクラウドイング・アウトであろう。2節で述べたような実証実験で示されている他のモチベーション・クラウドイング・アウトの根拠に関しては、戦略的行動あるいはシグナリングという側面から説明したほうが、より説得力があることも確かであろう。

註

1) 動機とは、行動を指示し、活気づけ、持続させる心理的プロセスや、努力したいという内なる欲求のこととさ

- れる (Grant (2008))。
- 2) 内発的動機とは、自己の内面に沸き起こった興味・関心や意欲などから発生してくるやる気のことであり、Murray (1964) は、内発的動機として、感性動機 (過度な刺激を求める欲求)、好奇動機 (新しい経験や、好奇心の満足を求める欲求)、活動性動機 (活発な行動を楽しむことを求める欲求)、操作動機 (様々な試行を楽しむことを求める欲求)、認知動機 (頭を使って問題解決する楽しみを求める欲求) の5つを挙げている。また、内発的に動機付けられた行動は、外発的報酬が意識されることにより、急速に低下することが多くの実験で確かめられている。このことに関し、Deci and Ryan (1985) は、自発的行動であっても、外発的報酬が与えられると、その行動をコントロールしているのは自分ではなく、外部にあるものであると認知するようになる結果、内発的動機付けが弱められることがあるとしている。
 - 3) 心理学では、アンダー・マイニング (undermining) と呼んでいる。
 - 4) 向社会的行動とは、必ずしも自分の利益にならずとも他に人のためする行動であり、他の人や集団を助けたり、人々のためになることをしようとして行われる自主的な行為のことである。例えば、寄付行為やボランティア活動への参加などが挙げられる。
 - 5) Grant (2008) は、この2つの動機形態は、その原動力に関する前提条件に違いがあるとしている。内発的動機は、快楽や楽しみを努力の原動力とする自己の欲に基づくヘドニック (hedonic) な視点を持つものに対し、向社会的動機は、他人のために何かをしたいという、意味や目的を努力の原動力とするユーダイモニック (eudaimonic) な視点を持っている。さらに、この2つの動機では、努力を費やす理由が異なる。すなわち、内発的動機づけを持つ人は、興味や楽しみに基づいて努力をし、向社会的動機づけを持つ人は、他人に利益を与えたいという欲求に基づいて努力をする。この基本的な区別は、2つの動機における3つの違い、すなわち、自己規制 (自律的なものと内発的なもの)、目標指向性 (プロセスと結果)、時間的焦点 (現在と未来) にあるとしている。
 - 6) Ryan and Deci (2000) によれば、義務感や使命感は内発的動機ではなく外発的動機としている。
 - 7) Titmuss の主張を検証するために、Mellström et al. (2008) は、フィールド実験を行った。その結果、男女間で著しく異なり、女性ではかなりのクラウディング・アウト効果が見られたとしている。
 - 8) この概念では、外発的報酬は2つの側面から内発的動機づけに影響を及ぼすとする。1つは制御的な側面で、もう1つは情報的な側面である。Deci (1975) によれば、外発的報酬には「報酬による行動をコントロールするための手段」としての側面と、「報酬の受け手の有能さを示す情報」としての側面があるとしている。前者は、行為者が認知した因果性の所在。すなわち、ある課題に対する動機づけが内発的であるか、それとも外発的であるかという認知プロセスに影響する。この場合、「報酬のために何かしている」「報酬によりコントロールされ

- ている」と認知すると内発的動機づけは低下する。他方、後者は、行為者のある課題に対する有能感ならびに自己決定感という感情プロセスに影響する。そして、このような過程は、外発的報酬のもつ2つの側面のどちらが優位であるかによって、択一的に働くと考えられる。物質的報酬 (金銭や品物) は、多くの場合、外発的報酬のもつ情報の側面よりも制御的側面の方が優位となる結果、課題に対する内発的な動機よりも外発的な動機の方が多く見積られ、内発的な動機づけが低下するとされる (桜井(1984))。
- 9) Andreoni (1989,1990)。
 - 10) warm glow は、心理学的には外発的動機に分類される。Ryan and Deci (2000) によれば、warm glow な行動は、「内発的ではあるが、実際には自己の一部として経験されていない」ので、これらの行動は、外部環境で名声や友情を獲得する方向けられているという意味で、外発的動機と考えている。また、彼らによれば、内発的動機と外発的動機は連続性があり、コントロールの所在が内部的で、かつ自己決定性 (自律性) の程度が非常に高度の場合のみを内発的動機としている。
 - 11) Gagné and Deci (2005) は「内発的動機付けと外発的動機付けは、加法的というより、正の相互作用および負の相互作用の双方が認められる」と述べている。

参考文献

- Andreoni, J. (1989). Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence, *Journal of Political Economy*, 97(6): 1447-1458.
- Andreoni, J. (1990). Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving, *The Economic Journal*, 100: 464-477.
- Bénabou, R. and J. Tirole (2003). Intrinsic and Extrinsic Motivation, *Review of Economic Studies*. 70: 489-520.
- Bénabou, R. and J. Tirole (2006). Incentives and Prosocial Behavior, *American Economic Review*. 96(5): 1652-1678.
- Besley, T. and M. Ghatak (2018), Prosocial Motivation and Incentives, *Annual Review of Economics*, 10: 411-38.
- Bolle, F. and P. E. Otto (2010). A Price Is a Signal: on Intrinsic Motivation, Crowding-out, and Crowding-in, *KYKLOS*, 63(1): 9-22.
- Bruno, B., Faggini, M. And A. Parziale (2017). Motivation, Incentives and Performance: An Interdisciplinary Review, *International Journal of Business and Management*, 12(12): 29-52.
- Deci, E.L. (1971). Effects of Externally Mediated Rewards on Intrinsic Motivation, *Journal of Psychology and Social Psychology*. 18(1): 627-668.
- Deci, E.L. (1975) *Intrinsic motivation*. New York, Plenum Press. (E.L. デシ 安藤延男・石田梅男 (訳) (1980) 『内発的動機づけ・実験社会心理学的アプローチ』誠信書房)
- Deci, E. L. (1980). *The Psychology of Self-Determina-*

- tion. Lexington, MA: D. C. Heath: Lexington Books. (E.L. デシ 石田梅男 (訳) (1985) 『自己決定の心理学—内発的動機づけの鍵概念をめぐって』 誠信書房)
- Deci, E. L. and R.M. Ryan (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Publishing Co.
- Deci, E.L., R. Koestner and R.M. Ryan (1999). A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation, *Psychological Bulletin*. 125(3): 627–668.
- Fabes, R. A., Fultz, J., Eisenberg, N., May-Plumlee, T., & Christopher, F. S. (1989). Effects of rewards on children's prosocial motivation: A socialization study. *Developmental Psychology*, 25(4), 509–515.
- Fehr, E., S. Gächter, and G. Kirchsteiger (1996), Reciprocity as a Contract Enforcement Device, *Econometrica*, 65, 833–860.
- Frey, B.S. (1994). How Intrinsic Motivation is Crowded out and in, *Rationality and Society*. 6: 334–352.
- Frey, B. S., and F. Oberholzer-Gee (1997). The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-out, *American Economic Review*, 87(4), 746–755.
- Frey, B.S. (1997). *Not Just for the Money*. Edward Elgar Publishing, Cheltenham.
- Frey, B.S. and R. Jegen (2001). Motivation Crowding Theory: A Survey of Empirical Evidence, *Journal of Economic Surveys*. 15(5): 589–611.
- Frey, B. S., & Benz, M. (2004). From Imperialism to Inspiration: A Survey of Economics and Psychology. In J. B. Davis, A. Marciano & J. Runde (Eds.), *The Elgar Companion to Economics and Philosophy* (pp. 61–84). Cornwall, UK: Edward Elgar.
- Gagné, M. and E.L. Deci (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior* 26(4): 331–362.
- Grant, A.M. (2008). Does Intrinsic Motivation Fuel the Prosocial Fire? Motivational Synergy in Predicting Persistence, Performance, and Productivity, *Journal of Applied Psychology*, 93(1): 48–58.
- Gneezy, U. and A. Rustichini (2000a). A Fine is a Price, *Journal of Legal Studies*. 29(1): 1–17.
- Gneezy, U. and A. Rustichini (2000b), Pay Enough or Don't Pay at All. *The Quarterly Journal of Economics*, August, 115(3)
- Harvey, S.J. (2005), Why did you do that? An economic examination of the effect of extrinsic compensation on intrinsic motivation and performance, *Journal of Economic Psychology* 26: 549–566.
- Mellström, C. and M. Johannesson (2008), Crowding out in Blood Donation: Was Titmuss Right?, *Journal of the European Economic Association*, 6(4): 845–863.
- Murray, E.J. (1964) *Motivation and Emotion*. Englewood Cliffs, NJ, Prentice-Hall.
- Ryan, R. M., and E.L. Deci (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55: 68–78.
- Schnedler, W. and C. Vanberg (2014). Playing 'hard to get': An economic rationale for crowding out of intrinsically motivated behavior, *European Economic Review*, 68: 106–115.
- Titmuss, R.M. (1970). *The Gift Relationship*. London: Allen and Unwin.
- Warneken, F. and M. Tomasello (2008). Altruistic Helping in Human Infants and Young Chimpanzees. *Science*, 311(3), 1301–1303.
- 鵜野好文, 井上正, 高橋与志 (2018) 「外的介入による内発的動機づけのクラウドニング・アウト効果の再現モデル」 mimeo.
- 桜井茂男 (1984) 「内発的動機づけに及ぼす言語的報酬と物質的報酬の影響の比較」『教育心理学研究 (日本教育心理学会)』第 32 卷, 第 4 号。