

成果主義賃金に関する行動経済学的分析

大洞 公平

(関西学院大学専任講師)

本稿では、成果主義的な報酬体系を契約理論的に分析する。さらに、行動経済学的視点から、報酬が労働者の心理的側面に与える影響も考慮したモデルを考察する。具体的には、まず、自分の報酬だけでなく他者の報酬にも依存する社会的選好を労働者がもつ場合、成果主義的な報酬体系で賃金格差が拡大するのはどのような状況かを考える。次に、労働者のモチベーションを高めるべく与えられたインセンティブが、むしろそれを低下させてしまうというモチベーションのクラウディング・アウト効果について経済学的に分析したモデルを紹介する。最後に、成果主義が機能するための条件の一つとして権限委譲の問題が取り上げられることがある。そこで、権限委譲に起因する心理的効用を考え、成果主義の有効性を考察した最近の研究を紹介する。

目次

- I はじめに
- II 成果主義と賃金格差
- III モチベーションのクラウディング・アウト
- IV 権限委譲と成果の評価
- V おわりに

I はじめに

バブル経済の崩壊とその後の長引く不況の中、それまでは日本の経済成長を支えてきたとして賞賛されていた日本の経営の諸特徴の評価も一変し、企業の置かれているビジネス環境の変化に対応して、その改革の必要性に迫られていた。その流れの中で、年功賃金をはじめとする報酬体系の見直しも始まり、労働者の成果に基づいて報酬額を決定する成果主義的な賃金制度が多くの企業で採用され始めた。しかし、一見単純に運用できそうなこの制度が現実にはなかなかうまく機能せず、最近ではその問題点が指摘されることが多い¹⁾。

本稿のテーマは成果主義的賃金制度を契約理論的に分析することである。さらに、インセンティ

ブがエージェントの心理的側面に与える影響にも焦点を当てるために、近年進展が目覚ましい行動経済学の成果をプリンシパル-エージェント・モデルに組み込んだモデルを考察する。ここで、行動経済学的視点が重要になってくるのは、成果主義的賃金制度の導入が直接的に労働者のインセンティブとして機能するだけでなく、労働者の心理的側面にも影響を与える可能性があり、それが企業の本来意図していた成果主義の効果と実際に生じる効果との乖離を生み出す原因の一つになっていると考えられるからである。

例えば、成果によって賃金が異なってくるため、成果主義的賃金制度の下では、当然、同じ職場の労働者の間で賃金格差が生まれてくる。その結果、この賃金格差が労働者のモチベーションに影響を与えることで、成果主義的賃金制度が企業の意図とは異なる方向へ作用する可能性が生まれてくる。また、報酬を決定する前提となる成果の評価がそもそも困難であるという問題もある。評価する側とされる側で、同じ成果であっても異なった評価を下すケースが多いことは想像に難くない。その場合、両者の評価の乖離が労働者のモチベーショ

ンに影響を与える可能性が生まれてくるということがここでも考えられる。このように、成果主義的賃金制度の導入は、労働者の仕事に対する直接的なインセンティブとして機能するだけでなく、そのあり方によって労働者の心理的な側面にも影響を及ぼす可能性が存在し、それが導入当初の意図とその成果の乖離を生み出してしまうのである。

このような点から、成果主義的賃金制度に関しては、インセンティブとしての報酬の直接的な効果のみならず、それがエージェントにもたらす心理的な効果も考慮することでより有意義な分析ができると考えられる。こうした考えから、本稿では、そのようなインセンティブの心理的な効果を考慮すると、スタンダードなモデルで得られる結果と比較して、どのような違いが生まれてくるかということいくつかの簡単なモデルに従って考察していくことにする²⁾。

本稿の構成は以下の通りである。まず、IIでは、エージェントの効用が自分の報酬だけでなく他者の報酬にも依存するような状況を考える。そのために、Fehr and Schmidt (1999) で定義された社会的選好 (social preferences) を説明する。その後、社会的選好も組み入れたモラル・ハザードのモデルを分析し、社会的選好を考慮することによって最適な報酬体系がどのような影響を受けるかを確認し、社会的選好の程度やプロジェクトの難しさなどに依存して、報酬を成果にどの程度依存させるべきかが決ってくることを見ていく。賃金格差の拡大を成果主義の広まりと捉えることができるならば、この結果からどのような状況では成果主義が好ましいかという点を指摘することができるだろう。次に、IIIでは、Frey (1997) らによって主張されている、インセンティブによるモチベーションのクラウディング・アウト効果 (crowding-out effects) について、簡単なモデルを交えて説明する。このクラウディング・アウト効果とは、心理学における外発的報酬と内発的動機づけの関係についての理論に基づく考え方であり、高橋 (2004) などでも成果主義に対する批判の根拠として用いられている。そこでIIIでは、クラウディング・アウト効果を定義して、それが存在する場合には、成果主義的なインセンティブはむしろ好

ましくない可能性が生まれることを Frey (1997) のモデルを紹介しながら確認する。最後に、IVでは、権限委譲の問題にとまなう心理的効用を考慮し、固定賃金契約と成果主義的賃金契約の比較を行う。成果主義的な報酬体系が一概に良いのか悪いのか判断することは実は現実的に困難なのかもしれない。しかし、労働経済学の実証的な研究を通して、成果主義がうまく機能するために必要ないくつかの条件が提示されている。その中の一つとして権限委譲が挙げられる場合がある。そこでは、成果主義的賃金制度は権限委譲をとらない、また、(権限委譲したことによってエージェントが高めた成果を報酬に反映させるという意味で) 公正な評価が行われる状況で、固定賃金契約よりも好ましくなることを示している。

II 成果主義と賃金格差

成果主義的賃金制度と一口に言っても、その詳細な内容はそれが導入される企業によって千差万別であろう。しかし、共通する成果主義的賃金制度の主要な特徴として、奥西 (2001) は次の3点を挙げている。

- (i) 賃金決定要因としてプロセスよりも結果を重視すること。
- (ii) 長期よりも短期の結果を重視すること。
- (iii) 賃金の格差をより拡大すること。

この三つ目の点に関して、賃金格差の広がりや成果主義の強まりとして解釈できる、と奥西 (2001) でも指摘されているように、成果主義的賃金制度では、よい成果が生まれたときとそうでないときでは賃金を大きく乖離させることが重要であり、それが労働者のインセンティブとして機能するような状況で有効に機能するということになる。つまり、成果主義的賃金制度を導入するにあたり、賃金格差が広がるようなインセンティブ・メカニズムとしての報酬体系が、企業にとって本当に好ましいのかどうかということが問題になってくるのである。

しかし、この議論は報酬が労働者にもたらす金銭的な効果にのみ基づいているということを前提としたものである。ここで注意すべきことは、成

果主義的賃金制度の結果としての賃金格差が、労働者の効用にそれとは別の影響を与える可能性があるということである。つまり、労働者は自分が受け取る賃金の大きさだけでなく、他者（例えば職場の同僚）の受け取る賃金と比べたときの相対的な大きさにも関心をもっている点に留意することが重要になる。Fehr and Schmidt (1999, 2003) は、通常の経済学のモデルで想定される金銭的な利得のみならず、他者の利得の大きさも反映させた効用関数を考え、このような効用関数をもつ経済主体を、社会的選好を有する経済主体と呼んで、その効果を分析している。

本節では、破産制約 (limited liability constraints) のあるモラル・ハザードのモデルの中で、一種の社会的選好をエージェントの効用関数に取り入れたモデルを分析する。それを通して、社会的選好を考慮することにより、インセンティブ・メカニズムとしての報酬体系がどのようなものになるか確認する。その結果、最適な報酬体系はプロジェクトの成功可能性や社会的選好の程度といったパラメータに依存して決まり、それゆえ、どのような状況で賃金格差をつけるべきか否か、つまり、成果主義的賃金制度がより好ましくなるのかといった問題を考えていく。

まず、スタンダードなモデルを考える³⁾。リスク中立的なプリンシパルとエージェントがいて、エージェントが一つのプロジェクトに従事し報酬がその成果に依存して支払われるケースを考える。エージェントがそのプロジェクトの遂行に費やす努力水準を e とし、その結果として生まれる成果を x とする。ここでは簡単に、エージェントの努力水準は高水準か低水準のいずれか、成果も成功 (x_s) か失敗 (x_f) のいずれかであるとする。エージェントが負担する努力のコストは、努力水準が高水準の場合は $d > 0$ 、低水準の場合は 0 としておく。また、成果に関しても $x_s = x > 0$ 、 $x_f = 0$ としておく。当然、努力水準が高いほど成功する確率も高くなると考える。具体的には、エージェントの努力水準が高水準であれば、成功する確率は p_1 、低水準であれば p_0 とし、 $0 < p_0 < p_1 < 1$ とする。プリンシパルはエージェントの努力水準は観察できず、そのため報酬は成果に

依存したものとなる。成果主義的賃金のもっとも単純な形式として、ここではプロジェクトが成功すれば報酬 w_s 、失敗すれば w_f を支払うものとし、簡単化のため $w_s = w \geq 0$ 、 $w_f = 0$ と仮定する。

このケースにおいて、プリンシパルがエージェントに提示する報酬 w がどのようになるか分析してみよう。エージェントの効用は、報酬から努力のコストを差し引いたものとして定義される。一方、プリンシパルの効用はプロジェクトの成果から報酬を差し引いた分で表されるとする。いま x は十分大きく、プリンシパルはエージェントの努力水準が高水準であることを望んでいるとしよう。その場合、プリンシパルは以下の条件を満たす w の中で彼女の期待支払額がもっとも小さくなるようなものを選べばよいことになる⁴⁾。

$$p_1 w - d \geq p_0 w, \quad (\text{IC})$$

$$p_1 w - d \geq 0, \quad (\text{IR})$$

$$w \geq 0. \quad (\text{LL})$$

(IC) の左辺は、エージェントの努力水準が高水準の場合の彼の期待効用、一方、その右辺は努力水準が低水準の場合の期待効用である。よって、(IC) はエージェントが低水準でなく高水準の努力水準を選ぶことを選好するための条件 (誘因整合性条件) である。(IR) は、エージェントがこの契約を受け入れないときに外部機会を利用して得られる留保効用をゼロとした場合に、エージェントがこの契約に同意するための条件 (個人合理性条件) である。最後に (LL) は、エージェントの報酬を負にするような契約は強制できないという破産制約を表している。

このケースの分析は容易である。まず、(IC) を整理すると、 $w \geq d/\Delta_p$ という条件を得る (ただし、 $\Delta_p \equiv p_1 - p_0$)。ここで、プリンシパルは可能な限り w を小さくしたいということ、また、 $w = d/\Delta_p$ は常に (IR) を満たしていることに留意すると、このケースの最適な報酬は $w^* = d/\Delta_p (> 0)$ となる。

次に、このモデルにエージェントの社会的選好を組み入れたモデルを考察する。賃金をエージェント間で比較するという点を考えるために、いま、2人のエージェントが存在するとしよう。ただし、

分析をごく簡単にするために、2人のエージェントは同質的で生産関係に相関がなく、それぞれ独立にプロジェクトに従事し、それぞれの成果に応じて独立に報酬が支払われるものとする⁵⁾。

ここでは、Fehr and Schmidt (1999) に従って社会的選好を定式化することにしよう。その場合、エージェントの効用関数の項目として、報酬、努力のコストに加え、次のような自分の報酬 (w) と他者 (同僚) の報酬 (\bar{w}) を比較したときに得る効用 S を考える⁶⁾。

$$S \equiv -\alpha \max\{\bar{w} - w, 0\} - \alpha\lambda \{w - \bar{w}, 0\}. \quad (1)$$

$\alpha > 0$ は社会的選好の程度を表す係数である。第1項は、もし自分の報酬よりも同僚の報酬のほうが高い場合、その差額分に α を掛けた分だけエージェントは心理的不効用を受けることになる。 λ はその値に従っていくつかのケースが考えられる。まず、 $\lambda > 0$ の場合、エージェントは自分と同僚との間に賃金格差があることをそもそも好まないという性質を表している。これは不平等回避 (Inequity Aversion) と呼ばれている。 $1 > \lambda$ の場合、自分の報酬が同僚に比べて高かろうが低かろうがとにかく賃金格差の存在を嫌っているものの、自分が同僚よりも報酬が多いときに感じる不効用は、自分の報酬が低いときに感じる不効用よりは小さいという性質を表すことになる。この性質は損失回避 (Loss Aversion) と呼ばれている。次に、 $0 > \lambda > -1$ の場合、エージェントは同僚に比べて自分の報酬が上回った場合にはむしろ心理的効用を感じるという状況を表していると解釈できる。以下では、 $-1 < \lambda < 1$ と仮定しておく。

このケースにおける最適な報酬を求めていこう。まず、このケースにおけるエージェントの誘因整合性条件は次のようになる。

$$\begin{aligned} p_1 w - d - \alpha w(1-p_1)p_1 - \alpha\lambda w p_1(1-p_1) \\ \geq p_0 w - \alpha w(1-p_0)p_1 - \alpha\lambda w p_0(1-p_1). \end{aligned} \quad (ICS)$$

(ICS) の左辺の第3項は、自分が確率 $1-p_1$ で失敗し報酬が得られず、同僚は確率 p_1 で成功して報酬 w を得た場合に被る心理的不効用を表す。

また同じく第4項は、自分が確率 p_1 で成功して報酬 w を得て、同僚が確率 $1-p_1$ で失敗して報酬が得られない場合に発生する心理的効用を表す⁷⁾。(ICS) の右辺も同様に考えられるが、この場合、このエージェントの努力水準が低水準であることから彼の成功確率が p_0 になっている。

一方、このケースのエージェントの個人合理性条件は、

$$p_1 w - d - \alpha w(1-p_1)p_1 - \alpha\lambda w p_1(1-p_1) \geq 0, \quad (IRS)$$

である。ここで、(ICS) と (IRS) を書き換えるとそれぞれ次のようになる。

$$w \geq \frac{d}{\Delta_p} \cdot \frac{1}{1 + \alpha\{p_1 - \lambda(1-p_1)\}}, \quad (ICS')$$

$$w \geq \frac{d}{p_1} \cdot \frac{1}{1 - \alpha(1-p_1)(1+\lambda)}. \quad (IRS')$$

ここで問題となるのは、(ICS') と (IRS') の右辺の大小関係である。 α が一定の値よりも小さい状況では、(ICS') の右辺は (IRS') の右辺よりも大きくなる⁸⁾。ここでは、社会的選好を考慮していないケースと同様に、誘引両立条件が有効になるこのケースに着目する。このとき、最適な報酬は (ICS') が等号で満たされる状況で決定し、次のようになる。

$$w^{**} = \frac{d}{\Delta_p} \cdot \frac{1}{1 + \alpha\{p_1 - \lambda(1-p_1)\}},$$

ここで、 w^{**} と w^* を比較すると、その大小関係は p_1 と λ の値に依存する。まず、 $\lambda < p_1/(1-p_1)$ のケースでは、 $w^{**} < w^*$ となる。つまり、社会的選好を考慮するならば、インセンティブを弱めるべく成果に依存した報酬はむしろ低くしたほうがよいということである。特に、 $\lambda < 0$ 、もしくは、 $1/2 < p_1$ のケースでは、常にこの関係が成立する。まず、 $\lambda < 0$ の場合、つまり、自分の報酬が同僚に比べて下回ることを嫌い、自分の報酬が同僚に比べて上回ることを好むようなエージェントの場合には、成果に依存した報酬はむしろ低くしたほうがよいということである。なぜなら、相手の成果にかかわらずとにかく自分が成功することを望むというこのタイプのエージェントの選好が、高い努力水準を選択するように仕向け

ており、プリンシパルがこの点を加味するならばインセンティブとしての報酬が低くてもエージェントは高い努力水準を選択してくれるのである。また、 $1/2 < p_1$ の場合、つまり、努力さえすれば高い確率でプロジェクトが成功するような場合には、低水準の努力水準を選択することで、相手が成功して報酬を得ているにもかかわらず、自分だけ失敗して報酬を得られない可能性が高くなってしまふ。そのとき、 $\lambda < 0$ であれば、当然、エージェントは自分の報酬のほうが低くなるのを嫌うし、また、 $0 < \lambda$ であっても自分と同僚との間に賃金格差があることを嫌う。そのため、エージェントはたとえインセンティブとしての報酬が低くても高水準の努力水準を選択するのである。よって、 $\lambda < p_1/(1-p_1)$ のケースでは、高い成果を上げたものとそうでないものとの賃金格差は縮小するのが好ましく、成果主義的な傾向はむしろ弱めたほうがよいということになる。

一方、 $p_1/(1-p_1) < \lambda$ のケースでは、 $w^{**} > w^*$ となる。つまり、社会的選好を考慮するならば、成果により強く依存した報酬が好ましいということである。この条件が成立するためには、少なくとも、 $p_1 < 1/2$ でなければならない。つまり、努力水準が高くてもプロジェクトの成功確率が低い状況に該当するといえる。また、このケースが成立するためには、当然、 $0 < \lambda$ (つまり、不平等回避) でなければならない。この場合、エージェントには、同僚が報酬をもらって自分だけがもらえない状況避けるために努力しようとするインセンティブと、自分だけが報酬をもらう状況避けようとして努力を怠ろうとするインセンティブという、相反するインセンティブが存在することになる。しかし、エージェントが自分だけが報酬をもらう状況を嫌う度合いが強くなるほど、後者のインセンティブが強くなる。また、そもそもプロジェクトが成功する可能性が極めて低いことから、同僚が失敗し報酬を得られない可能性が高く、 $0 < \lambda$ である限り、自分も失敗したほうが心理的不効用を軽減できるのである。このようなエージェントに対して、高水準の努力水準を選択するようにインセンティブを与えるために、成功した場合の報酬をより高くしなければならな

いということである。この結果から、労働者間で横並び意識が強い状況では、より困難なプロジェクトを成功に導くために、あえて賃金格差を拡大すべく、成果主義的な傾向を強めたほうがよい、というような示唆が得られるであろう。

Ⅲ モチベーションの クラウディング・アウト

成果主義的賃金制度に対する批判の一つとして最近よく耳にするのが、モチベーションのクラウディング・アウトに関する議論である。これは外発的な報酬が労働者の内発的動機づけを駆逐しそのパフォーマンスを低下させてしまう、という Deci (1975) をはじめとする心理学の研究から得られた知見の依拠するものである⁹⁾。例えば、ある労働者が自分の担当する職務を遂行すること自体に喜びを感じ満足しているような状況で、急に成果に依存した賃金が支払われた場合、まるで自分がお金のために働いていると思われて、逆に仕事に対する満足度が低下しモチベーションが下がってしまうということである¹⁰⁾。

この点に関してはいくつかの経済実験によっても確認されている。例えば、Fehr and Gächter (2002) や Irlenbusch and Sliwka (2005) は、固定的な賃金体系と比較して、成果主義的な報酬体系を導入したときに、エージェントの努力水準が低下し、効率性を阻害してしまうという実験結果を得ている。また、Gneezy and Rustichini (2000) は、成果に対する金銭的インセンティブとその成果は単調な関係にはないことを確認している。つまり、もともと成果主義的な報酬が与えられているのであれば、インセンティブを高めるほど労働者のパフォーマンスは高くなる。しかし、もともと成果主義的な報酬が与えられていなかった労働者にそのような報酬を適用すると、パフォーマンスはむしろ低下してしまうということである。Pokorny (2004) も Gneezy and Rustichini (2000) と同様にインセンティブとパフォーマンスが非単調な関係にあるという結果を得ているが、そこではインセンティブがない状況から次第にそれを強めていくに従って、パフォーマンスが最初

は上昇し、ある点を境にして低下するという逆U字型の関係にあることが示されている。

Frey (1997) は、報酬がもつこのような性質を報酬の隠されたコスト (hidden cost of reward) と呼んでいるが、Ⅲでは、彼のモデルに従ってこの点を見ていこう。エージェントの効用は、便益を B 、コストを C とすると、 $U = B - C$ で表される。また、 B 、 C ともに成果 x とプリンシパルから与えられるインセンティブ w に依存しているとすると $B(x, w)$ 、 $C(x, w)$ とそれぞれ表すことができる。なお、成果に関して B は凹関数、 C は凸関数とする¹¹⁾。このとき、エージェントにとって最適な成果 x^* は、 $B_x = C_x$ を満たす点で決まる。さらに、 x^* がプリンシパルによって与えられるインセンティブに依存する場合、 $B_x = C_x$ を w で偏微分することによって、その効果を表す dx^*/dw が次のように表現できる。

$$\frac{dx^*}{dw} = \frac{B_{xw} - C_{xw}}{C_{xx} - B_{xx}} \quad (2)$$

ここで、(2)の右辺の分母は正である。一方、分子の C_{xw} は価格効果を表し、 B_{xw} はクラウディング効果を表している。つまり、クラウディング効果とは、インセンティブが限界便益に与える影響をここでは指している。Frey (1997) は、 $B_{xw} > 0$ のケースをクラウディング・イン効果が存在するケース、 $B_{xw} < 0$ のケースをクラウディング・アウト効果が存在するケースと呼んでいる。ここで $C_{xw} = 0$ と単純化して考えると、クラウディング・アウトのケースでは、 $dx^*/dw < 0$ となる。この結果は、インセンティブとしての報酬を高めることで、エージェントの最適な成果水準がむしろ低くなってしまふことを示している。さらに、Frey (1997) では、 $dx^*/dw < 0$ のケースでは、最適な報酬は低下すべきであることを説明している。このことから、インセンティブとしての報酬がエージェントの限界便益に負の影響を及ぼす場合、インセンティブを強めることは、エージェントの成果をむしろ低めてしまう可能性があり、それを考慮するならば、プリンシパルはインセンティブを弱めることが望ましいというように解釈できることになる。これは、成果主義によってインセンティブを強めることが内発的

動機づけの減退を通してパフォーマンスを低下させる可能性を指摘しており、成果主義的賃金制度に対するこの視点からの批判をサポートしていると解釈できるであろう。ただし、 $B_{xw} > 0$ のケースではもちろん逆のことが起こっている。この結果は、インセンティブとしての報酬がそもそもエージェントの行動への動機づけとなるよう本来的な機能を果たしているときは、やはりインセンティブがエージェントの成果を高めることを示しているというように理解できるだろう¹²⁾。

Frey (1997) と同様に、Grepperud and Pederson (2006) は、プリンシパル・エージェント・モデルの中で線形契約を用いてクラウディング・アウト効果の分析を行っている。ここで、線形契約とは、固定給的部分と成果に依存した歩合給的部分の和として表現される賃金契約をさしている。例えば、固定賃金を γ 、成果 (x) に依存して支払われる賃金の度合い (インセンティブ強度: incentive intensity) を β とすると、線形契約は $w = \gamma + \beta x$ で表すことができる。これまで線形契約を前提とした契約理論の多くのモデルでは、その分析の主眼はインセンティブとリスクのトレード・オフであった。つまり、プロセスを評価せず (評価できず)、結果としての成果に基づく報酬がリスクを含んだものであると見るとき、このようなリスクを嫌うエージェントに対しては、成果に依存した契約は好ましくないという議論である¹³⁾。これに対して、Grepperud and Pederson (2006) は自分たちのモデルをインセンティブとリスクのトレード・オフではなく、インセンティブとモチベーションのトレード・オフを分析するモデルと位置づけている。本質的な点では Frey (1997) のモデルと共通するので詳細は省略するが、Grepperud and Pederson (2006) はエージェントの費用関数が彼の選択する努力水準だけでなく、プリンシパルが設定するインセンティブである β にも依存し、 β が大きくなるに従って、費用が増加し、また、努力水準に関する限界費用も増加するケースをクラウディング・アウトのケースとして考えている。この状況で、スタンダードなモデルと比較して、最適な努力水準のインセンティブに対する反応度が弱くなること、ま

た、最適なインセンティブが弱くなることを示している。

もっとも、このような結果はクラウドイング・アウトに関する直接的な仮定に強く依存している¹⁴⁾。これに対して、Sliwka (2003) はモチベーションのクラウドイング・アウト効果を内生的に分析するモデルを考えている。そこでは、二種類のタイプのエージェントを想定している。一つは、スタンダードなプリンシパル-エージェント・モデルで想定されているタイプであるが、もう一つは、機会主義的行動に走ることなく、プリンシパルが指示した努力水準を必ず選択する信頼できるエージェント (reliable agents) である。当然、信頼できるエージェントに対しては、プリンシパルは固定賃金を提示しておけばよい。そのとき、信頼できるエージェントの割合が高ければ、最適なインセンティブはスタンダードなケースよりも弱くなり、また、成果を評価するコストが大きいときには、成果主義的な報酬よりも、固定賃金が好ましくなることを示している。さらに、エージェントの利得が他のエージェントの利得だけでなく、彼らの信頼性に関する信念にも影響を受けるケースを考察している。その場合、より強いインセンティブは他のエージェントは信頼できるタイプでないということのシグナルとして機能し、その結果、エージェントの努力水準を低下させてしまうということも示している。

IV 権限委譲と成果の評価

ここまでの考察からわかったことは、心理的要因を考慮したケースでは、スタンダードなケースとは異なるインセンティブを与えることが好ましい状況が生まれてくるということである。ただ、最適なインセンティブは心理的要因の作用の仕方、また、その程度によってさまざまである。そう考えると、成果主義的賃金制度に関してさまざまな評価がなされているのは、理論的に考えてもそれほどおかしなことではないのかもしれない。一方、実証研究によって、成果主義的な報酬体系が有効に機能するための前提条件がいくつか指摘されている。例えば、大竹・唐渡 (2003) はホワイトカ

ラー、ブルーカラーに共通する条件として、(1) 仕事の分担や役割の明確化、(2) 仕事に対する責任の付与、(3) 能力開発の機会の提供、を挙げている¹⁵⁾。また、ホワイトカラーに関しては、仕事の裁量の程度を大きくすることも、重要な要因であると指摘している。

そこでIVでは、特に権限委譲という点に着目し、権限委譲の可能性にともなう心理的効用を考慮した場合の契約形態に関して分析している Daido (2006) を紹介する。そこではリスク中立的なエージェントがあるプロジェクトを遂行する状況を考え、その際に、プリンシパルは固定賃金契約 (fixed wage contracts: 以下、FWC) か成果主義的賃金契約 (pay-for-performance contracts: 以下、PPC) のいずれかを提供する。また、PPC の場合には、権限委譲するか否かを定める。ここでは、プリンシパルがプロジェクトの遂行法まで指示するケースを権限委譲のないケース、エージェントが遂行法を自ら見いだすケースを権限委譲されたケースと考える。この場合、エージェントの努力水準がたとえ同じであっても、権限委譲されるか否か、つまり、どのような遂行法でプロジェクトを行うかによって、その成果は異なってくるということが考えられる。この点を考慮すると、エージェントにとって、通常考えられる効用に加え、心理的要因に基づく効用が発生する可能性が見いだされる。ここで、権限委譲されないケースと権限委譲されたケースそれぞれに心理的効用が生まれる可能性がある点に注意したい。

まず、権限委譲されないケースを考える。このケースにおけるエージェントの心理的効用は、もし権限委譲されていたらより高い成果をもたらすような効率的な遂行法を見だし、それによってより高い報酬を得られていたにもかかわらず、権限委譲されていないために低い成果に甘んじざるをえないという点に起因する。つまり、権限委譲されていれば得られたであろう成果に対する報酬と権限委譲されていない場合のそれとの差が、心理的 (不) 効用をもたらすのである¹⁶⁾。次に、権限委譲されたケースの心理的効用について考える。この場合、エージェントは自らより効率的な遂行法を見だし、より高い成果を生み出す可能性が

あるのだが、問題はプリンシパルがその高まった成果をどのように評価するかという点である。プリンシパルが権限委譲したために高まった分の成果を十分評価しない場合、その過小評価に対してエージェントは心理的不効用を感じると考えられるのである¹⁷⁾。このような権限委譲の有無、また、それにとまなう心理的効用を考慮したPPCの各ケースをFWCと比較し、成果主義的な報酬の評価をモデル分析を通して行っている。

簡単にモデルを紹介しておこう。リスク中立的なプリンシパルとエージェントを考え、エージェントが1つのプロジェクトに従事する。プロジェクトの成果 x はエージェントの努力水準 $e \geq 0$ に依存し、 $x = e$ とする。この成果は立証可能である。プリンシパルは契約形態として、前述の固定賃金契約(FWC)か成果主義的賃金契約(PPC)のいずれかを選択できる。また、PPCのケースでは、業務の遂行法まで指示するか(権限委譲するケース)、それをエージェントに委ねるか(権限委譲しないケース)という選択もする。契約形態に関しては、それぞれ次のように考える。まず、FWCに関しては、プリンシパルが指示した努力水準をエージェントが固定賃金 w で選択するものとする¹⁸⁾。次にPPCについては線形契約 $w(x) = \gamma + \beta x$ を考える。

権限委譲に関しては次のように考える。エージェントが自ら選んだ遂行法をとった場合のプロジェクトの成果を μe で表されるとしよう。 $\mu > 1$ ($\mu < 1$)の場合、エージェントに権限委譲したほうがより高い(低い)成果が実現すると考えられる。以下では、 $\mu > 1$ のケースに着目することにする。この場合、前述した権限委譲の可能性にとまなう心理的効用は以下のように考えられる。まず、権限委譲されないケースを考える。権限委譲されないケースでの心理的効用の源泉は、もし権限委譲されていれば実現したであろう成果に基づく報酬とそれがなされなかったために実現した成果に基づく報酬の乖離にある。 $\mu > 1$ の場合、前者が後者よりも大きければ、エージェントは負の心理的効用を受ける。権限委譲されなかったことによる心理的不効用を加味した場合のエージェントの期待効用は

$$U = \gamma + \beta e - c(e) - \alpha(\mu - 1)\beta e,$$

となる。 $c(e)$ はエージェントの努力のコストを表している。また、最後の項が権限委譲されていれば得られたであろう成果に依存した報酬と実際の報酬の差から受ける心理的不効用を表している¹⁹⁾。

次に、権限委譲されたケースの心理的効用について定式する。このケースで実現する成果は $x = \mu e$ である。しかし、これはエージェントが自ら選んだ遂行法による成果によるものであり、それをどの程度評価するかはプリンシパル次第であるとする。そこで、プリンシパルのこの成果に対する評価の度合いを $\theta (> 0)$ で表すことにする。もし、 $\mu > \theta$ であれば、プリンシパルのエージェントの成果に対する評価は過小評価であり、エージェントからすれば、せっかく実現した高い成果を十分評価してもらえないとして不満を抱くと考えられる。一方、 $\mu < \theta$ であれば、プリンシパルは実際の成果以上に過大評価していることになる。以下では、 $\mu > \theta$ のケースのみを考える。ここでは、 θ が大きいほど、プリンシパルは公正な評価をしていると解釈することができよう。そのとき、エージェントの期待効用は次のように表される。

$$U = \gamma + \beta \theta e - c(e) - \alpha(\mu - \theta)\beta e.$$

なお、プリンシパルの期待利潤は成果の期待値からエージェントの報酬の期待値を差し引いたものである。つまり、FWCのケースは $e - w$ 、権限委譲しないPPCのケースは $e - (\gamma + \beta e)$ 、権限委譲するPPCのケースは $\mu e - (\gamma + \beta \theta e)$ となる。

以上のモデルを通して得られる結果を紹介しよう。まず、エージェントの努力水準が観察可能でそれをFWCで強制できるケースと心理的効用を加味していないPPCのケースを比較すると、エージェントがリスク中立的である限り、エージェントの最適な努力水準、および、プリンシパルの期待利潤は一致する。つまり、このような環境においては、プリンシパルからすればFWCとPPCいずれを選択しても実現する期待利潤は等しくなるのである²⁰⁾。

次に、FWCと権限委譲しないPPCを比較してみよう。まず、最適な努力水準に関してはFWCのほうが高くなる。つまり、心理的な効用を考慮すれば、インセンティブとしてのPPCを導入することが努力水準をむしろ押し下げてしまうのである。これは、インセンティブが努力水準に負の効果を与えるという意味でのモチベーションのクラウディング・アウト効果として捉えられよう²¹⁾。また、プリンシパルの期待利潤に関しても、FWCのほうが高くなる。つまり、心理的効用を考慮しなければ、権限委譲しないPPCとFWCは等価であったにもかかわらず、心理的効用を加味すると、プリンシパルにとってはFWCのほうが権限委譲を伴わないPPCよりも好ましくなるということである。以上より、権限委譲をとまわらない成果主義的賃金制度の導入は、プリンシパルにとって好ましくない結果を生む可能性があるということがわかる。

最後に、権限委譲をとまわうPPCを前の二つと比較する。このケースでは、プリンシパルからすれば評価が公正であればあるほど(θ が高いほど)、エージェントへの支払額は高くなるが、彼の心理的な不効用を軽減することができ、その結果として高い努力水準を引き出すことができる。また、 θ が高くなるにつれて、後者の効果が大きくなることから、プリンシパルの期待利潤は θ に関する増加関数となる。この結果、権限委譲をとまわうPPCは、 θ の値が大きくなるに従って、権限委譲をとまわらないPPCよりも、さらに、FWCよりもプリンシパルにとって好ましくなる。ちなみに、 $\theta > 1$ 、つまり、プリンシパルが自分が指示した遂行法で実現する成果の分だけは最低限評価するようなケースであれば、権限委譲をとまわうPPCは常にもっとも好ましくなる。この結果から、権限委譲をとまない、また、成果を公正に評価することによって、成果主義がそれを導入する企業にとってより好ましい制度として機能するというように理解できるのである。

V おわりに

本稿では、エージェントの効用に心理的側面を

組み入れたプリンシパル-エージェント・モデルを分析し、スタンダードなモデルで得られる結果との比較を通して、心理的効用がインセンティブに与える影響を考察してきた。その結果から、成果主義的賃金制度に関する評価に関して、行動経済学的な視点からどのような解釈ができるかということ述べてきた。エージェントが社会的選好をもつ場合には、その程度やプロジェクトの難しさに応じて、賃金格差を広げるべきか否か、また、高い報酬が常にエージェントのパフォーマンスを引き上げるとは限らないこと、さらに、成果主義を導入するのであれば権限委譲をとまわう必要があること、を簡単なモデルに従って分析した。

行動経済学の急速な発展の中、本稿で取り上げたテーマと関連するより興味深いモデルも存在する。例えば、IIの社会的選好のモデルに関して、脚注で触れたように、Itoh (2004)は、2人のエージェント間でそれぞれの成果に依存して報酬が決まるケースを分析し、社会的選好の性質に応じて最適な契約がどのようになるかを示している。さらに、Neilson and Stowe (2004)は、社会的選好をもつエージェントに線形契約を提示するケースに関して同様の分析を行っている。また、IIIの内発的動機づけのクラウディング・アウトに関連する文献として、Bénabou and Tirole (2003)は、職務の難しさ、もしくは、プリンシパルがエージェントの能力に関する私的情報を持っている場合、プリンシパルが提示する賃金がエージェントの能力に関するシグナルとして機能し、高い賃金がエージェントのモチベーションをむしろ引き下げてしまう可能性があることなどを示している。

このように、行動経済学では、心理学などから得られる知見を基にして、これまでに蓄積されてきた経済理論を拡張しより広範なテーマを扱えるモデルを構築すべく多くの研究が輩出されてきている。今後は、これらの成果をより包括的に説明できる、ある程度統一されたモデルが要求されてくるかもしれない。それによって、本稿で取り上げた成果主義をはじめとする労働経済学における重要なトピックスの分析に、これまでになされなかった有意義な解釈を与えられる可能性が生まれてくることが期待される。

- 1) 例えば, 城 (2004), 高橋 (2004) など参照。
- 2) このような目的から, II以降で説明するモデルは, 厳密性を多少犠牲にすることを厭わず, できるだけ平易に説明できるように詳細は割愛している。ただし, より詳細な分析を行っている参考文献をそれぞれ挙げていたので, 厳密な議論に興味ある読者はそちらも参考にしていただきたい。
- 3) 契約理論の詳細に関しては伊藤 (2003) を参照。また, ここでのモデルは其中でも, 5.1 破産制約のモデル, がベースになっている。
- 4) 以下はこの分野の慣例に従って, プリンシパルを「彼女」, エージェントを「彼」と呼ぶことにする。
- 5) Itoh (2004) は, エージェント間でそれぞれの成果に依存して報酬が決まるケースを分析し, 社会的選好の状況に応じて最適契約がどのような特徴を有するものになるか分析している。
- 6) 以下, 本稿では, Sのような心理的要因を加味した際に追加的に考慮している効用を「心理的効用」と呼ぶことにする。
- 7) なお, 同僚は高水準の努力水準を選択しているという前提である点に留意すること。また, 社会的選好を考慮するケースでは $w_j = 0$ となる条件を確認しておく必要があるが, ここでは割愛する。
- 8) 厳密には,

$$\frac{1}{(1-p_1)(1+\lambda)+l-1} \geq \alpha,$$
 ただし $l \equiv p_1/p_0 > 1$, のときである。
- 9) この文脈でクラウドイング・アウトという用語を使ったのは, 私が知る限り Frey (1997) が最初である。
- 10) その他, モチベーションのクラウドイング・アウトに関する議論は, Frey (1997), Frey and Jegen (2001) などを参照のこと。
- 11) つまり, $B_x > 0$, $B_{xx} < 0$, $C_x > 0$, $C_{xx} > 0$ を仮定する。なお, B_i は B の i についての偏導関数, B_{ij} は B_i の j に関する二次偏導関数を表している。ただし, $i, j = x, w$ で, C に関しても同様とする。
- 12) このクラウドイング・アウト効果に基づく成果主義に対する批判は高橋 (2004), またその批判に対する問題の指摘は中村 (2006) を参照。
- 13) このようなインセンティブとリスクの負の相関に関して疑問視する研究も最近ある。例えば, Prendergast (2002) はこの二つの関係が正の相関をもつケースを分析している。なお, 石黒 (2005) はこの二つのケースの理論モデルを簡潔にまとめている。
- 14) この点は, Grepperud and Pederson (2006) も, 自分たちのモデルが内発的動機づけとそのクラウドイング・アウトに関する初歩的なものでありさらなる研究が必要であることを脚注で断っている。
- 15) これに先立って, 玄田・神林・篠崎 (2001) でも同様な指摘がなされている。
- 16) エージェントが実際に権限委譲されても良い遂行法を自分では見つけられないということも考えられる。その場合, このケースはエージェントが自信過剰 (overconfidence) な状況として捉えることもできよう。
- 17) プリンシパルが支払いをしない理由としては, 成果の増加分は立証不可能であることなどが考えられる。
- 18) このような契約が可能になるためには, 本来, エージェントの努力水準が観察可能であったり, モニタリングの可能性などを考える必要がある。しかし, ここではその点は無視して, 前述した Sliwka (2003) で想定されているような信頼

できるタイプのエージェント (reliable agents) を考えておくことにする。

- 19) ここでは, 基本的に(1)の定式化に従っており, $\alpha > 0$ は心理的効用の程度を表している。さらに, (1)に従えば, 当然 $\mu < 1$ のケースも考慮することができるが, ここでは割愛する。
- 20) 先述したように, ここでは努力水準や成果のモニタリング・コストは無視している。なお, Prendergast (2002) では, 成果に関するモニタリング・コストのほうが努力水準に関するそれよりも大きいと仮定している。一方, Barth *et al.* (2006) では, FWCにおける努力水準のモニタリング・コストのみを考慮している。このような点を考慮すると, 2つのモニタリング・コストの大小関係はそれほど明確ではないといえるであろう。
- 21) なお, ここでいうクラウドイング・アウト効果とは, あくまで FWC と PPC を比較したときに努力水準が低下するという意味であって, PPCの中でインセンティブ強度が強くなると努力水準が低下するという意味ではないことに注意したい。

参考文献

- 石黒真吾 (2005) 「職務発明の経済分析——契約理論的接近」『日本労働研究雑誌』No. 541, pp. 24-33.
- 伊藤秀史 (2003) 『契約の経済理論』有斐閣。
- 大竹文雄・唐渡広志 (2003) 「成果主義賃金制度と労働意欲」『経済研究』Vol. 54, No. 3, pp. 1-20.
- 奥西好夫 (2001) 「「成果主義」賃金導入の条件」『組織科学』Vol. 34, No. 3, pp. 6-17.
- 玄田有史・神林龍・篠崎武久 (2001) 「成果主義と能力開発——結果としての労働意欲」『組織科学』Vol. 34, No. 3, pp. 18-31.
- 城繁幸 (2004) 『内側から見た富士通「成果主義」の崩壊』光文社。
- 高橋伸夫 (2004) 『虚妄の成果主義——日本型年功制復活のススメ』日経 BP 社。
- 中村圭介 (2006) 『成果主義の真実』東洋経済新報社。
- Barth, E., B. Bratsberg, T. Hægeland, and O. Raaum (2006) “Who Pays for Performance?” IZA Discussion Paper, No. 2142.
- Bénabou, R. and J. Tirole (2003) “Intrinsic and Extrinsic Motivation,” *Review of Economic Studies*, Vol. 70, pp. 489-520.
- Daido, K. (2006) “Incentives, Delegation, and Performance Evaluation,” mimeo.
- Fehr, E. and S. Gächter (2002) “Do Incentive Contracts Crowd out Voluntary Cooperation?” Institute for Empirical Research in Economics, University of Zürich, Working Paper No. 34.
- Fehr, E. and K. M. Schmidt (1999) “A Theory of Fairness, Competition, and Cooperation,” *Quarterly Journal of Economics*, Vol. 114, No. 3, pp. 817-868.
- Fehr, E. and K. M. Schmidt (2003) “Theories of Fairness and Reciprocity: Evidence and Economic Applications” in Dewatripont, M., L. P. Hansen, and S. J. Turnovsky, eds., *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress, Volume 1*, Cambridge University Press, NY, 208-257.
- Frey, B. S. (1997) *Not Just For the Money: An Economic*

- Theory of Personal Motivation* Cheltenham, UK and Brookfield, USA, Edward Elgar Publishing.
- Frey, B. S. and R. Jegen (2001) "Motivation Crowding Theory," *Journal of Economic Surveys*, Vol. 15, No. 5, pp. 589-611.
- Gneezy, U., and A. Rustichini (2000) "Pay Enough or Don't Pay at All," *Quarterly Journal of Economics*, Vol. 115, No. 3, pp. 791-810.
- Grepperud, S. and P. A. Pederson (2006) "Crowding Effects and Work Ethics," *LABOUR*, Vol. 20, No. 1, pp. 125-138.
- Irlenbusch, B. and D. Sliwka (2005) "Incentives, Decision Frames, and Motivation Crowding out: An Experimental Investigation," IZA Discussion Paper No. 1758.
- Itoh, H. (2004) "Moral Hazard and Other-Regarding Preferences," *Japanese Economic Review*, Vol. 55, No. 1, pp. 18-45.
- Neilson, W. S. and J. Stowe (2004) "Piece-Rate Contracts for Other-Regarding Workers," mimeo.
- Pokorny, K. (2004) "Pay-But Don't Pay Too Much: An Experimental Study on the Impact of Incentives," mimeo.
- Prendergast, C. (2002) "The Tenuous Tradeoff between Risk and Incentives," *Journal of Political Economy*, Vol. 110, No. 5, pp. 1071-1102.
- Sliwka, D. (2003) "On the Hidden Costs of Incentive Schemes," IZA Discussion Paper No. 844.

だいどう・こうへい 関西学院大学経済学部専任講師。主な論文に "Formal and Relational Incentives in a Multitask Model" forthcoming in *International Review of Law and Economics*. 契約理論, 法と経済学, 行動経済学専攻。