

評価制度の経済学

——設計上の問題を理解する

大湾 秀雄

(東京大学教授)

これまで経済学で扱われてきた評価制度の設計に係わる諸問題を、(1) 主観的指標が機能する条件、(2) 評価指標の組み合わせ、(3) 複数任務に起因するエージェンシー問題、(4) 評価におけるバイアス、(5) 評価指標の操作（ゲーミング）という5つのテーマに分類して整理した。こうした問題を包括的に議論することで、客観的指標と主観的指標、アウトプット指標とインプット指標、絶対的指標と相対的指標、集団業績指標と個人業績指標といった様々な特性を持ちうる評価指標をどう組み合わせ、どのような点に配慮して制度設計を行ったらよいか、経済学文献の提供する知見をまとめてみた。

目次

- I はじめに
- II どのような状況で主観的指標は有効か？
- III 異なる評価指標をどう組み合わせるべきか？
- IV 評価制度設計の基本——マルチタスク問題
- V 評価におけるバイアス
- VI 評価指標の操作（ゲーミング）
- VII 最後に

I はじめに

評価制度を設計することは難しい。多くの人事制度改革の実施段階で、従業員の不満の原因となるのが評価制度である。設計上の欠陥や評価バイアスへの対処の欠落が勤労意欲を下げ、望ましくない行動を引き起こす場合も少なくない。他方、失われた20年を経て、成長に伴う成功の機会の提供ではなく、正しい貢献の評価に基づく報奨によって従業員のやる気を高めることが必要という認識が高まったことで、多くの企業が試行錯誤を繰り返している。そうした中、評価制度に絡む問題と対処法が理路整然と解説された手引き本が巷

には見当たらないことは問題であると感じる。

経済学は抽象的な学問であり、組織に関する多くの研究もエージェンシー理論など数理モデルを使ったメカニズムの解明が基礎となっている。しかし、過去20年程度の研究の蓄積を見ると、評価制度一つを取っても多くの知見を提供してくれる。これを整理することで、人事制度に携わる研究者、専門家、実務家の方々の参考の用に供するというのが本稿の目的である。

評価制度の設計を行う上で、2つの重要な事実を認識して細部に入ることが望ましい。まず、評価制度と一口に言っても、報酬制度に密接に結びついているものから、人材配置や人材開発を主目的に使われるものまで、目的は様々である。使われる技術や業務の性格、職場組織の設計によって、かなり客観的な業績指標が存在するケースから、ほぼ上司の主観的な評価に頼らざるをえない状況まで、制度設計の制約条件も大きく異なる。したがって、文献で紹介される知見を正しく応用するためには、評価の目的、評価技術、職場組織についてどのような仮定に基づいて論じられているのか、常に念頭に置きながら読み進めることが

大事であると考え。二番目に、通常、企業は、評価制度だけを単体で見直すことは稀であり、組織体制変更、昇進制度、報酬制度の変更などを含む人事制度全体の改革を意図して評価制度に修正を加えるケースがほとんどである。したがって、評価制度に関する多くの経済学の論文も、報酬制度設計、職の設計、企業の所有構造まで含めて論じられることが多い。すべてが内生的に定まる最適人事制度設計を常に目指しながら、評価制度に目を向ける、「森を見て、木を見る」姿勢が重要であると考え。

これまで経済学で扱われてきた評価制度の設計に係わる諸問題を、以下の5つに分類して議論を進めたい。

- (1) 主観的指標が機能する条件
- (2) 評価指標の組み合わせ
- (3) 複数任務に起因するエージェンシー問題
- (4) 評価におけるバイアス
- (5) 評価指標の操作（ゲーミング）

評価指標には、大きく分けて、客観的指標と主観的指標、アウトプット指標とインプット指標、絶対的指標と相対的指標、集団業績指標と個人業績指標という4つの軸での分類が可能である。任務の数や評価の次元の数が複数あるために、複数の評価指標を組み合わせる望ましい努力配分を促す必要が出てくるケースも多い。多くの先行研究がどの指標を使うべきか、どう組み合わせるべきか、報酬と業績指標の関係はどの程度強めたらよいかについて分析を行ってきた。その主な議論を網羅したいと考えている。

契約理論で長年注目されてきたのは、主観的指標を使った雇用契約をどう違反なしに履行できるかという問題である。まずその問題から議論を始めたい。

II どのような状況で主観的指標は有効か？

1 多くの先行研究は、評判のメカニズムを重視

契約理論上、客観的指標と主観的指標の違いは、契約で合意事項の実施が強制できるかどうか

であると一般に認識されている。例えば、ある目標を達成した時にボーナスを支払うという労使間の合意があった時に、実際に目標が達成されたかどうか、裁判所など外部機関が確認することが容易であれば、その目標は客観的であるし、そうでなければ主観的である。販売社員と雇用主の間で「販売によって会社が得た粗利益の10%をコミッションとして支払う」という契約が存在する場合に、仮に経営側が10%ではなく5%分しか支払わなければ、従業員は裁判所に訴えることが出来るが、「会社に対し卓越した貢献を行った場合に特別賞与を支払う」という会社の約束に対し、その不履行を裁判所に訴えて勝訴することは難しい。

そのため、金銭報酬と結びついた主観的業績指標が意味を持つためには、それに基づく合意に企業をコミットさせる何らかの評判のメカニズム(reputation mechanism)が必要である。つまり、雇用関係が長期であるか、もしくは一人の社員に対する企業の不誠実な行動が社内の他の社員に即座に知れ渡ることを前提にして、正しく業績を評価しない企業は従業員の勤労意欲低下によって罰されるというメカニズムが働く必要がある。したがって、主観的業績指標を前提とする契約の研究は、繰り返しゲームを使うのが通例である(Bull 1987; MacLeod and Malcolmson 1989; Pearce and Stacchetti 1998; Levin 2003)。これらの文献に共通する仮定は、(裁判所は確認出来ないけれども)契約の当事者である従業員あるいは彼/彼女の同僚は、雇い主が公平に業績を評価したかどうか正確に監視することが出来るということである。つまり従業員がある一定の業績基準を満たしたかどうかについて、労使双方が同じ認識を持つ必要がある。

2 衝突コストによる説明

こうした仮定は、実際の主観的指標に基づく評価制度を念頭に置いた時、かなり強すぎる仮定と言える。何故なら、従業員の自らの貢献に対する認識と上司の評価が全く異なるという状況は現実には頻繁に発生しうるからである。その場合、上述の評判のメカニズムは上手く機能しない。MacLeod (2003)は、雇用主と従業員が後者の業

績について異なるシグナルを受け取るという仮定の下で、双方の主観的評価の報告に基づく最適契約がどのような構造をもつか分析を行った。彼のモデルは繰り返しゲームではないので、雇用主と従業員双方に正直な報告を動機づけるため、前者が支払う賃金と後者が受け取って消費する収入が必ずしも一致しない（経済学用語で言うと、予算均衡条件が成立しない）。つまり、雇用主が業績が低いと判断して従業員の収入を減らしても、従業員が自らの業績が高いと報告すれば雇用主が支払う人件費は従業員の受け取り収入を上回る。差額がどちらのポケットにも入らないということにコミット出来れば、第三者に支払われようと溝に捨てられようと問題ではない。こうした性格を持つ契約を許容することは一見奇妙であるが、その差を労使間の意見の不一致に伴うコスト（衝突コスト）、つまり勤労意欲低下、サボタージュ、組合活動活発化による生産性低下などと解釈することによってある程度の妥当性が見えてくる。結果は示唆に富んでいる。

まず、従業員の受け取る賃金収入は雇用主の主観的評価のみに依存するが、両者の評価がずれる場合、衝突コストが増えることで、従業員の自己評価は雇用主に公正な評価の報告を動機づける。2番目に、雇用主と従業員の受け取るシグナルが完全に相関している場合、主観的評価に基づく最適契約は、客観的評価に基づく最適契約（つまりシグナルの報告ではなく、シグナルそのものに基づく契約）と一致する。3番目に、雇用主と従業員の受け取るシグナルが不完全に相関している場合、衝突コストを抑えるため、従業員の賃金収入の評価による格差は縮小する。これは人事の実務で広く知られた中心化傾向と整合的である（この点については、Vで再度触れる）。最後に、両者の報告の間に全く相関がない場合、最適契約は2段階の賃金レベルしかないボーナス契約となり、雇用主は最悪のシグナルを受け取った時だけボーナスの支払いを拒否する。こうした評価報酬契約の形は、客観的指標を前提とした最適契約に関する先行研究結果に比べ、現実に観測される慣行により近い。

3 昇進審査における主観的評価の利用

これまでの議論では、主観的評価を用いた報酬制度が有効に機能するためには、長期雇用関係に基づく評判のメカニズムが働くことや、評価が不公正に捻じ曲げられた時に何らかの衝突コストが発生することが不可欠であった。しかし、主観的評価結果を昇進や解雇の意思決定に用いる場合は、公平な運用を促す他のメカニズムが働く。例えば、Prendergast (1993) は、従業員の企業特殊的人的資本の蓄積が主観的評価によって図られ、それが昇進を決める際の情報として使われる状況を分析した。仮に企業が事前に雇用契約で決められた職階と賃金の組み合わせにコミット出来る場合、公平な評価をしないことが従業員の昇進の意思決定を歪め、それは非効率な人材配置という形で企業利益を下げるので、より公正な評価を誘因すると分析した。

しかしこの場合でも、企業特殊的人的資本の蓄積が十分に高い場合のみ正しい昇進が決定され、必ずしもすべての昇進審査が効率的に行われるわけではない。特に、人的資本が一般的である場合、昇進によって従業員の能力が外部に知れ渡るため、企業は市場にシグナルされた能力に見合った賃金の支払いを強いられる。このため、評価を下げ昇進を遅らせることによって賃金の上昇を抑えようとする誘因が働く (Waldman 1984)。この場合、昇進させないで雇用関係を維持する選択が常に許される通常の昇進制度よりも、ある一定の時期に同年入社の従業員全員の昇進審査を行い、昇進できないものはすべて解雇するという「昇格か解雇の二者択一を企業に迫る契約」(up-or-out contract)の方が優れている (Ghosh and Waldman 2010)。つまり通常の昇進制度の下では非効率な人材配置が発生しうるが、その頻度は企業特殊的人的資本が大きい時は小さい。他方、二者択一契約は、昇進させず低いポジションで雇用し続ける方が望ましい場合でも解雇を強制するので非効率な離職が発生する。企業特殊的人的資本が小さい時は、この離職コストは小さい。いずれの場合も、企業が適切な昇進制度設計をするのであれば、評価指標が主観的であるが故に昇進審査を通じ非効

率な人的資源配分が発生する危険性は比較的小さいと言えるかもしれない。

Ⅲ 異なる評価指標をどう組み合わせるべきか？

1 客観的指標と主観的指標

よく聞かれる議論は、客観的指標と主観的指標の二つの指標が利用可能である時に、どちらを使うべきか、あるいはどう組み合わせるべきかという問題である。この問題については、客観的評価と主観的評価の間に補完性が働いているという研究結果が複数あり参考になる。

Pearce and Stacchetti (1998) は、雇用主が様々な不確定要素を含む客観的業績指標で捉えられるよりも多くの情報を持っている場合、客観的指標に基づく給与と主観的指標に基づくボーナスを両方活用することが効率的であることを示した。客観的指標に基づく給与のみによってインセンティブを与える場合、客観的指標は従業員のコントロールできない多くの不確定要素を含むため、従業員がリスク回避型であれば、効率的な努力水準を達成できない¹⁾。他方、主観的指標に基づくボーナスのみによって高いモチベーションを図る場合、高い評価と低い評価の場合のボーナス支給額の差を十分に大きく設定する必要が生じるため、高い努力水準を確認したにもかかわらず低い評価を与えようとする誘因が雇い主側に働く。両方を組み合わせることで、初めて、従業員の高い努力水準と企業の誠実な対応を実現することが出来る。

Baker, Gibbons, and Murphy (1994) も客観的評価と主観的評価の間に補完性が働くメカニズムを分析している。Pearce and Stacchetti (1998) との最大の違いは、従業員の行動と客観的業績指標の関係について事前にはわからない不確定要素が存在し、従業員は行動を選択する前に、その不確定要素についての私的情報を得るということである。また、従業員はリスク中立型であることが仮定されている。従業員の行動が客観的業績指標に与える影響を雇い主が観察出来ない場合、客観

的指標にのみ基づいてインセンティブを高めようとする、従業員の行動は企業利益最大化のために求められる行動から乖離する。

例えば、営業担当者は、店舗を訪れる客の数を毎日見ているが、本社の経営陣は観察できないとする。営業担当者が毎月の受注目標の達成率で評価される場合、店舗を訪れる客が多く、月半ばで受注目標を達成した営業社員は、月の残りを遊んで過ごすかもしれない。この時、この社員がどのくらい顧客に電話をかけているか、実際に顧客企業を訪問しているか、顧客は満足しているか、など様々な視点からかれの頑張り具合を評価した上司の主観的指標を組み合わせることが出来れば、上のモラルハザードは解決出来る。

実際、Gibbs *et al.* (2009) は、自動車ディーラーで働く営業マネージャーの報酬制度に関するデータを使って、評価指標と企業利益へのインパクトとの間の乖離が懸念されるほど、そして評価指標の操作が容易なほど、主観的評価が使われる傾向が高いことを示している。

2 インプット指標とアウトプット指標

一般に、複数の業績指標がある時、それらをどう組み合わせるべきかというのは、それらがどの程度企業目的に合致しているか、不確定要素や計測値に含まれるノイズの大きさはどの程度かに依存する (Baker 2000, 2002)。インプット指標とアウトプット指標のどちらを使うか、あるいはどのように組み合わせるのか検討する場合は、それに加えて、プリンシパルつまり経営側が、インプットとアウトプットの関係についてどの程度情報を持っているか、そしてその情報を基にどの程度モニタリング (監視) が可能かに大きく依存する。

通常、インプット指標としてまず想起されるのは、労働時間であろう。ただし、労働時間は努力や人的資源の活用指標としては必ずしも正確ではない。その他使える情報としては、モニタリングの結果として得られた情報が含まれる。例えば、顧客にかけた電話の本数、顧客の訪問回数、レストランにおけるトイレの掃除回数のように特定の作業の実施の確認、あるいは部下の頑張りに対する上司の主観的な評価などである。アウト

プット指標としては、実際の生産高、売上げ、利益など企業利益と直結する業績を指す。通常、アウトプット指標よりもインプット指標の方が不確実性あるいは従業員にとってのリスクは小さい。なぜなら、インプットには従業員の行動が直接反映されるのに対し、アウトプットには従業員の支配の及ばない、例えば、景気、消費者の嗜好、機械の故障、競合の行動、といった様々な要因が影響を与えうるからである。ということは、仮に従業員がリスク回避型で、経営側がインプットとアウトプットの間を完全に把握しているのであれば、インプット指標のみを使った方が効率的である。

例えば、小売店舗がどうすれば売上げを最大化することが出来るか、本社が完全に理解しているのであれば、店舗の社員がやるべき作業項目のチェックリストあるいは陳列すべき商品のリストを作成し、その努力項目がどの程度徹底されているかをモニターし、その結果を基に店舗マネージャーの報酬を定める方が、いたずらに店舗の売上げとリンクしたインセンティブボーナスを払うより、企業利益にはプラスである。しかし、仮に店舗によって顧客のニーズが異なったり、競争環境が異なる時には、顧客や競合に関する情報をより多く持っている店舗マネージャーに、どの商品を調達陳列すべきか、どういうサービスに力を入れるべきかについて、決定の権限を与え、彼らの創意工夫を高めるために、彼らの収入を店舗の売上げや利益とリンクさせた報酬制度を構築することが、望ましい。

つまり、インプットとアウトプットの間について、従業員の方が私的情報を持っているほどアウトプット指標のウエイトを高くすべきであるし、そうした私的情報が少なければインプット指標のウエイトを高めるべきである (Raith 2008)。実証研究で利用される不確実性の指標は、多くの場合従業員が獲得するこうした私的情報とも相関が強い。多くの実証研究で、不確実性とアウトプット指標を利用したインセンティブ契約の強度の間に契約理論から導かれる負の関係ではなく、正の関係がしばしば観測されるのは、不確実性が従業員の持つ私的情報の重要性を高め、かつそれは権限移譲の必要性を高め、アウトプット指標の

ウエイトを高くすることに原因があると説明できる (Prendergast 2002)。

また、後に述べるゲーミングと関連するが、従業員が持つ私的情報は必ずしも経営側に役立つ情報とは限らない。例えば、顧客の購入タイミングについて事前の情報を持っている営業社員は、交渉を通じて、自らに有利なタイミングでの購入へと顧客を誘導する可能性がある。そうしたアウトプット指標が操作される危険性がある時には、インプット指標のウエイトを高めることが望ましい (Baker 2002)。

3 絶対的指標と相対的指標

昇進審査において上位のポストの数があらかじめ決まっている場合は、そこで使われる評価指標は、必然的に相対的なレベルにおいて解釈されるから、相対的指標と考えて差し支えない (Lazear and Rosen 1981)。報酬制度において相対的評価指標が選択される場合、それが客観的な指標か主観的な指標かで理由は全く異なる。

まず契約理論の古典的論文 Holmstrom (1982) で議論されたように、個々の従業員の元の客観的業績指標に、共通の要因によって正の相関関係が生じている場合、最適契約は相対的評価指標を用いるものである。例えば、営業社員の業績を売上げで測る場合、その指標は市場全体の動向、あるいはもっと正確に言うならば、雇用動向、金利、政府の財政政策、世界情勢など、特定製品市場に影響を与えるすべての要因から共通の影響を受ける。これら共通の要因は不確実な収入の変動をもたらすので、絶対指標である売上成績だけを報酬にリンクさせるのは、従業員がリスク回避型である場合、好ましくない。この時、同じ製品を販売する営業社員全員の平均売上げと各自の売上げを比較し、その差に基づき報酬を支払うことで、両方に含まれる共通のリスクを相殺し、収入の変動を抑制することが出来る。

他方、主観的指標において相対評価が用いられるのは、多くの場合、後に述べる中心化傾向によって、社員の業績に関する情報を経営陣が得ることが難しくなっているというのが理由である。この場合、評価ランクそれぞれの割合を強制的に

上から与えることによって、上司に部下のランキングを迫ることが出来る。ただし、職場によって能力や業績に大きなバラつきがある場合、職場ごとのランキング情報を報酬にリンクさせるのは、新たな不確実性をもたらすので、好ましくない。こうした相対評価は、職場組織ごとの能力のバラつきがある程度一様だと思ふ場合を除いては、避けた方が賢明であろう²⁾。

4 集団業績指標と個人業績指標

多くの企業で、集団業績指標と個人業績指標の両方が使われている。例えば、個人の給与あるいは給与の引き上げ率が、彼の人事考課あるいは目標管理制度の下での目標到達度に応じて決められる一方、ボーナスは利潤分配の考えの下、企業全体の単年度利益とリンクしているというケースは多い。上場企業では、持ち株会参加を通じ、企業価値増大の恩恵を受けられる人も多い。集団業績指標と個人業績指標を組み合わせた報酬制度を企業が提供するの、それぞれ長所短所があり、単独では最適契約とはならないからであろう。

個人業績指標の長所は、個人の行動や努力に対する感応度が高く、また集団の指標に比べれば、他人の行動や意思決定の影響を受けにくいので、不確実性やノイズが比較的小さい。その分、強いインセンティブを与えることが出来る。短所は、他人の業績や他部門への影響を考慮する必要が薄

れるので、仮に他人や他部門との協力やコーディネーションが重要でも、そうした望ましい行動を取るよう十分なインセンティブを与えられない。これは、後で議論する組織目標と評価指標とのズレ (distortion) として理解出来る。

他方、集団業績指標の長所は、組織目標と合致しており、また従業員が評価指標を操作することも難しいので、好ましくない行動を誘発する危険性は低いということである。短所は、個人の行動や努力に対する感応度が低く、他人の努力にただ乗り (free-riding) するインセンティブを与えてしまうことや、他人、他部門、競合の行動やマクロ経済の影響を受ける可能性が高くなり、不確実性やノイズが増すことであろう。

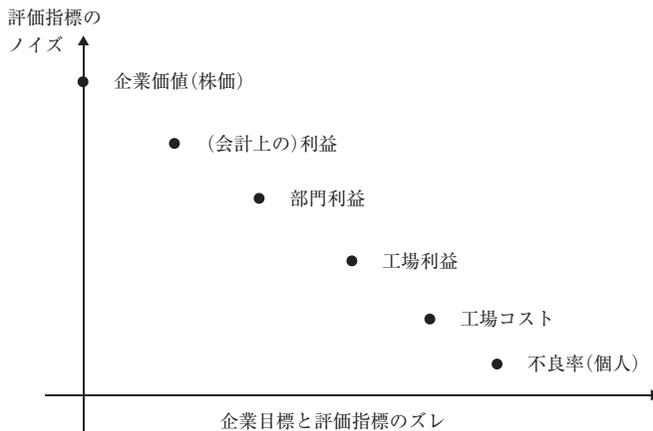
Baker (2002) に、ノイズとズレ (distortion) のトレードオフという視点から分かりやすく上記の関係を描いた図がある。図1に転載するので、参考にされたい。

IV 評価制度設計の基本——マルチタスク問題

1 マルチタスク問題とは

ある程度客観的な業績指標が存在するにもかかわらず、業績給ではなく固定給が支払われるケースが多いのは何故だろうか？ あるいは、より企

図1 集団指標と個人指標



出所：Baker (2002) に筆者が一部加筆。

業目的に近いと見られる業績指標の利用が可能な中で敢えてそれとは異なる業績指標が使われるのは何故であろうか？ かって研究者を悩ましたこうした問題は、今では、Holmstrom and Milgrom (1991) らによって分析されたマルチタスキング・エージェント問題と総称される現象によって説明出来ることが多い（以下略してマルチタスク問題と呼ぶ）。マルチタスク問題を理解することは評価制度を設計する上での最も重要な基礎となるので、ここで詳細に説明を行いたい。

この世に存在する職業のほとんどは複数の業務をこなすことを要求されている。あるいは、極めて限られた範囲の作業であったとしても、成果の次元は複数存在するケースがほとんどである。例えば、多くの生産労働者の成果は、製品を何個完成させたかという量と不良品がいくつあったかという質の2つの次元で測ることが可能である。

営業職の職務には、売るという業務（販売）と顧客のニーズに関する情報を集めるという業務（マーケティング）と顧客の苦情を処理するという業務（クレーム処理）の3つがあるかもしれない。顧客のニーズに関する情報は商品開発部門に伝えられ、将来発売される商品の価値を高める。また顧客の苦情に誠意を持って適切に対処することは、同じ顧客への将来の販売のみならず、市場における同企業の評判を高めることで、将来の利益にプラスに働くであろう。

このように、従業員が複数の業務あるいは複数次元の課題を抱えている場合、企業にとって価値を生み出すすべての業務について評価指標が存在し、かつ評価が正確でノイズやバイアスを含まないのであれば、企業にとって価値を生み出すすべての業務およびすべての次元において、適切なウェイトを置いて評価することが望ましい。Kerr (1975) はかつて、「Bを望みながらAに対して報奨を与える愚かさについて（“On the folly of rewarding for A while hoping for B”）」と題する論文のなかで、企業目的（あるいは企業利益）と業績指標のアラインメント（行動を動機づける方向の合致）が出来ていないために、いかに望ましくない行動が誘因されるか数多くの事例を挙げて説明した。

アラインメントの重要性は、仮に営業職社員に

販売額に基づくコミッションのみを支払った場合を想定すれば明らかだ。こうしたインセンティブ契約の下では、彼らは、販売額を伸ばすことだけに時間と注意を払い、顧客に会ってニーズを聞きだしたり、彼らの不満に耳を傾け満足させるための方策を練るためには、（それが短期的な販売増につながると予想される場合を除いては）ほとんど時間と労力を注がなくなるであろう。これは会社にとっては大きな損失となり得る。

2 評価指標のズレと不確実性が原因

上に述べたような不完全なアラインメントは、必ずしも経営陣の無策が原因とは言えない。そもそも実際の企業価値への貢献と評価指標で計測されるものの間には、多くの場合、必ずズレ（distortion）がある。Baker (2000, 2002) は、このズレを、従業員の様々な行動が企業価値と評価指標に与える限界効果のベクトルのズレとして表現している。つまり、ある行動の組み合わせが企業価値に与える影響と評価指標に与える影響の間に違いがあるために、企業価値を最大化する最適行動と、評価指標を最大化する（従業員にとっての）最適行動が一致しなくなる。加えて、従業員がリスク回避型である場合、評価指標の持つ不確実性や計測誤差も、最適契約によって導かれる努力水準を歪める要因として働く。

例を使って説明しよう。先に挙げた営業職の事例で、簡単のため、売るという業務（販売）と顧客の不満を解消する業務（クレーム処理）の2つしか職務が与えられていないとする。仮に、販売活動に費やす努力とクレーム処理に費やす努力の限界効果を、それぞれ B_S 、 B_C とすると、評価における2つの業務のウェイトをどのように設定したらよいだろうか。少し厳密な議論をするために、販売活動に投入した努力を t_S 、クレーム処理活動に投入した努力を t_C とする。この時、この社員による獲得企業価値 V は、従業員に支払う賃金を w として

$$V = B_S t_S + B_C t_C + \varepsilon - w \quad (1)$$

と表される。 ε は、企業価値を取り巻く不確実性（確率変数）である。

今、インセンティブ契約のミスアラインメント、つまり企業利益最大化行動から実際の行動が乖離する要因を見るために、1つの評価指標が存在する場合と、2つの評価指標が存在する場合、の2つのケースについて最適契約を見てみよう。

① 1つの評価指標だけが利用可能な場合

仮に、営業社員の成功にとって顧客との長期的な関係の下での継続的販売が重要であれば、売上げという1つの業績指標だけでも、ある程度クレーム処理に時間を費やすインセンティブを与えることが出来る。販売成績指標を p とおき、社員の行動選択との関係の以下のように定める。

$$p = g_s t_s + g_c t_c + \eta \tag{2}$$

g_s と g_c は、それぞれ従業員の販売努力およびクレーム処理努力の評価指標に対する限界効果である。 η は、業績結果における不確実性や計測誤差を表す確率変数であり、その分散を σ_η^2 とおく。つまり、これらの業績評価指標は、従業員本人のコントロールの及ばない多くの要因の影響を受ける。 ε と η の間には一般に相関がある。

企業価値と業績指標の間には通常ズレ (distortion) がある。つまり、完全にはアラインされていない。このズレは、業績指標によって導かれる努力の方向の、企業価値最大化に必要な努

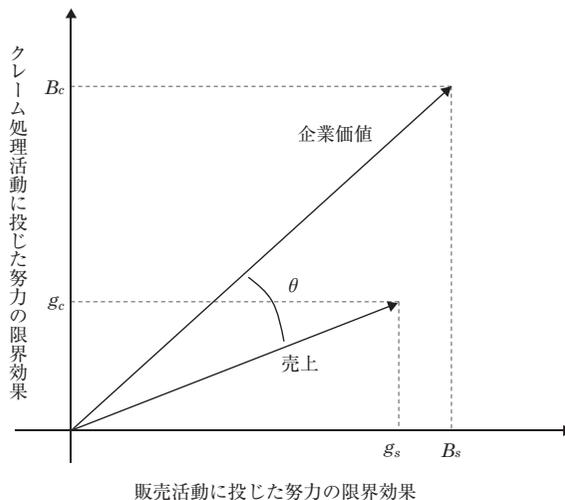
力方向からの乖離によって表すことが出来る。具体的には、努力の限界効果の示す方向を比べることである。式 (1)、(2) から、 V と p に対する努力 t_s 、 t_c の限界効果は、それぞれ、 (B_s, B_c) 、 (g_s, g_c) であり、図2に示すように、 $\frac{B_c}{B_s} \neq \frac{g_c}{g_s}$ である限り、これらのベクトルは一致しない。2つのベクトルの角度 θ は、企業価値と業績指標の間のズレを測る尺度になる。

指数効用関数 $u(w, t_s, t_c) = -e^{-r(w - c(t_s, t_c))}$ を仮定し、努力費用関数を $c(t_s, t_c) = \frac{1}{2}t_s^2 + \frac{1}{2}t_c^2$ と仮定する。これは後に述べるように2つの努力の間に補完性も代替性も存在しないケースである。企業は、従業員に対し、 $w = \alpha + \beta p$ という線形のインセンティブ契約に基づき賃金を支払うとすると、最適契約のインセンティブ強度 β^* は、比較的容易に求められる (Baker 2002)。

$$\beta^* = \frac{\sqrt{B_s^2 + B_c^2} \sqrt{g_s^2 + g_c^2} \cos \theta}{g_s^2 + g_c^2 + r\sigma_\eta^2} \tag{3}$$

θ は、先に述べた企業価値と業績指標のズレである。 $\cos \theta$ は、三角関数であり、 θ がゼロから 90° に増えるにつれ、 $\cos \theta$ は1から0に減少することにご注意頂きたい。 σ_η^2 は η の分散であり、評価指標 p の持つ不確実性 (リスク) の大きさを

図2 企業価値と評価指標の間のズレ



表す。 r は効用関数に含まれるパラメーターであるが、絶対的リスク回避係数と呼ばれ、個人のリスク回避度の大きさを表す。

$\sqrt{B_s^2 + B_c^2}$ は、図2において企業価値に対する社員の努力の限界効果を表すベクトルの長さであり、企業価値の努力に対する感応度である。これをベクトルの長さを表す表示方法を使って $|\mathbf{B}| = \sqrt{B_s^2 + B_c^2}$ とおく。同様に、評価指標の努力に対する感応度を $|\mathbf{G}| = \sqrt{g_s^2 + g_c^2}$ とおくと、(3)式は

$$\beta^* = \frac{|\mathbf{F}| \cos \theta}{|\mathbf{G}| 1 + r(\sigma_\eta / |\mathbf{G}|)^2} \quad (4)$$

と簡素化出来る。

この結果のポイントは3つある。まず、企業価値と評価指標のズレ θ が大きければ大きいほど、 $\cos \theta$ は低下し、最適契約のインセンティブ強度は弱まる。つまり、報酬の評価指標に対する感応度 β^* が下がる。これは、 θ が大きくなるほど、努力配分の歪みのコストが大きくなるので、インセンティブを弱めて効率的努力配分比率からの乖離が大きくなり過ぎないように調整する必要が出てくるからだ。極端な場合、 $\theta = 90^\circ$ であれば、固定給が最適であり、評価指標は役に立たない。ただし、 $90^\circ < \theta < 270^\circ$ であれば、評価指標は望ましくない行動を測っていることになり、インセンティブ強度 β^* はマイナスとなり、企業価値を下げる行動が罰される。

次に、評価指標の不確実性 σ_η^2 が大きくなればなるほど、最適契約のインセンティブ強度は弱まる。これはリスク回避型の従業員にとって不確実性の増大は収入の変動上昇を通じて効用の低下をもたらすので、それを部分的に相殺するため、評価指標と報酬の結びつきを弱める必要が出てくるからである。

3番目に、最適契約は、企業価値と評価指標の間の社員努力を条件とする条件付相関 $Corr(V, p | t_s, t_c) = Corr(\varepsilon, \eta)$ に全く依存しない。これは、企業目標と評価指標の間に観察上大きな相関が見られたとしても、それは評価指標の有効性を示す証拠とは必ずしも言えないことを示している。見掛け上の相関は ε と η の間の相関から来ている可能性があり、実際のズレを決定

するパラメーターである (B_s, B_c) と (g_s, g_c) の間には大きな開きがあるかもしれないからである。

例えば、顧客満足度とその顧客への売上げの間に高い正の相関があるということは、必ずしも顧客満足度が報酬制度で使用する評価指標として適切であることを意味しない。顧客満足度を効果的に押し上げる行動と、売上げを高める行動は全く異なるかもしれないからだ。

② 2つの評価指標が利用可能な場合

仮に、販売活動とクレーム活動に注がれる努力をそれぞれ測る2つの評価指標が利用可能であった場合、上述のミスアラインメントは無くなると期待されるかもしれないがそうではない。例えば、以下の2つの評価指標を想定する。

$$p_s = t_s + \eta_1 \quad (5)$$

$$p_c = t_c + \eta_2 \quad (6)$$

販売成績は、例えば目標売上達成率によって評価し、クレーム処理に関する成績は、報告されたクレームごとに顧客に処理後の満足度を尋ねるメールアンケートを行い、その結果に基づき評価しているとする。これらの業績評価指標は、従業員本人のコントロールの及ばない様々な要因によって定まることは明らかであろう。そうした不確定要素を表す η_1, η_2 はお互いに独立な確率変数であり、その分散を σ_s^2, σ_c^2 とおく。この時、やはり $w = \alpha + \beta_1 p_1 + \beta_2 p_2$ という線形の契約を前提に、最適なインセンティブ強度 β_1^* および β_2^* を求めると、それぞれ

$$\beta_1^* = \frac{B_s}{1 + r\sigma_s^2}, \quad \beta_2^* = \frac{B_c}{1 + r\sigma_c^2} \quad (7)$$

と表せる (Holmstrom and Milgrom 1991)。

この時、仮に $B_s = B_c$ であったとしても、クレーム処理後のアンケート調査の不確実性 σ_c^2 の方が販売業績のそれ σ_s^2 よりも大きいとすると、 β_2^* は β_1^* に比べ低く設定され、販売業績の方により大きなウエイトを置くことになる。したがって、2つの評価指標が存在する場合でも、2つの業務の間の努力配分が望ましい比率よりも大きく乖離する可能性は排除できない。

マルチタスク問題が深刻となり、客観的評価指

標が望ましいインセンティブを与えることが困難になる場合、主観的評価を使うことが効果的であるかもしれない。なぜなら、上司が部下の行動をモニターすることによって、彼女が組織にとって望ましい行動をとっているか直接評価することが可能となるため、従業員本人のコントロールの及ばない様々要因によって生じる不確実性を排除できるからである。ただし、その場合、評価者には正しく評価する十分なインセンティブが与えられている必要がある。

高橋・都留・上原（2011）は、大手自動車販売会社の人事データを用い、実際に、マルチタスク問題が深刻化する可能性が高い職場において、主観的評価と客観的評価が補完的に運用されているか検証した。具体的には、(1) 営業スタッフ一人当たりの新入社員数割合とランクの高い社員の業績考課結果の間に正の相関関係がある、(2) 年間総顧客数に占める法人客の割合と業績考課の間には正の相関関係がある、という2つの仮説を店舗単位のデータで検証している。つまり、新入社員数が多いほど新入社員の教育という追加的な任務が課され、法人客の割合が高いほど法人客のメンテナンスのためのサービス提供という追加的な任務の比重が高まるという仮定の下、主観的評価と客観的評価の間に予想される相関の変化が見られるかを検証した。結果は、(1) については仮説が支持されたが、(2) については有意な結果を得ていない³⁾。

最後に、ここでは単純化のためモデルでは可能性を排除したが、販売のための努力とクレーム処理のための努力の間に補完的あるいは代替的な関係がある場合、インセンティブ強度の設定には更なる注意が必要である。販売のための努力水準を高めることが、クレーム処理のための限界費用を低下させる時、両者の間には補完性があるという。例えば、販売活動に注力することで、顧客とのコミュニケーションが改善したり、顧客のニーズをより正しく理解することが出来るようになるかもしれない。それは、クレームの適切な処理を容易にし、顧客満足を得るために必要な従業員の時間を下げる方向で働くだろう。こうした補完性は、両方の評価指標の強度 β_1^* および β_2^* を押し上

げる方向で働く。

逆に、販売活動に注力すればするほど、クレーム処理に割く時間がなくなり、その効果的解決が難しくなる場合も考えられる。その場合は、クレーム処理に十分な時間を充てるインセンティブを作り出すため、販売活動に対するインセンティブ強度 β_1^* が大きく押し下げられる場合もある。次に述べるケースは、2つの活動の努力の間に代替性があるために、 $\beta_1^* = \beta_2^* = 0$ つまり固定給が選択される場合である。

3 固定給が望ましいケース

多くの職で、客観的な業績指標があるにもかかわらず、固定給が支払われている。マルチタスキング・エージェントモデルは、この現象を説明するのに有効である。例えば、先の販売社員の例で、今度は販売活動とマーケティング活動のみを考え、努力費用は、販売活動とマーケティング活動の努力コストの和、 $t_S + t_M$ によって決まるとする。Holmstrom and Milgrom (1991) に倣い、図3に描かれた次の費用関数を想定する。

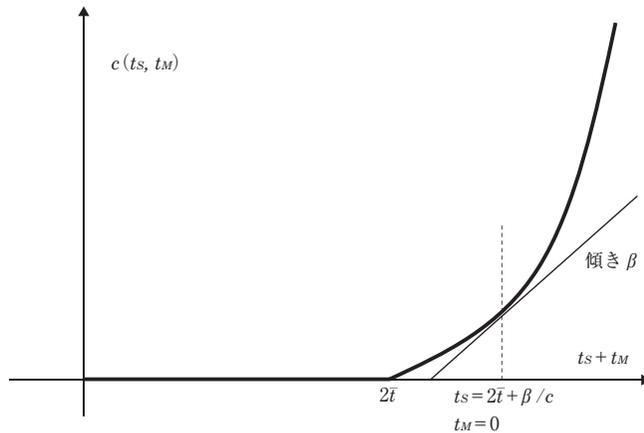
$$c(t_S, t_M) = \begin{cases} \frac{1}{2} c \cdot (t_S + t_M - 2\bar{t})^2 & \text{if } t_S + t_M > 2\bar{t} \\ 0 & \text{if } t_S + t_M \leq 2\bar{t} \end{cases}$$

この費用関数の下では、 $t_S + t_M$ が $2\bar{t}$ に達するまでは、努力コストはゼロであり、何らインセンティブ契約がなくとも、従業員は上司の指示に従い合計 $2\bar{t}$ までの努力配分を行うものとする。

企業利益は $\Pi = B t_S t_C - w$ と仮定する。つまり、マーケティング努力を全くしなければ、顧客のニーズに合った商品・サービスは提供できず、販売活動のレベルに関係なく、間もなく売上げはゼロになるという、マーケティング活動が極めて重要な状況を想定する。この時、マーケティング活動の業績指標が存在せず、販売業績指標のみが存在するとする。マーケティング活動が不可欠であるこうした状況で、販売業績に基づくインセンティブを導入すべきであろうか？ 答えはノーである。

販売業績指標を $p_S = t_S + \varepsilon_S$ 、報酬を $w = \alpha + \beta p_S$ と表す。この報酬制度の下では、従

図3 仮定された努力費用関数（固定給が最適となるケース）



業員は、限界効用＝限界費用となる努力を選択し、 $t_s = 2\bar{t} + \beta/c$ 、 $t_m = 0$ となる。いかに β を小さく取っても、それが正である限り、マーケティング活動より販売活動に時間を使うことが得で、誰も前者に時間を割かなくなる。したがって、期待企業利益は、 $E\Pi = -\alpha - \beta \cdot 2\bar{t} < 0$ 、つまり負となり、この企業は赤字に転落する。

仮に固定給を支払うとどうなるか？ つまり $\beta = 0$ とし、固定給 $w = \alpha$ を支払う。この場合、 $t_s + t_m \leq 2\bar{t}$ である限り努力コストはゼロなので、従業員は会社の指示にしたがって努力の分配を図る。上司が、 $t_s = t_m = \bar{t}$ を指示すれば、企業利益は、 $\Pi = B\bar{t}^2 - \alpha$ となり、 $B\bar{t}^2 > \alpha$ である限り、企業利益は正となる。

この事例は、インセンティブ契約を導入することで、従業員の努力配分が企業の望む方向とは逆に捻じ曲げられ、結果的に企業業績が悪化するメカニズムを端的に示している。

V 評価におけるバイアス

1 えこひいきのモデル

評価におけるバイアスについての経済理論はあまり多くない。よく知られたモデルは、Prendergast and Topel (1996)によるえこひいき (favoritism) のモデルであろう。経営者が評価者である上司と評価される部下の2人のインセ

ンティブ契約を設計する比較的単純なプリンシパル・エージェントモデルである。いくつか特筆すべき特徴がある。まず、上司の効用が、自身の賃金と部下の賃金の2つに依存し、部下の賃金に依存している度合い（部下に対する好意と呼ぶ）はマイナス（差別）でもプラス（えこひいき）でもありえ、平均値0の正規分布が仮定されている。この好意のパラメーターは上司の私的情報であるが、経営者はその分布を知っている。2つ目は、経営者は上司とは別に部下の業績を観察することで上司をモニターし、仮に経営者の観察と上司の評価の間に乖離があれば、その大きさに応じて、上司の賃金は減額される、つまりえこひいきは罰される。どの程度罰されるかは、経営者のインセンティブ設計の一部である。第3に、上司の評価は部下の報酬を決めるだけでなく、その後昇進を決める際にも利用される。

モデルの構造から当然であるが、上司は気に入った部下と働く場合、この部下の評価を甘くつける。それによって生じるバイアスは、部下に対する好意（あるいは悪意）が強いほど、部下の報酬が評価により敏感なほど、そしてえこひいきに対する罰金が小さいほど、大きくなる。こうしたえこひいき（あるいは差別）には2つのコストがある。まず、部下にとっては自らの支配の及ばない賃金の決定要因が増えるので、不確実性が増し、リスク回避型の部下の効用が下がる。経営者はそれを相殺するためにより高い賃金を払う必要

が出てくる。第2に、えこひいきの結果、本来昇進させるべきではない人の昇進が決定され、企業利益を減少させる。それでは、えこひいきの罰金（経営者と上司の部下に対する評価の差によって決まる報酬減額分）を無制限に引き上げれば良いかというそうでもない。上司は、好きな人は甘く評価し、嫌いな人には厳しい評価を課すことである一定の効用を得ている。Prendergast and Topel (1996) はこれを評価権限から上司が得る余剰と呼び、評価権限への欲求を生み出すと説明する。この余剰の分だけ企業は上司に支払う賃金を節約できる。したがって、えこひいきの罰金を上げれば上げるほど、企業はより高い賃金を上司に支払わなければいけない。

また、えこひいきの罰金は、経営者が観測する部下の業績の計測誤差によっても変動する。したがって、上司がリスク回避型であれば、えこひいきの罰金が増えれば増えるほど、その不確実性によって彼の効用は下がり、それを相殺するためさらに企業が上司に支払う賃金は増える。したがって、企業はえこひいきのコストとえこひいきをやめさせるためのコストが釣り合うところで、えこひいきの罰金を定める。

このモデルは3つの重要な結果と知見を提供してくれる。まず、えこひいきが企業利益にマイナスの影響を及ぼすかどうかは、経営者がどれだけ効果的に上司をモニタリング出来るか、つまり経営者が部下の業績を評価し上司の評価にバイアスがかかっていないかチェック出来るかに大きく依存している。かりにこのモニタリングのコストがゼロで、上司がリスク中立的である場合、えこひいきの企業利益に与える影響はゼロである。部下にオファーされるインセンティブ契約も、えこひいきが全く存在していない場合に提案されるものと同一で、えこひいきによってインセンティブが低下するという事もない。非効率性が全く生じないというこの結果は、先に述べた理由からえこひいきの罰金はえこひいきを無くすほど高くは設定されず、上司は部下の評価にある一定のバイアスをかけ続けるという結論と併せ見ると、驚くべきである。この均衡では、えこひいきによって生じる部下の効用低下と、上司の得る評価権限から

の余剰が完全に一致しており、それらを相殺するための賃金の増減も、上司と部下で一致するため、企業利益への影響は中立となる。

2番目の知見は、論文では十分には議論されていないが、上の結果は現実には様々な要因から成り立たず、均衡では部下のインセンティブ契約における上司の評価のウエイトは低下し、インセンティブは弱められる可能性が高い。1つは、経営者によるモニタリングのコストは実際には高く、えこひいきに対する罰金も低く設定されるためである。2つ目に、現実には上司がリスク回避型であることが考慮される可能性が高く、やはりえこひいきに対する罰金と部下のインセンティブ契約における上司の評価のウエイトを下げる方向で働く。3つ目には、モデルではえこひいきは上司の私的情報であり、部下はそれを観測できない（あるいは今のモデルは静学的なモデルなので、事後的に知っても結果に影響を与えない）。仮に、部下が上司の差別や同僚へのえこひいきを知ることで、会社を辞めたり、将来の勤労意欲が下がるのであれば、企業はえこひいきの弊害を避けるため、追加的な措置を迫られる。こうした他の要因が強まれば、企業は上司の評価は出来るだけ報酬には反映させず、昇進などの意思決定にのみ利用することになる。

3番目の知見は、仮に上司が、部下の業績について経営者が得る情報やその予想についてある程度事前に知ることが出来れば、自らの情報を経営者の情報に近づけることで、えこひいきの罰金を下げようとするだろう。これは、情報の圧縮を意味し、いわゆる評価の中心化傾向につながる可能性が高い。上司の持つ情報が十分に活用されない結果、部下のインセンティブは弱められ、また効率的な昇進を妨げることに繋がる。こうした評価の中心化傾向が、報酬制度にどのような影響をもたらすか明らかではない。まず、評価が圧縮されることによってインセンティブが弱まることを避けるため、報酬と上司の評価の関係を逆に強めることが望ましいという直接的なインパクトがある。他方、中心化傾向をなくすため、えこひいきの罰金を減らす必要があり、その結果バイアスが増すため、報酬の上司の評価への依存度を下げる

必要が出てくる。これは間接的なインパクトである。どちらが強くなるかによって、報酬制度への影響は変わってくる。

2 評価に見られるバイアス

評価にバイアスが存在するという事実は多くの心理学や経営学の文献によって指摘されてきたが、実験に基づくものやケーススタディに基づく研究が多く、ビジネスの現場における定量的な証拠はあまりない。また、多くの研究は人種問題に焦点を合わせており、学歴、経験などより広い個人特性の違いがどのようなバイアスを引き起こしているかという研究は極めて限定的だ。とはいえ、主観的評価にバイアスが生じるという事実を定量的に計測した論文が増えてきたことは、バイアスの潜在的な大きさを測り、それを補正する政策を検討する上で、基本的なデータを提供することになる。

Elvira and Town (2001) は、ある米国大企業の1部門の販売社員を対象に彼らの人種を含む個人属性と上司の人種が人事評価結果にどのような影響を与えるか分析を行った。販売社員の評価の主要な評価基準は目標販売額の達成度であったことから、この客観的業績指標もコントロール変数に加えて推定を試みた。アジア系やヒスパニック系社員の比率は著しく低かったことから、これらマイノリティの社員は除いて、白人と黒人社員からなるデータで検証を行っている。

結果は、白人の上司と黒人の部下、あるいは黒人の上司と白人の部下という異なる人種の組み合わせの際に、評価が有意に下がることを確認している。例えば、典型的な社員の場合、白人と白人の組み合わせでは最上位の“Outstanding”がつけられる確率は13.0%あるのに対し、白人上司の黒人部下の場合は4.3%、黒人上司の白人部下の場合は3.5%まで低下する。こうした結果は、上司と部下の人種は、採用、自発的離職、解雇、昇進いずれにおいても、同じ人種間の方が相手にとり都合の良い結果をもたらす傾向(own-race bias)が幅広く観測されたというGiuliano, Levine, and Leonard (2009, 2011)らの研究と整合的である。興味深いのは、勤続年数、年齢、婚姻状況も評価

結果に有意な影響を与えており(例えば独身者に対するバイアスは負)、上司の部下の経済状況への配慮や上司部下の人間関係が評価結果に影響を与えていることがうかがわれる。

Elvira and Town (2001)のように現実の企業の人事評価結果を使った研究はむしろ稀で、多くの研究は、客観的なパフォーマンスと審判による主観的な評価の組み合わせが比較的容易に利用できるスポーツからのデータを使っている。例えば、Price and Wolfers (2010)の場合は、米国プロバスケットボールにおいて、選手と審判の人種の一致がファウルの頻度にどのような影響を与えるか、Parsons et al. (2011)は米国大リーグ野球の判定データを基に、投手と審判の人種の一致が投球判定にどのような影響を加えているか、について分析を行っている。いずれの研究でも同じ人種の選手により有利な判定を下すバイアス(own-race bias)が有意に検出されているが、Parsons et al. (2011)は更に興味深い研究を行っている。監視の目が厳しくバイアスの心理的コストが高い状況ではバイアスが消失すること、及びこうしたバイアスが有利に働く同じ人種の投手は、より自由な裁量が働きやすいストライクゾーンの部分への投球や変化球を投げる確率が高まることを示している⁴⁾。

ビジネス組織においてもバイアスが生じ、バイアスによって不利な扱いを受けている人がそれを認識出来るのであれば、彼らの行動はより業績が見えやすく、主観的な判断が効きにくい業務により多くの努力を注力する傾向が出てくるかもしれない。

3 中心化傾向と報酬制度との関係

MacLeod (2003)やPrendergast and Topel (1996)によって示唆されている中心化傾向、および上司の評価を報酬制度に使わないという傾向は、多くの経営学の文献によって指摘されてきた。しかし、中心化傾向が、MacLeod (2003)のモデルにあるように上司と部下の衝突コストを下げる意図から生じているのか、Prendergast and Topel (1996)にあるように、えこひいきに対する罰則への反応として生じているのか、梅崎・中嶋

(2005)らの主張するように評価者負担によって生じているのか、あるいはJacob and Lefgren (2008)が示したように、分布の中間に属する多くの部下達の業績の違いを区別する能力が上司に欠けているために生じているのか、その形成のメカニズムを明らかにした実証研究はまだない。

また、Prendergast and Topel (1996)は、えこひいきによるバイアスの弊害が大きい時、企業は部下の報酬と上司の評価の間の関係を弱めることで、よりバイアスの小さい評価を提出させ、より効率的な昇進・職への配置を図ろうとすると議論している。しかしながら、上司がある部下に対し好意を持っている場合、なぜ甘い評価によって部下の昇進の確率を高めようとしないのか、明確な議論はない。1つの説明は、報酬は本人が口にしない限り他人には漏れないが、昇進は誰の目にも明らかになるので、えこひいきに対する社会的な圧力がかかりやすいということであろう。いずれもより包括的な実証研究が必要な分野と言える。

VI 評価指標の操作（ゲーミング）

企業目的と評価指標のズレ（distortion）は、報酬制度とリンクして利用されるまでは目につかないことが多い。それは、ズレが従業員の特定の行動によって発生し、その行動を取らない限り、評価指標が企業目的の優れた代理変数となっている場合が多いからである。報酬最大化を目指す従業員によって引き起こされる評価指標のズレをもたらす操作（manipulation）はしばしばゲーミング（gaming）と呼ばれる。

多くの営業社員や企業経営者らは、ある一定期間の売上げや利益に基づき報酬を支払われるインセンティブ契約のもとで働いている。こうした業績給やボーナスが支払われないケースでも、営業ノルマを達成出来なければ、昇進・昇格・異動において不利な待遇を受ける場合、やはり一種のインセンティブ契約と見ることが出来る。この時、報酬のプロファイルが線形ではなく、凸型あるいはターゲットボーナスのようにある閾値を越えた時に報酬がジャンプするような仕組みを持つ場合、指標を操作するインセンティブが高まる

(Oyer 1998)。

例えば、多くの米国企業の経営者達にオファーされるボーナス契約の多くが、支払額のキャップ（上限）やフロア（下限）を持っている（Healy 1985; Holthausen, Larcker, and Sloan 1995）。この場合、利益がキャップを超える好況の会計期やフロアに達しない不況の会計期には、利益を先送りするインセンティブが高まる。同様に、コミッション契約でありながら最低賃金が保証されている営業社員の場合にも、最低賃金以上に稼げない悪い評価期間から最低賃金を上回るコミッションが支払われる良い評価期間へ取引日を動かす操作が生じる（Owan and Tsuru 2011）。また、売上げに応じてコミッション率が上昇して行くインセンティブ契約のもとでは、出来るだけ高いコミッション率が期待できる評価期間へ取引を集中させることが従業員にとって得である（Larkin 2007; Owan, Tsuru, and Uehara 2010）。

会計利益、売上げといった本来企業価値と密接に相関しているはずの指標を使っても、従業員が取引計上日を操作出来れば、組織にとって望ましい行動と誘導される行動の間に乖離が生じる。こうした行動のズレは、本来の目的に合致していないばかりでなく、本来の目的とは逆に企業価値を下げる方向で働く。取引計上日を操作するゲーミングのコストは、3つある。まず取引日を移動させる最も「安全な」方法は顧客にこちらにとって望ましい取引日を選んでもらうことであるが、そのために値引きが使われることが多い。例えば、購入を急いでいない顧客に評価期間内に発注してもらうためには（pull-in 操作と言われる）、その意思決定が顧客にとって有利であるよう、値段の引き下げが必要となるが、それは会社にとって利益機会の逸失である。第2に、ゲーミングが成功した結果、会社は従業員に余計に報酬を支払わなければいけないため、人件費の増大を招く。最後に、取引日の操作は、評価期間の最終日あるいは最終週に受注が集中する傾向を生み出すため（Oyer 1998）、生産設備の稼働率が大きく変動したり、経営トップの利益予想が外れて市場の信頼を失うと言った弊害が生じる。

こうした多くの弊害が指摘されるにもかかわらず

ず、非線形報酬は依然として広く使われている。理論的・実証的裏づけのない仮説であるが、その理由として、非線形報酬が通常、目標設定、その達成、達成に対する報酬というフィードバック効果を持っていることに原因があるのかもしれない。達成感のない様なインセンティブよりも、達成すべきスタンダードが設定され、それを越えた時に報酬が支払われる、あるいは不連続に報酬が増える構造の方が、モチベーションを与える上で優れているかもしれない。ちなみに、Owan and Tsuru (2011) らが分析した日本の自動車ディーラーでは、業績給決定のために使用する指標として、毎月の獲得粗利益の代わりに、その2カ月移動平均を使うことで、車両登録日の操作がほぼ消失した。しかし、その代わり、営業社員の生産性も有意に低下した。上述の仮説と整合的な結果である。

上に挙げたゲーミングは、営利組織だけの問題ではない。むしろ組織の目的が曖昧、あるいは複数のプリンシパルあるいは複数の目標が設定される傾向がある政府組織の方がゲーミングのリスクは高い。例えば、旧社会保険庁において、保険料納付率向上を目標として課したことにより、多くの事務所が、保険料の不正免除手続きを進め、分母の縮小による納付率引き下げに動いた。保険料収入最大化という組織の本来の目標のために、それに最も近い評価基準である納付率が選択されたにもかかわらず、不正手続きという行動が可能だったために、目的と評価指標の間にズレが生じた好例であろう⁵⁾。

Ⅶ 最後

本稿の中で、主観的指標と客観的指標の違いを始め、様々な評価指標の利点・欠点、そして報酬と業績評価をリンクさせることによって生じる問題など、評価制度を設計する上で知っておくべき知見を整理した。世に出ている多くのケーススタディや社会人学生との会話に基づき判断すると、日米の企業を比較した際、日本企業の方が米国企業よりも客観的指標やアウトプット指標を避け、主観的指標やインプット指標に頼る傾向があるよ

うに感じる。その理由として以下の5点が考えられる。

まず、人事制度が中央で管理され、同一の制度をすべての職能、職階に当てはめようとする傾向があるため、比較可能な評価基準を事前に明確にして運用しなければいけない客観的指標よりも、柔軟に基準を差別化できる主観的指標の方が運用は楽である。2つ目に、日本企業では、ブルーカラーもホワイトカラーも職務が曖昧で能力のある人には職域が広く設定される傾向がある。このため、マルチタスク問題が顕在化し易く、それへの対処が難しい客観的指標やアウトプット指標は敬遠される傾向がある。

3つ目に、長期雇用慣行の下、評判のメカニズムが機能しやすく、履行が法的に保証されない主観的指標に基づく暗黙の契約が破られにくいのかかもしれない。4つ目に、長期雇用慣行の下で定期的に人事異動が行われるため、主観的指標で問題となる評価のバイアスが複数評価者の平均を取るうち正されていくというメリットがあるのではないか。最後に、長期雇用慣行の下で、昇進・昇格・異動を通じたインセンティブの役割が報酬を通じたインセンティブよりも重要であり、前者の目的のために使われる場合に比較的弊害が少ない主観的評価指標がより重宝される傾向がある。

以上の見方が概ね的を射ているのであれば、日本企業の人事管理システムと補完性のある評価システムは何かという議論もある程度意味があるものかもしれない。成果主義の下で、多くの企業が経験した主観的指標の弊害やマルチタスク問題の本質について、より多くのデータを集め理解を深め、実証分析に裏づけられた知見を提供していくことが研究者や人事の専門家に求められている。

- 1) エージェントがリスク回避型である場合、不確定要素の含まれる指標を使ったインセンティブ契約は取入リスクの増大を通じてエージェントの効用を下げる。そのため、企業は指標と報酬の関係を弱める必要が生じ、効率的な努力水準を達成できない。この結果は、インセンティブと保険のトレードオフとして、契約理論で広く知られた結果である。Milgrom and Roberts (1992) または大湾 (2008) を参照されたい。
- 2) 主観的業績評価に相対評価を導入する場合に生じる問題については、大湾 (2009) が詳細にまとめている。
- 3) (1) (2) の仮説が妥当かについては、いささか疑問が残る。追加的な任務があれば、それに対するウエイトが高まるとい

うことであって、単純にそれが主観的評価を押し上げるということではないのではないか。検証すべき仮説は、(1) 営業スタッフ一人当たりの新入社員数割合が高いほど、ランクの高い社員の販売成績と業績考課結果の間に正の相関が弱まる、(2) 年間総顧客数に占める法人客の割合が高いほど、販売成績と業績考課結果の間の正の相関が弱まる、ではないかと思われる。

- 4) Parsons *et al.* (2011) については、本誌9月号の論文 Today において、高橋主光氏が解説を加えているので、興味のある方は参照されたい。
- 5) 米国では、障害者の雇用機会拡大を狙って、ジョブ・トレーニング・パートナーシップ法の下、各地にトレーニングセンターがつけられた。センターのパフォーマンスを評価するために、就職率やトレーニング後の賃金を基に評価基準が作成されたが、これが障害の程度が低い者ばかり受け入れる操作につながったのではないかという議論がある (Heckman, Heinrich, and Smith 1997)。

参考文献

- Baker, George (2000) "The Use of Performance Measures in Incentive Contracting" *American Economic Review* 90(2) 415-420.
- (2002) "Distortion and Risk in Optimal Incentive Contracts." *Journal of Human Resources* 37(4) 728-751.
- Bull, Clive (1987) "The Existence of Self-Enforcing Implicit Contracts." *Quarterly Journal of Economics*, 102(1): 147-159.
- Elvira, Marta, and Town, Robert (2001) "The Effects of Race and Worker Productivity on Performance Evaluation." *Industrial Relations*, 40(4): 571-590.
- Gibbs, Michael J., Merchant, Kenneth A., Van Der Stede, Wim A., Vargus, Mark E. (2009) "Performance Measure Properties and Incentive System Design" *Industrial Relations*, 48(2): 237-264.
- Ghosh, Suman, and Waldman, Michael (2010) "Standard Promotion Practices versus Up-or-Out Contract" *RAND Journal of Economics*, 41(2): 301-325.
- Healy, P. (1985) "The Effect of Bonus Schemes on Accounting Decisions." *Journal of Accounting and Economics*, 7, 85-107.
- Heckman, James, Heinrich, Carolyn, and Smith, Jeffery (1997) "Assessing the Performance of Performance Standards in Public Bureaucracies" *American Economic Review*, 87(2): 389-395.
- Holmstrom, Bengt (1982) "Moral Hazard in Teams." *Bell Journal of Economics*, Vol.13, pp.324-340.
- Holmstrom, Bengt and Paul Milgrom (1991) "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership and Job Design" *Journal of Law, Economics and Organization* 7, special issue, 24-52.
- Holthausen, R. W., D. F. Larcker, and R. G. Sloan (1995) "Annual Bonus Schemes and the Manipulation of Earnings." *Journal of Accounting and Economics*, 19, 29-74.
- Jacob, Brian A. and Lefgren, Lars (2008) "Can Principals Identify Effective Teachers? Evidence on Subjective Performance Evaluation in Education" *Journal of Labor Economics*, 26(1): 101-136.
- Kerr, Steven (1975) "On the Folly of Rewarding A While Hoping for B" *Academy of Management Journal*, 18(4): 769-83.
- Lazear, Edward P., and Rosen, Sherwin (1981) "Rank-Order Tournaments as Optimal Labor Contracts" *Journal of Political Economy*, 89(5): 841-64
- Levin, Jonathan (2003) "Relational Incentive Contracts" *American Economic Review*, 93(3): 835-57.
- MacLeod, W. Bentley (2003) "Optimal Contracting with Subjective Evaluation" *American Economic Review*, 93(1): 216-40.
- MacLeod, W. Bentley and Malcomson, James M (1989) "Implicit Contracts, Incentive Compatibility, and InVoluntary Unemployment" *Econometrica*, 57(2): 447-80
- Milgrom, Paul and John Roberts (1992) *Economics, Organization & Management*, Prentice-Hall, New Jersey (ポール・ミルグロム, ジョン・ロバーツ『組織の経済学』NTT出版, 1997年).
- Owan, Hideo, Tsuru, Tsuyoshi, and Uehara, Katsuhito (2010) "Incentives and Gaming in a Nonlinear Compensation Scheme: Evidence from North American Auto Dealership Transaction Data" manuscript.
- Owan, Hideo, and Tsuru, Tsuyoshi (2011) "Integrating High-Powered Performance Pay into a Seniority Wage System" manuscript.
- Oyer, P. (1998) "Fiscal Year Ends and Nonlinear Incentive Contracts: the Effect on Business Seasonality." *Quarterly Journal of Economics*, 113, 149-185.
- Pearce, David G. and Stacchetti, Ennio (1998) "The Interaction of Implicit and Explicit Contracts in Repeated Agency" *Games and Economic Behavior*, 23(1): 75-96.
- Prendergast, Canice (1993) "The Role of Promotion in Inducing Specific Human Capital Acquisition." *Quarterly Journal of Economics*, 108, 523-534.
- Prendergast, Canice, and Topel, Robert H. (1996) "Favoritism in Organizations" *Journal of Political Economy*, 104(5): 958-978.
- Schnedler, Wendelin (2008) "When Is It Foolish to Reward for A While Benefiting from B?" *Journal of Labor Economics*, 26(4): 595-619.
- Waldman, Michael (1984) "Job Assignments, Signaling, and Efficiency" *RAND Journal of Economics*, Vol.15: 255-267.
- 梅崎修・中嶋哲夫 (2005) 「評価者負担が評価行動に与える影響」『日本労働研究雑誌』No.545: 40-50.
- 大湾秀雄 (2008) 「組織と人事制度の経済学第3回——契約を通じて企業組織を見る」『経済セミナー』No.638: 98-103.
- (2009) 「日米のビジネススクール事例で学ぶ人事評価の落とし穴」『労政時報』, No.3742.
- 高橋信吾・都留康・上原克仁 (2011) 「主観的業績評価の理論と現実」『経済研究』62(4).

おおわん・ひでお 東京大学社会科学研究所教授。最近の主な著書に、"Market Characteristics, Intra-Firm Coordination, and the Choice of Human Resource Management Systems: Theory and Evidence," with Takao Kato, *Journal of Economic Behavior and Organization*, forthcoming in 2011. 人事制度と組織の経済学および産業組織論専攻。