

# ポテンシャル効用モデルの一般化 — 「おせっかい」「だまし」問題への適用

松村良平

本論文は、著者らが開発したポテンシャル効用モデルの基礎論を検討しなおすことで、「おせっかい」、「だまし」という行為・現象について、数理的な表現を行い、具体例を用いて説明することを目的としている。このモデルを用いた過去の研究では、プリンシパル（当局などの、情報完備主体）の効用は、エージェント（市民などの情報不完備主体）の効用そのものであると仮定してきた。しかし、本論文では、そのような限定をせず、通常のエージェンシー問題と同様に、プリンシパル自身の利己的な効用関数を想定することで、「おせっかい」や「だまし」を数理的に表現することを可能にした。エージェント自身のみならず、プリンシパルもエージェントのポテンシャル効用関数を誤認識している場合が「おせっかい」であり、エージェント自身はポテンシャル効用を認識していないが、プリンシパルはそれをわかっており、間違えた情報を教え込むことが「だまし」である。これを一般的な関数形を用いてではあるが、数理的に表現したことが本論文の限界貢献である。

keyword：ポテンシャル効用、啓蒙、おせっかい、だまし、数理モデル

## 目次

- 1 はじめに
- 2 ポテンシャル効用モデルについての見直し
  - 2-1 従来の理論との違いのレベルについて
  - 2-2 モデルの基本構造の見直し
- 3 具体例
- 4 数理モデル化
  - 4-1 「おせっかい」ケース
  - 4-2 「だまし」ケース
- 5 今後の展望

## 1 はじめに

本論文は、著者らが、情報の誤認識を表現するために開発したポテンシャル効用モデルの基礎論を検討しなおすことで、日常の人間関係レベルから、当局のような大きな存在による情報流布のレベルまでの間に広く見られる、「おせっかい」、「だまし」という現象について、数理的な表現を行い、具体的な例を用いて説明することを目的とする。本来、ポテンシャル効用モデルは、意思決定主体が、自分自身の効用関数について誤認識していて、なおかつ少なくとも、短期・中期的には、獲得する（感じる）効用を認識できないケースを扱うモデルである。ここでは、このモデルを用いて、「おせっかい」、「だまし」について特定の視点からの表現を提案する。

著者らは、過去にポテンシャル効用モデルを用いて、さまざまな問題を分析してきたが、基本的には、すべて、情報完備な主体（主に当局）による、情報不完備な主体（主に市民）への啓蒙問題をさまざまな社会問題にあてはめて分析するというスタイルであった [1, 2, 3, 4, 5]。本論文では、ポテンシャル効用モデルの基礎的な見直しをすることで、実は、啓蒙問題以外にも応用可能であることを示す。とくに、「だまし」、「おせっかい」という行為あるいは現象にこのモデルを応用することを目的とする。

通常の合理的意思決定理論では、意思決定主体は、意思決定時の自分の効用関数を基本的には正確に認識し、これを最大化するような意思決定を行うようにモデル化されている。しかし、実際の意思決定主体である人間について、これを仮定するには無理があるのではないかという議論は当初から存在していた。

ゲーム理論の文脈でいえば、意思決定主体自身ではないが、相手の戦略や効用について誤認識しているケースをハイパーゲームと名づけて、ある種のクラスの問題を扱っている [6, 7]。

また、個人の意思決定についても、ベイズ学習などがある。文献 [5] で、ポテンシャル効用とベイズ学習の違いについて、“これは、次の点で、

ベイズ学習などとは異なるものである。ベイズ学習では、自分が学習すべき内容というものを主体自身が認識していることが前提となるが、ポテンシャル効用モデルにおけるエージェントは、必ずしも自分自身の情報不完備な点についてわかっている必要はないのである。”と述べたが、さらに一点重要な違いが存在することを付記しておく。それは、ベイズ学習においては、情報を取得したり意思決定をしたりした後、さまざまなアップデートを考える点である。たとえば、ベイジアンゲームにおいて、自分のタイプを知ったときに他者のタイプについて確率計算を行ったりする。これはもともと与えられている自分のタイプについての確率分布から、何らかの情報更新が起こるたびに、ベイズの定理を用いて他者のタイプについて新しい確率を計算しているわけである。ポテンシャル効用では、これと同様のこと（つまり意思決定、情報更新後による効用などのアップデート）を考えていない。冒頭でも述べたように、ポテンシャル効用モデルは、短期・中期的には、獲得する効用を認識できないケースを扱うモデルなので、短期・中期的時間経過後では「現在、自分がどんな状態にいるのかわからない」のである。逆にいうと、このようなアップデートを考えることが現実的でない問題のクラスには、ポテンシャル効用モデルを用いるのが合理的であるというのが著者らの主張である。

いま、相手のタイプについてのアップデートを例としてあげたが、自分の効用関数自体に確率分布を振ることができる場合を考えてみよう。この場合、何かのイベントが起こるたびに、効用がアップデートするわけであり、その期の効用関数（これが真の効用関数であるかどうかにかかわらず）が決定し、割引因子を用いて、長期的効用というものを計算することができる。ポテンシャル効用モデルでは、効用関数自体に確率分布が振れない状態を扱っているのである。

著者らは、過去に、このモデルを用いて、情報不完備な主体に対する啓蒙問題を扱ってきた。たとえば、環境問題に対する意識が希薄な主体に、「われわれのちょっとした努力で、ある程度、環境悪化防止の効果がもたらせますよ」という具体

的な情報を与えることで、各意思決定主体が自律的に意思決定し（強制的な意思決定でなく、あくまで自分のもっている情報、効用関数に基づいて自律的に意思決定することが重要なポイントである）、結果として社会厚生の上につなげるというような問題をモデル化し、分析してきた [2]。環境配慮問題のほかにも、飲酒運転問題 [3]、ソフトウェアの違法コピー問題 [4]、ワクチン接種問題 [5] などにおける情報完備化コストについて分析してきた。このようにさまざまな問題をひとつひとつ扱ってきた理由としては、ポテンシャル効用モデルで扱える問題のクラスが特定化できていなかったということがあげられる。これは、ゲーム論などの確立したリサーチプログラムと違って、新しいモデル研究であることの必然的結果である。ひらたく述べると、ゲーム論のナッシュ均衡で説明できる例題を紹介しても、学術的な新規性、限界貢献はほぼないといえよう。ポテンシャル効用モデルにおいては、応用例を探すこと自体がモデル研究に大きく貢献する可能性をもつということである。

ところで、このモデルの本来的なポテンシャルは、啓蒙活動の表現・分析にとどまらない。自分の効用を誤認識していて、短期・中期的には、獲得する効用を認識できないケースであれば、このモデルが応用できるのである。とくに、本論文では、一般に広く見られる現象である、「おせっかい」、「だまし」という2つの対象について、これをポテンシャル効用モデルで表現することを試みる。具体的には、後の節で、過去に著者らが啓蒙問題として扱ってきた問題を、「だまし」、「おせっかい」の例としても用いる。同じ問題を別の構造・視点から扱うことで、ポテンシャル効用モデルの位置づけ、利用価値がより明確に伝わることを期待してこれらを題材に選んだ。

本論文の構成は以下の通りである。次節では、ポテンシャル効用モデルの位置づけの見直しを述べた後、モデルの基本構造について説明する。ポテンシャル効用モデルの基礎となるエージェント・モデルについても、新たな視点から簡単に説明する。3節では、著者らの過去の研究例と比較しながら、「おせっかい」と「だまし」という

ものがどういふものなのかを説明する。4節で、数理モデル化を行い（ただし今回は、具体的な関数形までは特定化しない）、最終節で今後の展望を述べる。

## 2 ポテンシャル効用モデルについての見直し

### 2-1 従来の理論との違いのレベルについて

著者らは、文献 [5] で、ポテンシャル効用モデルの限定合理性モデルの中での位置づけについて次のように述べた。少し長くなるが引用してみる。“ポテンシャル効用モデルは、意思決定主体の限定合理性を考慮したモデルのひとつといえる。著者は、このモデルが汎用性の高いものだとは思っているが、あらゆる問題に適用できると考えているわけではない。（中略）科学哲学の用語を用いるなら、スタンダードなミクロ経済学はリサーチプログラムの堅い核にあたり、それで説明できない現象に出くわしたときは、理論のアドホックな修正ではなく、前進的に保護帯を修正するというのが社会科学としての経済学、合理的意思決定理論および周辺理論の発展なのだと考えている。（中略）著者らが提唱し、発展させているポテンシャル効用モデルは、後で述べるように、従来の合理的意思決定モデルのひとつである2段階意思決定モデルを修正して使用している点などから、新しいリサーチプログラムというのではなく、従来の理論の修正を行うようなスタイルの研究であるということが出来る（引用中の参考文献は省略した）”

リサーチプログラム内での保護帯の修正か、新しいリサーチプログラムかという問題については、多分に主観が絡む問題であり、一概にはいえないのであるが、ポテンシャル効用の基礎付けを行った研究から、現在では、むしろ新しいリサーチプログラムへの移行と考えるのが自然であると認識している [8]。このあたりの哲学的、認識論的議論は、文献 [9] を参照されたい。

### 2-2 モデルの基本構造の見直し

ポテンシャル効用モデルは、非対称な情報をもった2主体の意思決定を扱うプリンシパル-エージェントモデル（以下エージェンシー・モ

デル）のひとつの発展形であることは何度か述べてきた。しかし、単に2段階意思決定であり、形が似ているからエージェンシー・モデルの一発展形なのであるという誤解を招いていた可能性もあるので、ここで、なぜポテンシャル効用モデルがエージェンシー・モデルの発展形といえるのかについて、より詳しく説明したい。実は、この基本的な部分の見直しから、今回のような適用例の拡張につながったのである。

通常のエージェンシー・モデルは、不完備情報2人ゲームのひとつのクラスと特徴づけることができる。これは、プリンシパルの意思決定の際の利得に、エージェントの意思決定が関わってくる、そして、エージェントの意思決定の際の利得に、プリンシパルの意思決定が関わってくるからである。このポテンシャル効用モデルにもそのことが必然的になる理由というものが存在する。それは、エージェント（情報不完備主体）が、自分自身の効用関数をわかっていないという点である。もし確率分布レベルでもわかっているのなら、それは単独主体の意思決定問題として表現できるのである。結局、情報不完備な主体をエージェントとみて、そのエージェントを上から眺めるかのような存在であるプリンシパルが必要になるのである。上から眺めるといっても、自分の望む方向にコントロールしようとする場合も、エージェントの真に望む方向にコントロールしようという場合も、プリンシパル自身もエージェントの効用を誤認識している場合も存在するが、ともかくそのような存在がどうしても必要になってくる。非常に極端なケースとしては、“将来のことに思いをはせる自分”がプリンシパルで、“せつな的な思考をする自分”がエージェントであるというようなモデリングも不可能ではない。

次に、ごく一般的なエージェンシー・モデルについて説明したい（より詳しい説明は、エージェンシー・モデルについてのすぐれた文献である、[10]などを参照されたい）。過去の文献で紹介した部分も多いが、本論文なりの立場から新しい解説を試みる。プリンシパル-エージェント関係とは、ある経済主体（プリンシパルとよぶ）が、自らの目的を、報酬と引き換えに、別の経済主体

(エージェントとよぶ)に遂行してもらおうとしているときの二者関係のことをいう。エージェンシー・モデルは、このような状況で、プリンシパルがエージェントを、いかに効率よく動機付けられるかについての方法を分析するための数理モデルである。典型的には以下のような2段階最適化問題として表現される。時系列的に示すと、まず、プリンシパルはエージェントに、成果と報酬の関係である契約を提示する。これはインセンティブシステム (IS) と呼ばれることもある。次に、エージェントはこの契約のもとで、自らの効用を最大化するように努力水準を決定する。そして、エージェントの努力に環境の影響が加わり成果が決定する。最後に、プリンシパルは成果を観察しこれを得る。ただし、エージェントの努力水準そのものは直接観察することはできない。これが2者の情報非対称性である。これを具体的な数理モデルとして表現すると

$$\begin{aligned} \max_{IS} & P(e_{opt}, IS) \\ s.t. & A(e_{opt}, IS) \geq B \\ & e_{opt} \in \arg \max_e A(e, IS) \end{aligned}$$

$P$ : プリンシパルの効用

$A$ : エージェントの効用

$B$ : 留保効用

$e$ : エージェントの努力水準

$IS$ : インセンティブ・システム (契約)

となる。

著者らは、エージェントの効用関数に、それまで考慮されてこなかった内発的動機付けを導入することで、さまざまな分析を行った。その到達点が動機付けコストモデルである。実はエージェントの効用関数において、外発的動機付けと内発的動機付け、それにコストを足し算するということは、微妙な問題をはらんでいることは文献 [11] で軽く触れた。特に重要な点は、なんといっても、内発的動機付けと外発的動機付けの相互作用の問題である。これはまさに内発的動機付けを導入したことによる固有の問題である。しかし実は、問題はそれだけではない。金銭にコストを加えてい

るという部分 (これは通常のエージェンシー・モデルでも行われていることである) などは、一般にはvNMの期待効用と同じようなものと誤解されることもあるが、両者の間には関係はない。むしろ、文献 [12] などのいう価値関数という用語の方が適切なくらいである。しかし、十分な理解のもとに「効用」という用語を使うこと自体は何ら問題のないことなので、今後、本論文でも、満足度のことを効用という用語で表現することにす

る。ポテンシャル効用における効用概念について、利己的効用と利他的効用にわけられる可能性があることについては文献 [5] で述べた。一方、別の分け方として、いわゆる手続き的効用と結果効用にわけられることも可能である。ポテンシャル効用モデルにおける手続き的効用は、上で述べたエージェンシー・モデルにおける内発的動機付けに近い役割をもつ。たとえば、内発的動機付けが絡むことによるクラウドディング効果と同じことが、ポテンシャル効用の手続き的効用でも生まれる可能性があるのである。たとえば、環境問題において、「自分が本当に環境に役立つことをしている」と確信すればするほど、直感的には、環境友好的な行為自体の手続き的効用が大きくなると考えられることなどは、まさに動機付け理論におけるクラウドディング効果に近いものがあるといえよう。このように、内発的動機付けを導入したエージェンシー・モデルとポテンシャル効用モデルは微妙な関係をもっている。

ポテンシャル効用モデルと、従来のエージェンシー・モデルの違いは、“エージェントは、自分自身の真の効用について分かっていない”、言い方を変えると、選択肢と効用の関係について、確率分布のレベルですらわかっていないということである。

また、著者らの過去の研究では、プリンシパルの効用は、エージェントの効用そのものであり、この点もポテンシャル効用モデルと通常のエージェンシー・モデルの違いであると説明してきた。しかし、実は、必ずしもそのように限定する必要はないことがわかった。むしろ、通常のエージェンシー問題と同様に、プリンシパル自身の利己的

な効用関数を想定することで、適用範囲が広がるのである。それによって「おせっかい」や「だまし」を表現しようというのが本論文の最大の目的である。留保効用制約は考えないという点は、従来までのポテンシャル効用モデルと同様の仮定を踏襲する。

現実のプリンシパル－エージェント関係では、エージェントの自律的、主体的決定がなされることも多いと考えられているので [13]、ポテンシャル効用モデルは、このような自律性、主体性という概念を考慮してエージェント・モデルを発展させたものと位置づけられる。このあたりの事情を、メタ的視点からくわしく説明したのが、文献 [14] である。そこでは、特に、経済学と心理学の複合領域のサーベイも丁寧に行った。

このモデルでは、プリンシパルは、情報を与えることでエージェントを間接的にコントロールし、エージェントは、プリンシパルから与えられた情報をもとに、自律的、主体的に意思決定をする。啓蒙のケースでは、プリンシパルが制御するというよりは、あくまで、エージェントの自律的、主体的決定をサポートするように情報完備化が行われているといえる。

### 3 具体例

この節では、「おせっかい」および「だまし」について、具体的な問題を用いて説明する。具体例としては進路指導、環境問題、飲酒運転、ソフトウェアの違法コピー、ワクチン接種といった、過去に啓蒙問題として扱ったものを取りあげる。

#### ア 進路指導

親が子供の進路について、あるいは、高校や大学の教員が生徒・学生の進路についてアドバイスをするというのはごく日常のことである。理想論を述べるなら、親や教員が進路指導する際には、親や教員の豊富な人生経験を生かして、本人にどんな適性があるのか、本人は潜在的にどんなことを望んでいるのかを、本人以上に正確に把握してやり、それを本人に伝え、主体的に望ましい選択肢をとらせてやるというのが、望ましい進路指導ということになろう。この問題において、“本人にどんな適性があるのか、本人は潜在的に

どんなことを望んでいるのか”こそが本人のポテンシャル効用を決定づけるものとなる。

しかし、現実には、上記のような理想論だけで指導が行われているとはいえないだろう。たとえば、本人は本当のところ、文系希望なのだが、親や教員からみると、理系を望んでいる、あるいは理系の素養がありそうなので、後で理系を選んだことを良かったと思うだろうと勝手に思い込み、そのように説得したとする。その結果、説得にかけたコストに応じて、進路を理系に決定する確率が高まったり、あるいは、文系と理系の中間、融合領域を希望したりするようになるというのが「おせっかい」問題の一例となる。

さらに、本人が文系希望であることを承知しているながら、親の都合、教員の外発的動機付けなどにより医学部を志望させるように説得するというようなケースが「だまし」にあたる。もちろん、将来医師になることできっとよかったと思えるときがあるだろうと思っているケースもあるだろうし、その場合は、「だまし」というよりは上記の「おせっかい」にあたるわけであるし、本当に数十年後に本人が医師になってよかったと思えるなら、「だまし」でも「おせっかい」でもなく、「啓蒙」であったということになる。

#### イ 環境問題

普通の市民は、どのような行為が長期的に地球環境に影響を及ぼすかについて、完全に正しい知識をもつことはほぼ不可能である。実際問題、専門家の間でも、CO<sub>2</sub>排出が本当に温暖化に結びついているのか、温暖化が本当に地球環境に悪いのかというレベルでも見解がわかれており、完全に正しい知識というのは、誰もわからないというのが実情であろう。しかしながら、一応、CO<sub>2</sub>を排出しないことが温暖化防止に結びつき、温暖化防止が地球の未来にとって望ましいという前提で話をすすめよう。

ある行為（たとえば電力を消費するなど）のCO<sub>2</sub>排出量について、本当の値を教えることで、市民が啓蒙され、自発的に環境友好的な行動をとるようになるというのが啓蒙モデルで分析したパターンであった。ところが、もし市民が、実は地球環境の将来よりも、身近な経済問題、家庭問題

を重要視していたとしたら、そういった市民に対して、きっと彼らも将来の地球の環境に思いをはせているに違いない、あるいはそのように思うべきと考えて、環境問題の重要性を訴えることは「おせっかい」となる。いってみれば、エージェントのポテンシャル効用関数では、環境改善効果からもたらされる効用が、他の問題の改善からもたらされる効用よりも相対的に小さいにもかかわらず、それを誤認識して、大きなコストをかけて啓蒙するというようなケースである。

一方、市民が環境問題などよりもっと身近な家庭、経済問題について考えることを重要視しているにも関わらず、環境有効的な製品を大量に販売するようなインセンティブをもつために、当局が市民に対して、CO<sub>2</sub>の排出と温暖化について重大な因果関係があるという情報を流し続けるなら、これは「だまし」となる。また、CO<sub>2</sub>の排出と温暖化の間の因果関係が不明であると仮定しよう。この因果関係が不明であるにも関わらず、当局自身のインセンティブのために、当局が市民に対して、過剰な情報流布を行うことも「だまし」ということになる。

#### ウ ソフトウェアの違法コピー問題

ソフトウェアの違法コピー問題の相当部分は、そうした行為の社会に対する影響（特にインセンティブや著作権に絡んだ社会厚生への損失）を大きく考えずに行われているものと推測できる。そこで、「違法コピーを行うことが著作権やインセンティブに対して重大な問題を引き起こし、社会厚生への損失につながる」ということを十分に市民に理解させることで、こうした行為を防ぐことができるだろうというのが啓蒙モデルである。

一方、市民がそれでもせつな的に、一見便利な方を選ぶというのに、「いや、そんなはずはない」と頑なにこれを「啓蒙」しようというのは、民主主義社会では「おせっかい」ということになろう。もちろん、いったん法律・ルールを作ったうえで、それを無視する市民を見過ごすのが良いといっているわけではない。法律・ルール作りの前段階での話である。また、ソフトウェア業界等からの圧力があつたとして、当局にインセンティブが存在した上で「啓蒙」しようというなら、「だまし」

ということになる。

## 4 数理モデル化

では、前節で述べた点をふまえて、実際に、モデルの適用範囲が拡大されることをみてみよう。具体的には、従来の啓蒙ケースのほかに、「おせっかい」ケース、「だまし」ケースに適用できることを、数理モデルで説明する。新しいモデルについて説明する前に、まず従来の啓蒙ケースについて説明しておく。

エージェント自身はポテンシャル効用を認識していないが、プリンシパルがそれをわかっており、プリンシパルがコストをかけて、エージェントに自分自身の真の効用を気づかせることを試みるようなケースがこれにあたる。これが、過去に著者が扱った啓蒙ケースの構造である。2者の意思決定問題の構造は以下ようになる。まず、プリンシパルが、正しい（「正しい」の意味は、エージェントの真の効用を決定づけるのに適切であるという意味である）情報をエージェントに与える。次に、情報完備化にかけたコストの大きさに応じてエージェントの効用関数が真の効用であるポテンシャル効用に近づく。そして、エージェントは、意思決定時の効用関数（誤認識している）のもとで自らの効用を最大化するような選択をする。プリンシパルは、エージェントが事後的に感じる効用であるポテンシャル効用とコストの差が最大になるように、かけるコストを意思決定する。

これを具体的な最適化問題として表現すると

$$\begin{aligned} \max_c & A^*(e_{opt}, c) - c \\ \text{s.t.} & e_{opt} \in \arg \max_e A(e, c) \end{aligned}$$

$A^*(e_{opt}, c) - c$  : プリンシパルの効用

$c$  : プリンシパルの意思決定変数（コスト）

$A(e_{opt}, c)$  : エージェントの意思決定時の効用

$A^*(e_{opt}, c)$  : エージェントのポテンシャル効用

$e$  : エージェントの努力水準

となる。

次に、これが本論文のテーマである、「おせっかい」、「だまし」ケースに適用できることをみて

みよう。

#### 4-1 「おせっかい」ケース

エージェント自身のみならず、プリンシパルもエージェントのポテンシャル効用関数を誤認識している場合がこれにあたる。エージェントが自分自身のポテンシャル効用について誤認識しておらず、プリンシパルのみ誤認識しているというケースも、現実には往々にして存在し、そのような場合も「おせっかい」と呼ぶのにふさわしいが、先に述べた理由で、この場合は、単独主体の意思決定問題としてモデル化するのが合理的であり、プリンシパル—エージェント関係を基礎にもつポテンシャル効用で扱う問題のクラスには入らない。

2者の意思決定問題の構造は以下のようになる。

- ① プリンシパルが、プリンシパル自身が正しいと考える情報をエージェントに与える。
- ② 情報完備化にかけたコストの大きさに応じて、エージェントの効用関数が、プリンシパルの誤認識しているエージェントの効用関数に近づく。
- ③ エージェントは、意思決定時の効用関数（誤認識している）のもとで自らの効用を最大化するような選択をする。
- ④ プリンシパルは、自分が誤認識しているエージェントのポテンシャル効用とコストの差が最大になるように、かけるコストを意思決定する。これを具体的な最適化問題として表現すると

$$\begin{aligned} \max_c & A^P(e_{opt}, c) - c \\ s.t. & e_{opt} \in \arg \max_e A(e, c) \end{aligned}$$

$A^P(e_{opt}, c) - c$  : プリンシパルの効用

$c$  : プリンシパルの意思決定変数（コスト）

$A(e_{opt}, c)$  : エージェントの意思決定時の効用

$A^P(e_{opt}, c)$  : プリンシパルが誤認識しているエージェントのポテンシャル効用関数

$e$  : エージェントの努力水準

となる。

#### 4-2 「だまし」ケース

エージェント自身はポテンシャル効用を認識していないが、プリンシパルはそれをおおまかにわかっており、さらにプリンシパルはエージェントに自分自身の真の効用を気づかせることよりも、プリンシパル自身の効用を上げるために、間違えた情報（「間違えた」の意味は、エージェントの真の効用を決定づけるのに適切でないという意味である）を教え込むことを試みるようなケースがこれにあたる。洗脳という用語を用いることも可能かもしれない。この場合、通常のエージェント・モデルとは逆に、プリンシパルの方にモラル・ハザードを起こす余地が出てくる。場合によっては、そのようなモラル・ハザードをブロックするための、より上位のメタ・プリンシパルとでもよぶような存在を考えることも可能である（例として、教員の利己的な教育を防ぐための教育委員会の存在などを想定されたい）。

次に数理的な表現を紹介しよう。今回は具体的な関数形までは踏み込まないが、以下の最適化問題こそがこのタイプの情報コントロール・モデルである。

- ① プリンシパルが、プリンシパル自身にとってもっとも都合のよい情報をエージェントに与える。
- ② 情報完備化にかけたコストの大きさに応じて、エージェントの効用関数が、プリンシパルにとってもっとも都合のよいエージェントの効用関数に近づく。
- ③ エージェントは、意思決定時の効用関数（誤認識している）のもとで自らの効用を最大化するような選択をする。
- ④ プリンシパルは、自分にもっとも都合のよい関数、つまりプリンシパル自身の効用関数の値とコストの差が最大になるように、かけるコストを意思決定する。これを具体的な最適化問題として表現すると

$$\begin{aligned} \max_c & P(e_{opt}, c) - c \\ s.t. & e_{opt} \in \arg \max_e A(e, c) \end{aligned}$$

$P(e_{opt}, c) - c$  : プリンシパルの効用

$c$  : プリンシパルの意思決定変数 (コスト)

$A(e_{opt}, c)$  : エージェントの意思決定時の効用

$P(e_{opt}, c)$  : プリンシパルにとって都合のよいエージェントの効用関数

$e$  : エージェントの努力水準

となる。

## 5 今後の展望

本論文では、ポテンシャル効用モデルの基礎を見直すことで、モデル分析を、「おせっかい」、「だまし」という行為にまで適用できることを説明した。過去の啓蒙問題のように、具体的な関数形の設定まではしておらず、比較静学分析のようなことは行えないが、「おせっかい」、「だまし」という行為・現象について、これらを数理的に表現できたのは一定の成果であろう。

最後に、今後の展望について触れたい。上に述べたように、具体的な関数設定を行うことが第一の課題である。著者らの過去の研究のように、解析的に解けるモデル化ができれば理想であるが、それが難しい場合、数値実験によるシミュレーションを行い、パラメータと解の関係について分析を行うことになろう。また、2節でも述べたように、ポテンシャル効用モデルは、それが応用できる問題のクラスを特定化することが、他の応用経済モデル以上に重要になるので、適用例の発見、適用範囲の拡大を目指しつつ、適用可能な問題のクラスが共通にもつ構造について分析していきたいと考えている。また、「おせっかい」は、パターンリズムと深い関係を持ち、パターンリズムについての人文・社会学的研究は多数存在するので、それらのサーベイを行いつつ、数理モデルの正当性、意義をより深く検証していきたい。

## 謝辞

本論文を執筆するにあたって、ポテンシャル効用モデルの共同開発者でもある、東京工業大学の小林憲正先生に、貴重なアドバイスを頂いた。こ

こに感謝の意を表したい。

## 参考文献

- [1] 松村良平、小林憲正、「ポテンシャル効用モデルの認識論的基礎」、計測自動制御学会第35回システム工学部研究会資料、pp.7-12、2005。
- [2] Matsumura, R. and Kobayashi, N., "The Theoretical Foundation of Potential Utility", The proceedings of the 1st World Congress of IFSR 2005 (CD-ROM).
- [3] 松村良平、「ポテンシャル効用モデルの飲酒運転問題への応用」、東洋大学経営論集 69号、pp.89-102、2007。
- [4] 松村良平、「新しい計量倫理アプローチ—ポテンシャル効用モデルの基本的考え方とその応用例について—」、現代社会研究、東洋大学現代社会総合研究所 4号、pp.33-40、2007。
- [5] 松村良平、「ポテンシャル効用モデルの新しい適用例と方向性について」、現代社会研究、東洋大学現代社会総合研究所 9号、pp.43-51、2012。
- [6] Benett, P. "Hypergames: developing a model of conflict", Futures, vol. 12, no. 6, 1980.
- [7] 高橋真吾、「ハイパーゲーム分析」、オペレーションズ・リサーチ、46 (2)、pp.73-78、2000。
- [8] Chalmers, A.F., "What Is This Thing Called Science?", Hackett Pub, 1976. (高田 紀代志、佐野 正博訳、「科学論の展開—科学と呼ばれているのは何なのか?」 恒星社厚生閣、1985。)
- [9] Kobayashi, N., "The Epistemic Foundation of Rational Choice", Ph.D. Thesis, Tokyo Institute of Technology, 2012.
- [10] 伊藤秀史、「契約の経済理論」、有斐閣、2003。
- [11] 一般関数モデルによる動機付けコスト問題の分析、東洋大学経営論集 79号、pp.113-123、2012。
- [12] Farquhar, P. and Keller, L., "Preference Intensity Measurement", Annals of Operations Research, 19, pp.205-217, 1989.
- [13] Kijima, K., "Intelligent poly-agent model and its application", Information and Systems Engineering, vol 2 1996.
- [14] Matsumura, R. and Kobayashi, N., "Incentive Design Utilizing Intrinsic Motivation", 'Handbook on Psychology of Motivation: New Research', Nova Publishers (in printing).