



長期的インセンティブにおける主観評価の役割

小笠原, 亨

(Degree)

博士 (経営学)

(Date of Degree)

2018-03-25

(Date of Publication)

2019-03-01

(Resource Type)

doctoral thesis

(Report Number)

甲第7105号

(URL)

<https://hdl.handle.net/20.500.14094/D1007105>

※ 当コンテンツは神戸大学の学術成果です。無断複製・不正使用等を禁じます。著作権法で認められている範囲内で、適切にご利用ください。



博士論文

長期的インセンティブにおける
主観評価の役割

2018年1月15日

神戸大学大学院経営学研究科

三矢 裕 研究室

経営学専攻

学籍番号 134B006B

氏名 小笠原 亨

目次

1	はじめに	1
1.1	本研究の重要性	1
1.2	本研究の概要	3
1.3	本研究の意義	4
1.4	各章の構成	5
2	先行研究	6
2.1	長期的インセンティブの重要性	6
2.2	業績評価と長期的インセンティブに関する経験的研究	10
2.3	長期的インセンティブにおける主観評価の役割	12
3	仮説の導出	14
3.1	過去の情報との関連	14
3.2	主観評価の識別力	16
4	研究方法	19
4.1	リサーチ・サイト	19
4.2	変数の説明	20
4.3	分析モデル	25
5	分析結果	29
5.1	サンプル選択	29
5.2	記述統計量	29
5.3	過去の情報との関連	33
5.4	主観評価の識別力	35
6	追加分析	46

6.1	追加分析の必要性	46
6.2	分析方法	48
6.3	分析結果	49
7	結論	52
7.1	分析結果の要約	52
7.2	本研究の貢献	53
7.3	本研究の限界	54

1 はじめに

はじめに、本章では本研究の重要性、概要、意義および各章の構成を提示する。まず、本研究の重要性について、関連する先行研究を踏まえたうえで、本研究の学術的立ち位置について明示する。次に、本研究の概要では、分析方法および分析結果について記述する。さらに、本研究の意義では、本研究の分析結果がもたらす学術的貢献について述べる。最後に、本論文における各章の構成について記述する。

1.1 本研究の重要性

経営者や従業員など企業に所属する人々が企業目的に適合した行動をとるように、適切な業績評価とインセンティブを組み合わせることは企業にとって重要な課題である (Jensen and Meckling 1976; Holmstrom 1979; Banker and Datar 1989; Feltham and Xie 1994; Lambert 2001)。適切なインセンティブ・システムの設計という課題について、これまで理論モデルをもちいた研究 (see Lambert 2001; Lazear and Oyer 2013)、実験室実験による研究 (Sprinkle and Williamson 2007) などの研究方法をもちいた多くの研究が行われてきた。しかし、理論モデルや実験室実験といった方法は、あくまで仮想的な環境下での分析である。そのため、これらの方法で検証された理論が、現実の経営実践を十分に説明できるかについては、実際の企業のデータをもちいた検証が必要となる。ただし、業績指標・業績評価および報酬に関するデータは企業の内部情報であるために一般に入手が困難である。そのため、実際の企業データをもちいた業績評価に関する研究は少ない (Moers 2007; Lazear and Oyer 2013)。このようなデータ入手の困難さを乗り越えたいくつかの管理会計研究は、これまで財務指標 (Banker1996)、非財務指標 (Banker2001)、主観評価 (Ittner et al. 2003; Gibbs et al. 2004)、相対的業績評価 (Matsumura and Shin 2006) といった様々な業績指標および業績評価に注目し、実際の企業データをもちいた経験的な検証を行ってきた。これらの一連の経験的研究は業績指標・業績評価とボーナス契約との関係に注目している。一方で、企業はボーナス契約など支払いが一度きりの

短期的インセンティブだけでなく、昇進・昇格・昇給といった支払いが複数期間におよぶ長期的インセンティブも取り入れている。これらの長期的インセンティブは、(1) 短期的インセンティブを代替する (Gibbons and Murphy 1992; Ederhof 2010)、(2) 業績評価の難しいホワイトカラーに対するインセンティブ (Baik et al. 2015) などの役割を果たしている。このような役割をもつ長期的インセンティブと業績指標・業績評価の関係を明らかにすることは、実務・研究の双方において重要な課題であろう。しかし、業績指標・業績評価と長期的インセンティブの関係をとり扱った経験的研究は、ボーナス契約を取り扱っている研究と比べてもさらに少ない。長期的インセンティブと業績評価の関係をとり扱った数少ない研究として Cichello et al.(2009) と Campbell(2008) があげられる。Chichello et al.(2009) は、財務指標と昇進・離職の関係について、Campbell(2008) は非財務指標とインセンティブとしての昇進の関係について、いずれも有用な知見を提供している。Campbell(2008) や Cichello et al.(2009) の研究は財務指標や非財務指標といった客観的な業績指標と長期的インセンティブの関係をとり扱っている。一方、現実の企業では昇進・昇格・昇給といった長期的インセンティブは、従業員の能力や成果に対する上司の査定といった主観的な評価により決定されることが多い。しかし、長期的インセンティブと主観評価の関係については、これまで十分に明らかにされてこなかった。長期的インセンティブではないものの、職務配置としての昇進に注目した経験的な研究として、Grabner and Moers (2013) があげられる。Grabner and Moers(2013) では、オランダにおける銀行一社を対象に分析を行い、法人部門から個人部門へ昇進する場合など職務内容の変更が大きい昇進については、客観評価よりも主観評価と昇進が強く相関していた。この結果は、主観評価が現在の職務での成果（客観指標）では測定できない従業員の能力を主観評価が評価していることを示している。しかし、Grabner and Moers(2013) の結果だけでは、昇格や昇給といった必ずしも職務配置の変更を伴わない長期的インセンティブに主観評価が採用されるのはなぜかという疑問には答えられない。この疑問に答えることできれば、主観評価が企業において果たす役割の一つを明らかにでき、一連の主観評価に関する研究 (ex. Bol 2008) に貢献があるだけでなく、財務指標や非財務指標の限界を明らかにすることにもつながる。本研究では、主観評価が長期的インセンティブに果たす

役割に関する一つの答えを提示するため、日本の上場企業一社における人事データをもちいて統計分析を行い、その経験的な証拠を提示することを目的とする。

1.2 本研究の概要

本研究の目的は、主観評価が長期的インセンティブに果たす役割について、実際の企業データをもちいて、経験的に明らかにすることにある。昇進や昇給といった長期的インセンティブが、ボーナスなどの短期的インセンティブと異なる点は、報酬の支払いが複数期間に渡ることである。支払いが一回きりのボーナスとは違い、昇格や昇給は一度決定されれば降格や減給は難しくなる。また、昇進については管理職への就任など職務配置の変更を伴うため、誰を昇進させるかという意思決定は企業の業績に大きな影響を与える (Bertrand and Schoar 2003)。こうした特徴をもつ長期的インセンティブでは、手持ちの情報が少ないゆえに誤って能力の低い従業員に長期的インセンティブを与えてしまったときのコストが短期的インセンティブと比べてはるかに高い。それゆえ、長期的インセンティブの決定においては、これまでに評価者が獲得した情報を総動員して能力の高い従業員と能力の低い従業員を区別することが重要になる。この従業員を識別するという役割に関して、主観評価が客観評価よりも長期的インセンティブに適していると、本研究では主張する。その理由は、主観評価が客観評価とは異なる次のような特徴をもつと考えるからである。すなわち、(1) これまでに評価者が獲得した過去の情報を利用できる、(2) 情報を蓄積することで高い能力をもつ従業員と低い能力しかもたない従業員を識別できるようになる、という二つの特徴である。本研究では、この二つの主観評価の特徴に注目し、仮説の検証を行う。

本研究では、日本の IT 業界に属する上場企業一社の人事データをもちいて、上述した主観評価の二つの特徴を明らかにする。主観評価に過去の情報が利用されているという前者の仮説については、主観評価を被説明変数とする線形モデルにより、仮説の検証を行う。主観評価に情報が蓄積されることで、主観評価の識別力が高まるという後者の仮説については、能力の高い従業員と低い従業員の主観評価を時系列にプロットし、時間が経つことで主観評価がこれらの従業員を区別できるようになる過程を可視化する。その後、統計的

検定を行う。本研究の分析結果は、次のとおりである。まず、主観評価は従業員の能力について情報をもつ前期の成果および前期の昇進と相関していた。この結果は、従業員の固定効果、当期の成果、および年度の影響などの変数をコントロールしたうえでも観測された。次に、時間が経過することで能力の高い従業員グループと能力に低い従業員グループの差は増加していき、かつ上位グループ内の主観評価に関する標準偏差は低下していった。さらに、本研究では、以下の二つの効果を合わせた効果を識別力の向上と定義し、統計的検証を行った。二つの効果とは、(1) 主観評価の値が従業員の能力を反映する値へと移動していく効果、(2) 主観評価の誤差が低下していく効果、である。本研究では、この識別力の向上について統計的検定を行ったところ、時間が経過することで、主観評価の識別力は高まるという仮説が支持された。一方で、客観評価については、時間経過により識別力が高まるという傾向は見られなかった。これらの結果が認知バイアスやえこひいきの影響を過度に受けていないことを示すため、主観評価と将来成果の関係についても分析をおこなった。この結果、分析の前半期間に受けた主観評価の平均値が高いほど、後半期間に挙げる成果が高くなるという傾向が見られた。これらの分析結果は、(1) 評価者が過去の情報を取り入れること、(2) 過去の情報を取り入れることで、主観評価の識別力が時間とともに向上していくこと、(3) これらの特徴をもつ主観評価は将来成果を平均的に予測する程度には適切なものである、と解釈できる。これらの分析結果は、時間が経過することで主観評価の識別力が高まるという本研究の主張を支持している。本研究はこれらの分析結果をもって、なぜ昇給・昇格・昇進といった支払いが長期に渡るインセンティブの決定に主観評価が利用されるのかという疑問に対して一つの答えを提示する。すなわち、長期的インセンティブは支払いが長期に渡るため能力の低い従業員に報酬を与えてしまうことによるコストが高いため、能力の高い従業員と低い従業員を区別することが重要であり、この能力を識別するという観点で主観評価が客観評価に比べて優れているのである。

1.3 本研究の意義

本研究の意義は、以下の三点である。まず、本研究は主観評価に関する一連の研究群 (see Bol 2008) に貢献がある。本研究では、過去の情報を蓄積することで、時間経過によ

り従業員の能力に対する識別力が高まるという主観評価の新たな役割を、実際の企業データをもちいた経験的な証拠をもって示した。本研究で明らかにした主観評価の役割は、先行研究では明示されていない主観評価の側面である。次に、客観評価と主観評価の性質について、時系列推移に注目することにより、その違いの一つを明らかにした。現実の業績評価では、評価が長期に渡って何回も行われる。こうした業績評価の時間的特徴は、業績評価システムの運用方法にも影響を与える (ex. Holmstrom 1999)。しかし、時間的制約の強い実験室実験、もしくはパネルデータを入手することが困難な質問票調査といった研究方法では、こうした数年に渡る業績評価の長期的な変化を捉えることは難しい。本研究では、アーカイバル分析という方法を採用したことにより、客観評価と主観評価の長期的な変動パターンが異なることを示すことができた。最後に、本研究は長期的インセンティブに関する一連の研究に貢献がある。これまでの業績評価と長期的インセンティブに関する研究では、職務配置を伴わない長期的インセンティブに、なぜ主観評価が使われるのかという疑問に答えることができなかった。この点について、本研究では過去の情報を取り入れることで、時間が経つことで従業員の能力を識別しやすくなるという主観評価の性質に注目することで一つの答えを提示する。すなわち、長期的インセンティブでは支払いが長期に渡るため、高い能力をもつ従業員が誰であるかの判断が重要であり、時間経過により識別力の高まる主観評価が客観評価と比べて有用なのである。

1.4 各章の構成

各章の構成は以下の通りである。まず、第 2 章にて先行研究の整理を行う。第 3 章では、仮説の導出を行う。第 4 章ではリサーチ・サイトとなった企業における評価制度の説明および仮説の検証方法について記述し、第 5 章では仮説の検証結果を記す。第 6 章では、前章での分析結果が認知バイアスやえこひいきといった効果の影響を強く受けていないことを示すために追加分析を行う。最後に、第 6 章では本研究の要約、貢献および限界について記述する。

2 先行研究

本章では、関連文献についてレビューを行い、本研究の学術的立ち位置および意義について明らかにする。まず、本研究が長期的インセンティブについて注目した背景について説明する。次に、業績評価と長期的インセンティブの関係について取り扱った先行研究をレビュー・整理することで、本研究の学術的立ち位置を明示する。最後に、主観評価と長期的インセンティブに関わる文献を整理することで、本研究で明らかにしたい課題を明示する。

2.1 長期的インセンティブの重要性

経営者や従業員など企業に所属する人々が企業目的に適合した行動をとるように、適切な業績評価とインセンティブの組み合わせを設定することは企業にとって重要な課題である (Jensen and Meckling 1976; Lambert 2001)。特に、企業で働く従業員に対して適切な業績指標・業績評価およびインセンティブを設定することは、マネジメント・コントロールの基礎でもあり、経営戦略を実現するうえで不可欠な仕組みである (Anthony and Govindarajan 2007)。このような従業員に対する業績評価とインセンティブの設計という問題に対して、理論モデルをもちいた研究 (ex. Banker and Datar 1989; Feltham and Xie 1994)、実験室実験による研究 (ex. Chow et al 1983; Frederickson 1992) といった方法論をもちいて多くの研究が行われてきた (see Lambert 2001; Sprinkle and Williamson 2007)。しかし、これらの研究蓄積により明らかになった理論が現実の経営実践を説明できるか検証するためには、実際の企業のデータをもちいた研究も必要となる。理論モデルや実験室実験は理論検証のために様々な条件が統制されている、いわば仮想的な環境下での検証である。しかし、これらの仮想的な環境のもとで支持された理論が、現実の企業でも観測される程に十分な説明力があるかどうかについては自明ではない。この点について明らかにするためには、理論モデルや実験室実験といった仮想的な環境下でのデータではなく、実際の企業データにもとづいた分析が必要になる。しかし、一般に業績評価や報酬

に関するデータは企業の内部情報であるため、入手が困難であり、実際の企業データをもちいた経験的研究は少ない (Moers 2007; Lazear and Oyer 2013)。

業績評価とインセンティブの関係について、実際の企業データをもちいた経験的研究の数は多くはないが、それでも、これまでの先行研究では様々な業績評価とボーナス・プランの関係が明らかにされてきた。例えば、Banker et al.(1996) および Banker (2001) では、小売業のデータをもちいて、財務指標にもとづくインセンティブ・プランの導入による効果について検証を行った。この小売業では、各店舗の売上高によってボーナス額が決定されるインセンティブ・プランを導入している。Banker et al.(1996) および、Banker(2001) の結果では、インセンティブ・プランを導入した店舗は、(1) 導入後に売上が上昇し時間とともにその効果は強まったこと、(2) 導入後に離職した従業員にくらべて、滞在を選択した従業員のパフォーマンスが高いことを明らかにした。これらの結果は、業績評価とインセンティブを組み合わせることが、企業の業績および経営にとって重要な影響を与えることを示している。一方、これまでの先行研究が店舗の売上高のような財務指標にのみ注目してきたわけではない。Banker (2000) では、顧客満足のような非財務指標に基づくインセンティブ・プランを導入したホテル・チェーンをリサーチ・サイトとし、そのインセンティブ・プランの効果について検証した。Banker (2000) の結果では、インセンティブ・プランの導入後に、これらの非財務指標が改善され、その結果として財務パフォーマンスも向上していることが明らかになった。また、Matsumura and Shin(2006) は、相対的業績評価に基づくインセンティブ・プランを導入した郵便事業を対象に分析を行い、比較される店舗との天候や地域事情による売上の変動などの共通の不確実性 (see Lazear 1981, Holmstrom 1982) が高いほどインセンティブ・プランの効果が高まることについて明らかにした。また、これらの客観的な評価だけではなく、上司の査定や裁量といった主観的な評価に注目した研究も存在する。Gibbs et al.(2003) は、自動車業界に対して主観評価に関する質問票調査を行い、評価や査定に使用される基準が短期的な利益に偏重している場合や当初の見込みとは異なり目標達成が困難な状況になった場合などは、主観評価が採用されやすいことを示した。これらの一連の研究は、財務指標、非財務指標、相対的業績評価、主観評価など様々な業績評価とボーナスとの関連について明らかに

している。また、Ittner et al. (2003) では、バランスト・スコア・カード（以下、BSC）に基づくインセンティブ・プランを導入した米国銀行のケースを取り扱っている。Ittner et al. (2003) の結果によれば、この銀行はインセンティブ・プランの導入後、BSC の各指標が主観的に重みづけられ、被評価者である従業員からの不満が続出した結果、BSC に基づくインセンティブ・プランを廃止している。

このように、これまでの先行研究は様々な業績指標・業績評価とボーナス契約の関係について明らかにしてきた。しかし、企業はインセンティブ・システムとして、ボーナスプランなどの短期的なインセンティブ契約だけではなく、昇給や昇進といった報酬の支払いが長期に渡るインセンティブも利用している。^{*1}企業が短期的インセンティブに加えて、長期的インセンティブを利用するのはなぜだろうか。先行研究では、企業における長期的インセンティブの役割として、(1) 短期インセンティブとの代替効果、(2) ホワイトカラーなど成果の測定が困難な場合のインセンティブとしての役割などが指摘されている。

まず、次期以降の報酬が当期の成果によって左右される場合、次期以降の報酬の増減がエージェントにとって当期のインセンティブの役割も果たす、いわゆるキャリア・コンサーンの影響が知られている (Holmstrom 1999)。そのため、長期的インセンティブが短期的インセンティブ契約の効果を一部代替できる。通常、エージェント（CEO や従業員といった評価される人物）のキャリア初期において、キャリア初期の成果はエージェントの能力を示すシグナルであるため、キャリア後期での職位や報酬を決定するうえでキャリア初期の成果は重要となる。それゆえ、キャリア後期の報酬はキャリア初期の成果の影響をうける。キャリア初期の成果は当期だけでなく次期以降の報酬額の増減も伴うため、キャリア後期の成果がもたらす報酬額の増減に比べてがリスクが大きい。それゆえ、リスクとインセンティブのトレード・オフの観点から、キャリア初期のインセンティブ係数はキャリア後期のそれと比べて小さく設定することが効率的となる (Gibbons and Murphy 1992)。言い換えれば、次期以降の報酬という長期的インセンティブが、キャリア初期の

^{*1} ここでいう長期的インセンティブは勤務年数に応じて自動的に賃金が増加するといった年功序列などの賃金プロファイルとは区別する。こうした前もった昇給の取り決めは、被評価者のインフルエンス活動を抑制するという効果をもっているものの、被評価者の努力や能力開発を引き出すインセンティブとしては機能していない (Prendergast 1999)。本研究では、あくまで被評価者の努力や能力開発を引き出すインセンティブとしての昇給や昇進に注目する。

報酬という短期的インセンティブの効果を代替しているのである。Ederhof(2011)は、国際企業のデータをもちいて、昇進の可能性が大きいキャリアの短い組織下層における従業員のインセンティブ報酬の割合は小さい一方で、昇進の可能性が小さいキャリアの長い組織上層における従業員のインセンティブ報酬の割合は大きいことを明らかにしており、この理論による予測を支持している。

次に、ホワイトカラーに代表される成果指標の測りにくい職種へのインセンティブとしても、長期的インセンティブが重要であることも近年指摘されている。ホワイトカラーに対して短期的インセンティブよりも長期的インセンティブが適している理由は、(1) マルチタスクであるため、ボーナスに連動させる適切な業績指標が存在しない、もしくは難しい (Holmstrom and Milgrom 1992)、(2) 研究開発など成果の発現が長期間に渡るため、短期では業績が測定できないことなどがあげられる。このような観点から、業績評価とインセンティブの関係について研究を行ったのが、Grabner(2014) および Baik et al.(2016) である。まず、Grabner(2014) の研究では創造性 (Creativity) とインセンティブのトレード・オフに注目する。創造性が戦略上重要である企業において成果ベースの報酬 (performance based pay) を利用することは (1) 内発的動機づけを損なう、(2) 従業員の行動を短期的にするといった問題がある。一方で、成果ベースの報酬という短期的インセンティブを全く使わないならば、従業員は採算度外視の製品をつくるという問題 (Grabner 2014 はこれを「芸術のための芸術 (art for art's sake)」と呼んでいる) が生じる。それゆえ、創造性が戦略上重要である企業においては、成果ベースの報酬という短期的インセンティブを利用することに関してトレード・オフが存在すると Grabner (2014) は主張する。このトレード・オフに関して、Grabner (2014) は創造性が戦略上重要である企業は、主観評価による査定を成果ベースの報酬と合わせて使うことで、トレード・オフを解消するとしている。すなわち、主観評価に基づく長期的インセンティブも併用することにより、短期視点になりがちである成果ベースの報酬のコストも主観評価を併用することで抑えつつ、かつ短期的インセンティブに加えて長期的インセンティブを使うことでインセンティブの水準も維持できる、という両者の補完的な関係について明らかにした。また、Baik et al.(2016) は、コンサルティング会社のデータをもちいて、企業間におい

ては明示的インセンティブ（ボーナスやストック・オプション型の報酬制度）と暗黙的インセンティブ（昇進の可能性）の関係について分析を行っている*²。Baik et al.(2016)の結果では、職務階層が上昇するにしたがって暗黙的インセンティブの割合が低下し、明示的インセンティブの割合が上昇していた。この結果について Baik et al.(2016)は、組織階層が上昇し権限が増えることで財務指標の情報価値が上昇することため、明示的インセンティブの割合が増えると解釈している。この Baik et al.(2016)の結果は、権限移譲が行われるほどインセンティブ報酬の割合は高まるという権限移譲に関する一連の研究結果とも整合的である (Nagar 2002; Abnerthy et al. 2004; Moers 2006)。逆にいえば、組織階層の下層に位置し、さほど権限をもたない従業員にとって、財務指標は情報として集約されすぎているため、明示的インセンティブを結びつけることが難しく、暗黙的インセンティブの価値が高まるのである。すなわち、組織下層の従業員にとっては長期的インセンティブの価値は比較的高い。

2.2 業績評価と長期的インセンティブに関する経験的研究

このように、昇給や昇進といった長期的インセンティブは企業において重要な役割を果たしている。しかし、企業が実際にどのような業績指標・業績評価と長期的インセンティブを組み合わせているかについて経験的な証拠を提示した研究は少ない。業績評価と昇進について取り扱った数少ない研究として、Cichello et al. (2009)、Campbell (2008)、Grabner and Moers (2013) があげられる。

Cichello et al.(2009)では、米国企業の公開データをもちいて事業部 ROA と事業部マネージャーの昇進および離職の関係について検証を行っている。Cichello et al.(2009)の分析結果では、事業部 ROA は昇進よりも離職を強く説明しており、昇進に対する会計指標の影響は相対的に小さいことを示した。

Campbell(2008)は、非財務指標と昇進のインセンティブ効果について検証を行ってい

*² Baik et al.(2016)では、ストック・オプション型の報酬制度を長期的インセンティブ (long-term incentive) と呼んでいるが、これは本研究の取り扱っている長期的インセンティブとは異なる。本研究の長期的インセンティブは支払いが長期に渡ることを前提としている一方で、ストック・オプション型の報酬は支払い自体が複数期間に渡るわけではない。

る。Campbel(2008)の分析結果では、(1)財務指標の影響をコントロールしたうえでも、昇進に対する非財務指標の影響は有意に正であったこと、(2)昇進後には従業員の非財務指標を向上させるための行動が低下したことが明らかになっている。これらの分析結果は、非財務指標が単に昇進に役立つ情報を提供してだけでなく、昇進というインセンティブと連動することで、非財務指標が従業員の行動を引き出すことを示している。

Cichello et al.(2009)およびCampbell(2008)は財務指標および非財務指標と長期的インセンティブの関係について有用な知見をいくつも提供している。しかし、昇進や昇格、昇給といった長期的インセンティブの決定において、現実では主観評価が利用されることが多い。例えば、労政研究所(2002)^{*3}の実態調査によれば昇給や昇格の決定には態度考課や能力考課といった基準が設定されている。学術面においても、Dohmen(2004)やDeVaro and Waldman(2004)の分析結果は、上司の主観的な業績評価が将来の昇進を予測することを明らかにしている。これらの実務面および学術面からの示唆にも関わらず、これまで長期的インセンティブにおける主観評価の役割については十分に注目されてこなかった。主観評価と昇進の関係に注目した数少ない研究としてGrabner and Moers(2013)があげられる。Grabner and Moers(2013)は、オランダの銀行をリサーチ・サイトを取り上げて分析している。彼女らの分析で興味深い点は、異なるタイプの昇進に注目したことである。この銀行では法人向けの業務と個人向けの業務という職務内容が異なる二種類の業務がある。したがって、昇進後に同じ職務内容に従事するタイプの昇進(例：個人向け業務から個人向け業務)と昇進後に異なる職務内容に従事するタイプの昇進(例：個人向け業務から法人向け業務)の二種類の昇進が存在することになる。彼女らの分析結果では、昇進前と昇進後の職務が同様のものである場合は財務指標が昇進と強く相関しているものの、昇進前と昇進後で職務内容が大幅に異なる場合は非財務指標が昇進と強く相関していた。この分析結果は、昇進前の職務での財務指標は現在担当している職務における被評価者の能力について有用な情報を提供するものの、昇進後の職務が異なる場合には現在担当している職務での財務指標は情報価値が低い。むしろ、現在の職務には現れない潜在能力のようなものを評価できる主観評価の方が情報価値は高いと彼女らは結

^{*3} 労政研究所, 2002年, 『最新人事考課事情：“成果主義”のカギを握る仕組みと実践』労働行政

論づけている。Grabner and Moers(2013) のこの主張は昇進後の職務内容が異なる場合、なぜ主観評価が利用されるのかという経営実践を明快に説明している。ただし、彼女らのこの主張は、昇進の長期的インセンティブとしての機能よりも、むしろどのような従業員をどのポストにつけるべきかという職務配置の機能に注目している。そのため、Grabner and Moers(2013) の結果では、昇格や昇給といった職務配置や職務内容の変更を伴わない長期的インセンティブにおいても、主観評価が利用されているのは何故かという疑問に答えることはできない。

2.3 長期的インセンティブにおける主観評価の役割

ボーナスなどの短期的インセンティブとは異なる長期的インセンティブの特徴は、報酬の支払いが一回きりではなく長期に渡ることである。さらに、一度昇格や昇給を決定してしまえば、降格や減給をすることは難しくなってしまう。事実、Baker et al. (1994a) の結果では、昇格に比べて降格は極めて少ない。そのため、従業員が将来にわたって生み出す成果が長期に渡って支払われる報酬に見合うかどうか判断することが重要となる。いいかえれば、従業員の昇給や昇格を決定するためには、従業員の生産性や能力を評価する必要がある (Becker 1964; Gibbons and Waldman 1999)。実際、労働経済学における一連の経験的研究では従業員の能力を反映する勤続年数や学歴といった指標が賃金額を決定する重要な要因であることが示されている (see Lazear 2009; Lazear and Oyer 2013)。

本研究では、従業員の生産性や能力といったものを評価するうえで、主観評価が客観評価よりも適していることを経験的に示す。そのために、主観評価の継時的な変化に注目する。財務指標や非財務指標といった客観評価は従業員の当期の努力や行動についての情報を提供する。しかし、これらの指標は客観的に測定されるがゆえに、客観評価では従業員の前期以前の行動や成果について考慮したりはしない。一方で、主観評価は上司など人間が行うため、これまでに獲得した過去の情報も考慮したうえで当期の成果について評価を行うことができる (Baker et al. 1994b)。例えば、(1) 予期できない要因により当期の実績が期待された目標を達成できない状況での目標を下方修正する (Gibbs et al. 2004; Merchant and Manzoni 1989)、(2) 新しいインセンティブシステムを導入時に旧インセ

ンティブシステムでの業績評価を参照する (Woods 2012)、といった主観評価の機能は、前期以前の情報を評価に取り入れるという主観評価の性質に起因している。もし、主観評価が適切に過去の情報を蓄積しているならば、主観評価は従業員の生産性や能力を判断するうえで、時間とともに正確になっていくと考えられる*4。

*4 過去の情報を蓄積するだけならば、過去の客観成果も含めた公式などにより客観的な評価を行うことも考えられる。しかし、従業員の生産性や能力を判断するうえで、過去の客観成果なども含めた公式によりできるだけ適切な評価を行うためには、従業員の生産性や能力を反映する指標を事前に想定する必要がある。これらの指標には従業員の学歴や勤続年数など事前に想定しやすいものもある一方で、景気状況、競合との関係、部署間の違い、従業員個別の事情など事前に想定することが困難なものも多い。したがって、評価に柔軟性をもたせるために主観評価を採用していると考えられる。ただし、この予想は本研究の分析結果から支持されるものではないため、本研究の限界に記述した。

3 仮説の導出

本章では、本研究における仮説の導出を行う。本章の前半では、主観評価が過去の情報を取り入れていることを示すため、主観評価と過去の情報との関連についての仮説を導出する。本章の後半では、主観評価が過去の情報を取り入れることで、能力の高い従業員と低い従業員を区別できるようになる、すなわち時間経過により主観評価の識別力は増加するという性質について、仮説の導出を行う。

3.1 過去の情報との関連

主観評価を行う際に、評価者は従業員の能力を評価する際に前年度の成果や昇進などいくつかのベンチマークとなる過去の情報を参照することが考えられる。評価者がこれらの情報を参照する理由は、これらの情報が従業員の能力を反映しているからである。実際に、多くの分析的研究では、そのような仮定が設定されている (ex. Gibbons and Waldman 1999)。すなわち、評価者は能力の推定精度を改善するために、被評価者の能力を反映しているこれらの情報を参照すると考えられる。本研究では、このような過去の情報として (1) 前期成果および (2) 前期昇進を考える。なお、勤続年数や学歴など、被評価者の能力を反映するベンチマークは他にも多数存在するにもかかわらず、前期成果および前期昇進というベンチマークに注目するのは次のような理由からである。本研究において注目すべき側面は、主観評価が過去の情報を取り入れることで、時間経過により主観評価の識別力が増すという性質である。この主観評価の性質において重要であるのは、時間経過により主観評価が含む情報量が増加するという点である。この点に関して、勤続年数や学歴といったベンチマークは、期間を通して不変であるため、時間が経過しても主観評価に含まれる情報量が増えることはない。一方、前期成果や前期昇進といった情報は毎年更新されるため、主観評価に含まれる情報量が時間とともに増加する。本研究では、時間経過により主観評価に含まれる情報量が増加するため、主観評価の識別力が増加すると

いう側面に注目するため、前期成果および前期昇進といったベンチマークに注目する^{*5}。

まず、評価者は従業員の能力を評価するうえで、前期の成果を参照すると考えられる。これは、能力の高い従業員は当期だけではなく、それ以前からも一貫して高い成果を挙げていると想定されるためである。この想定は一般的であり、多くの分析的研究でも高い能力をもつ従業員は一貫して高い成果挙げる可能性が高いと想定されている (Becker 1962; Prendergast 1999; Gibbons and Waldman 1999, 2006; Lazear 2009)。実際、過去の成果は当期成果の期待値とも考えられており、当期の成果を主観的に評価する際にも前期の成果が影響することが経験的に明らかにされている (Thorsteinson et al. 2008; Woods 2012)。したがって、以下の仮説をたてる。

仮説 1.1: 前期の成果が高ければ、当期の主観評価は高い。

また、評価者は評価対象である従業員が前年度に昇進している場合、主観評価を高くつけると考えられる。なぜならば、従業員が前年度に昇進したという事実は前年度までに高い成果を挙げていたことを反映するからである (Lazear 1981)。特に、昇進の意思決定を他者が行っている場合、昇進した従業員は他者がその従業員の能力を高いと判断したために昇進したと考えられる。そのため、昇進したという事実は、他者がその従業員の能力を高く評価したという情報を発信するシグナルとしての効果をもっている (DeVaro and Waldman 2012)。したがって、前年度に昇進している場合、評価者は高い主観評価をつけると考えられる。ゆえに、以下の仮説をたてる。

^{*5} なお、前期より前の成果もしくは、その累積が主観評価に影響することも考えられる。しかし、前期より前の成果や昇進といった変数が当期の主観評価に与える影響を検証するためには、前期より前の成果や昇進が主観評価にどのような形で影響を与えるのか、すなわち前期より前の成果および昇進と主観評価がどのような関数になっているのか特定する必要がある。このような関数を特定することは非常に困難である。なぜなら、前々期の成果は前期の昇進に影響を与えるため、前々期の成果をモデルに組み込むためには、前期に昇進したサンプルと前期に昇進していないサンプルとで主観評価に対する影響力を区別できるようにモデルを設定する必要がある。そして、このモデルの複雑度は1期前の成果および昇進を考慮するたびに、指数関数的に増加していく。そのようなモデルは設定することが困難なだけでなく、解釈も困難になる。本研究では、あくまで前期成果および前期昇進のみに注目する単純なモデルを採用することで、モデルの特定化の誤り、および解釈が困難になるといった問題を回避する。また、(1) 前期成果および前期昇進が当期の主観評価に取り入れられている、(2) 当期の主観評価と次期の主観評価が関連している、という二条件さえ満たされれば、次期の主観評価は二期以上前の成果および昇進を考慮できていることになる。本研究の分析結果では、上位および下位グループにおいてはこの二条件が満たされている。

仮説 1.2: 前年度に昇進していれば、当期の主観評価は高い。

3.2 主観評価の識別力

長期的インセンティブは支払いが長期に渡るため、能力の低い従業員を昇進・昇給させてしまったときのコストが非常に高い。それゆえ、従業員の能力を識別することが重要となってくる。本研究では、この識別力という観点に注目し、主観評価が能力の高い従業員と低い従業員を識別できるかどうかを明らかにする。前節の仮説が支持されていれば、主観評価は過去の情報を考慮・蓄積できるため、時間が経過することで従業員が優秀であるのかどうかを識別できるようになると考えられる。この時間とともに主観評価が能力の高い従業員と低い従業員を識別する力が高まるという効果を、本研究では、(1) 主観評価の値が従業員の能力を反映する値へと移動していく効果、および (2) 主観評価の誤差が低下していく効果が合わさったものであると考える。前者を評価の推移による効果、後者を精度の向上による効果と呼ぶことにする。

まず、評価の推移とは評価対象である従業員が予想された水準を上回る（下回る）はたらきをした場合に、その従業員に対する評価が向上（低下）する効果である。評価の推移がある場合、主観評価が適切に運営されるならば、主観評価の値は、その従業員の能力を反映した値に移動していくことになる^{*6}。それゆえ、時間が経過すれば、主観評価が推移することにより、能力の高い従業員の主観評価は高くなり、能力の低い従業員の主観評価は低くなる。そのため、能力の高い従業員と低い従業員の主観評価の差は、時間が経過することで増大していくと考えられる。したがって、以下の仮説 2.1 をたてる^{*7}。

^{*6} もちろん、認知バイアスなどの影響により、主観評価が適切に運営されておらず、従業員の本来の能力を反映した値に移動しないことも考えられる。この影響については、追加分析にて考慮している。

^{*7} なお、この効果は主観評価が本来の従業員の能力に移動していくという変化に加えて、主観評価の感度が向上したという効果を含む可能性がある。業績評価の感度 (sensitivity) とは、Banker and Datar (1989) によって提唱された概念である。Banker and Datar (1989) の理論モデルは業績指標とインセンティブ契約の関係をとり扱っており、リスク回避的なエージェントとリスク中立的なプリンシパルを想定している。この理論モデルでは、エージェントの行動が観察できないため、エージェントの行動を反映する業績指標をもちいて間接的にエージェントの行動をコントロールする契約をインセンティブ契約を結ぶ。このようなセッティングのもとでは、業績指標がエージェントの観察不能な行動を反映する程度に応じてプリンシパルはインセンティブ契約を結ぶ (Holmstrom 1979)。すなわち、業績指標の情報価

仮説 2.1: 時間が経過することで、能力の高い従業員と低い従業員の主観評価の差は大きくなる。

次に、精度の向上とは、時間が経過し利用できる情報量が増加することで、主観評価が正確になっていくという効果である。評価の推移との違いは、次のようなケースを考えればわかりやすい。例えば、評価対象である従業員がちょうど予想された水準のはたらきをした場合、主観評価は向上も低下もないため評価の推移はない。しかし、従業員が予想された水準のはたらきをしたという事実は、当初の主観評価が正しかったことを保証するため、評価者はその主観評価をより信頼に足るものだと判断し、主観評価の精度は高まることになる*⁸。したがって、以下の仮説 2.2 をたてる。

仮説 2.2: 時間が経過することで、主観評価の精度は高まる。

最後に、主観評価の識別力に関する仮説をたてる。本研究では、識別力を (1) 評価の推移による効果、(2) 精度の向上による効果という二つの効果に分類した。この考え方は、キャリア・コンサーンの代表的な理論研究である Holmstrom (1999) のモデルとも整合的である。Holmstrom (1999) のモデルでは、エージェントの生み出す成果がエージェントのもつ能力に依存するため、プリンシパルは報酬額を決定するうえで、エージェントがこ

値が高ければインセンティブ契約における業績指標の比重は高くなる。Banker and Datar (1989) は、Holmstrom (1979) の設定を弱めることで、インセンティブ契約における業績指標の比重が感度と精度の積に依存することを示した。感度とは、エージェントが行動 (a とする) を 1 単位変えたことによる業績指標 (y とする) の期待値の変化分であり、 $\partial E[y|a]/\partial a$ で定義される。すなわち、業績指標の感度とは、その業績指標がエージェントの行動変化をどれだけ鋭敏に捉えるかを表している。精度とは、業績指標が期待される成果からばらつく度合いであり、 $1/\text{Var}[y]$ で定義される。すなわち、業績指標の精度とは業績指標の正確さの度合いを表している。評価の感度が向上すれば、わずかな能力差であっても、その評価差を大きく捉えられることを意味する。分析期間において、両者の能力差が一定であると仮定すると、主観評価の感度向上により両者の評価差が増大すると予測される。時間が経つことで評価者の評価する能力自体が向上している場合、このような主観評価の感度向上といった効果が見られる可能性がある。本研究における能力の高い従業員と低い従業員の主観評価の差が時間とともに増大していくという分析結果が、主観評価が本来の従業員の能力に移動していくという効果によるものなのか、主観評価の感度の向上によるものなのかを区別することができない。この点は本研究の限界であるが、この限界が時間経過により主観評価の識別力が増すという本研究の主要な結論に影響を与えることはない。

*⁸ この精度の向上における精度という概念は、Banker and Datar (1989) における精度とも完全に対応しており、主観評価のばらつきの程度 (分散の逆数) として捉える。ただし、Banker and Datar (1989) の理論モデルでは推定すべきパラメーターがエージェントの努力量であるのに対して、本研究で想定しているべきパラメーターはエージェントの能力であるという点において文脈が少し異なる。

れまでに挙げた成果からエージェントの能力を推定する必要がある。Holmstrom (1999)の結果では、当期の成果のみから得られるエージェントの能力に関するシグナルを z_t 、前期までの累積成果から得られるエージェントの能力に対する推定値を m_t とした場合、当期におけるエージェントの能力に対する推定値 m_{t+1} は z_t と m_t の加重平均となる*⁹。したがって、 z_t が m_t よりも高い (低い) 場合、 $m_{t+1} > m_t$ ($m_{t+1} < m_t$) となるためエージェントの能力に対する推定値が上方 (下方) へ移動することになる。同時に、当期の成果のみから得られるエージェントの能力に関するシグナルを z_t は、エージェントの能力に対する推定値である m_{t+1} の分散の低下、すなわち精度の向上ももたらす。実際、Holmstrom (1999) のモデルでは、エージェントの能力に対する推定値 m_{t+1} の精度 (分散の逆数) h_{t+1} は、前期までの推定値の精度 h_t と当期のシグナルの精度 h_ϵ の和になる。これらの結果は、Holmstrom (1999) という特定のプリンシパル・エージェントモデルの設定の下だけでなく、次のような一般的な統計的推定の文脈においても成立する。すなわち、シグナル z_t がパラメータ η を平均、 $1/h_\epsilon$ を分散とする正規分布にしたがい、これまで得られたシグナル (z_1, \dots, z_t) から η の事後分布を推定するという文脈である。 t が大きくなり、シグナル (z_1, \dots, z_t) の流れが増えるほど、事後分布の平均 m_t はパラメータ η の値へ推移してゆき、事後分布の分散 $1/h_t$ は 0 へと近づいていく。すなわち、パラメータ η の事後分布が η へと収束していくことになる。これらの前提に基づけば、時間が経過し情報が蓄積されることにより、主観評価が推移し、かつ主観評価の精度が高まることで、主観評価の識別力が増すと考えられる。したがって、識別力を評価の推移および精度の向上という二つの効果が合わさった効果であると捉え、以下の仮説 2.3 をたてる。

仮説 2.3: 時間が経過することで、主観評価の識別力は高まる。

*⁹ 各変数に対する重みは、エージェントの能力推定に対する精度 (precision) となる。

4 研究方法

本章では、本研究の採用した分析方法について説明する。まず、本研究のリサーチ・サイトとなった企業について、その人事制度や評価制度の詳細について記述する。次に、分析に使用する変数について、それらの変数をどのように導出したのか、すなわち操作化の方法について説明を行う。最後に、仮説の検証に使用する分析モデルおよび手順について記述する。

4.1 リサーチ・サイト

本研究では、日本の一部上場企業 1 社における 2004-2013 年度の人事データをもちいて、前章で導出した仮説の検証を行う。以下では、この企業を A 社と呼ぶことにする。A 社は IT 業界に属しており、異なる製品・サービスを提供する二つの事業部から構成されている。各事業部において、従業員は主に営業職と技術職どちらか一方の職務を担当する。また、A 社では職能資格制度を採用しており、職位と職能資格に厳密な対応関係はないものの、特定の職位につくためには特定のランク以上の職能資格が必要となる資格先行・昇進追隨の昇格・昇進パターンをとる、典型的な職能資格制度を採用する企業であるといえる。職能資格制度を採用する企業において、従業員の給与は職位ではなく職能資格によって決まる。A 社においても、給与は 9 段階の職能資格にもとづき定められている。

A 社の評価制度のもとでは、夏・冬の賞与評価（7 段階評価）と年度末の昇給評価（6 段階評価）の計 2 種類の評価が行われる。これら 2 種類の評価は、事業部および等級を基準とする 7-8 人程度のグループごとに 1 人の評価者（上司）が割り当てられ、評価者が評価グループに所属する従業員全員の評価を行う。夏・冬の賞与評価では、従業員の当期成果を評価する方針である。これらの賞与評価は、営業職では営業成績、技術職ではバグの発生件数などの客観指標をベースに決定される。一方、年度末の昇給評価は賞与評価の決定後に行われ、評価方針は当期成果を反映する当期の賞与評価を 4 割、従業員の能力や勤務態度に関する評価を 6 割で評価する。この 6 割の評価内容は従業員の能力や勤務態度を

評価するという方針のみ決められており、具体的な評価方針（ex. どのような能力を重視するか等）は評価者である上司に一任されている。なお、本研究ではデータの制約から従業員の給与額に関するデータは利用できない。しかし、これらの賞与評価および昇給評価が従業員の受け取る給与額を決定するうえで重要な変数であることには変わりなく、各評価と給与額の関係はそれぞれ次のような対応関係にある。まず賞与と賞与評価の関係について、A社では賞与に関する予算が評価グループごとに毎期設定され、グループ内における各従業員の賞与額は賞与評価に基づき、その分配額が決定される*¹⁰。そのため、賞与と賞与評価は評価グループに割り当てられる予算を介して間接的に連動している。一方、基本給と昇給評価の関係について、A社における基本給は前述した9段階の職能資格で大枠が決定され、くわえて各段階の職能資格内でさらに細かなレベルが設定されている。基本給は従業員の所属するレベルに応じて決まっており、昇給評価は職能資格内のレベルを上昇させるかどうかに関する評価である。したがって、賞与評価と昇給評価のいずれも給与額と一对一の厳密な対応関係にあるわけではないものの、間接的には連動しており給与額を決定するうえで最も重要な変数であることには変わらない。なお、A社では上記の人事システムを2004年度より採用している。

4.2 変数の説明

4.2.1 主観評価

本研究では、A社における2種類の評価（賞与評価および昇給評価）をもちいて上司の主観評価を算出する。A社の人事制度のもとでは、昇給評価は賞与評価の決定後に上司により行われ、当期成果に関する評価（すなわち、賞与評価）が4割、能力に関する評価が6割という評価方針であった。そこで、本研究では次の式により従業員*i*の*t*期の主観評価 (*subjective1*) を定義する。なお、下式で6/7を乗じている理由は、昇給評価と賞与評価における評価尺度の違いを調整するためである。

*¹⁰ A社では、このような賞与に関する制度を設けているため、賞与評価は相対評価により決定される。賞与評価の平均は評価グループごとに0になるように、バランスを取らなければならない。この点に関して、本研究では等級などのコントロール変数を採用するなどの処置をとっており、このようなA社における賞与評価制度が、本研究の主要な結論を変更しうる程の影響を与えることはない。

$$subjective1_{i,t} = rise_{i,t} - 0.4 \times \frac{6}{7} \times performance_{i,t}$$

- $rise_{i,t}$:従業員 i の t 期における昇給評価 (6 段階評価)
- $performance_{i,t}$:従業員 i の t 期における賞与評価 (7 段階評価)

一方、「成果に関する評価が4割、能力に関する評価が6割」というA社の昇給評価の評価方針は、厳密に評価点数が分けられている訳ではなく、あくまで上司の裁量に委ねられている。それゆえ、評価にかかる手間や認知的な負荷を避けるため、上司は各評価を個別に集計し合算するのではなく、賞与評価を基準とした係留と調整 (Kahneman and Tversky 1974) により評価を行っている可能性がある。上司が係留と調整により昇給評価を行っていたとしても、昇給評価の調整過程で上司の主観性が入っていることには変わらない。しかし、A社の評価方針にもとづいた逆算を行う上記の操作化では、上司の主観性を適切に測定できないという問題が生じる。この点における分析結果の頑健性を確保するため、上司の係留と調整を踏まえた下式としても主観性 ($subjective2$) を操作化する。具体的には、まず賞与評価の7段階ごとに昇給評価の平均値 ($Mrise_p$) を算出する。ただし、添え字の p は賞与評価 ($performance_{i,t}$) を表している。そして、算出した平均値を昇給評価から減算することで主観性 ($subjective2$) を定義する。なお、この操作化は昇給評価を被説明変数、賞与評価の各評価に関するダミー変数を説明変数とした回帰分析を実施したときの残差と一致する。それゆえ、 $subjective2$ を昇給評価から賞与評価をコントロールしたものの解釈しても問題ない。

$$subjective2_{i,t} = rise_{i,t} - Mrise_{p(i,t)}$$

- $rise_{i,t}$:従業員 i の t 期における昇給評価
- $Mrise_{p(i,t)}$:賞与評価が p であるときの、昇給評価の平均値
- $p(i,t)$:従業員 i の t 期における賞与評価 (夏・冬の平均)

なお、Hoppe and Moers(2011) では、主観評価にも二つのタイプがあるとされている。

一つは客観評価とは別に主観評価を取り入れるタイプ、もう一つは客観評価をどのように組み合わせるかについて主観評価を取り入れるタイプである。本研究で取り扱う主観評価は、この研究で主張されているうちの前者のタイプの主観評価であるため、上記のような操作化を行っている。

4.2.2 前期成果

A社の評価方針のもとでは、賞与評価は夏・冬の年2回行われ、当年度の成果にもとづき評価される。本研究では、この賞与評価（夏・冬の平均値）を従業員の成果を表す変数（*performance*）として採用する。ただし、賞与評価は上司によって行われており、公式的・機械的に評価されているわけではない。そのため、本研究で成果の代理変数として採用する賞与評価が上司の主観性を含んでいるため、従業員の成果を適切に反映しないのではないかという懸念が生じる。この点について、リサーチ・サイトに対してインタビューを行った。A社において賞与評価は、昇給評価とは異なり、100%当期成果を評価するという方針である。そのため、上司により賞与評価が行われるとはいえ、客観的な数値（ex. 営業成績、プログラムのバグ発生件数など）に依存させる傾向が強いとの回答が得られた。それゆえ、賞与評価に上司の主観性が入り込む余地は非常に少ないと考えられる^{*11}。なお、ここでの前期や当期は会計年度における前期や当期を指すわけではない。本研究で使用しているデータは、人事データであり、前期や当期といった表現は人事評価における前期および当期を指している^{*12}

4.2.3 前期昇格

A社では、職能資格制度が採用されていた。本研究では、この職能資格が上昇することを昇格（*promotion*）として扱う。欧米では、職務等級制度を取る企業が多いため、先行研

^{*11} このような理由から、本研究では賞与評価（*performance*）を客観評価の代理変数として取り扱い、上記で定義した主観評価（*subjective1* および *subjective2*）との比較も行う。この主観評価との比較という文脈では、賞与評価を「完全な客観評価」というよりはむしろ「より客観的な評価」と解釈することで評価における主観性の役割について考察することが目的である。

^{*12} A社は3月決算の企業であるため、3月を決算とする会計データと夏および冬を基準とする人事データの間で数ヶ月分のずれが生じている。しかし、本研究の分析に使用しているデータは人事データのみで完結しており、こうした会計年度とのずれに影響を受ける変数は存在しない。

究で取り上げた欧米の研究における昇格 (*promotion*) は、職位の上昇を表すと言える。職能資格制度を取る企業において、給与と連動し、職位の上昇のための必要条件ともなる資格の上昇は、昇進よりも実質的な組織内での序列の向上であり、職務等級制度を採用する欧米の企業を対象とした先行研究における昇進と対応するものであると考えられる。本研究において、先行研究でいう昇進を昇格に置き換えることにより生じる違いは、ポストの数に制限があるか否かである。限られたポストに誰がつくか、についての昇進は、そのポストに就く可能性のある他者の数や成績による影響を受ける相対的な評価による選抜であるものの、昇格は個人の職務遂行能力をもとに決定されるため、定義上同じ部署の他者や同じ職能資格の他者の数や成績によって自身の昇格の有無が影響を受けることはない。そのため、本研究が注目する主観性と、階級の上昇の関係を検証するためには、トーナメントの要素がない（ライバルの数や成績の影響を受けない）昇格の方が有用であると考えられる。また、A社では職能資格の上昇、すなわち昇格については評価者である直属の上司の裁量で決定され、役員会は抜擢等の特別な事情がない限りは上司の決定をそのまま承認する。そのため、人事部や役員会は昇格の意思決定に事実上関与しておらず、人事部や役員会による裁定方針の影響をほとんど受けていない。したがって、昇格の意思決定は昇格を決定した上司の従業員に対する評価を直接反映している。

4.2.4 その他のコントロール変数

本研究の仮説 1.1 および仮説 1.2 において検証したい効果は、評価者が主観評価を行う際に、過去の情報を参照・基準として評価を行っているという効果である。この効果は、前期成果や前期昇進が当期成果など、他の情報には反映されなかった被評価者に関する能力の一部を反映しているために生じると、本研究では仮定している。しかし、前期成果や前期昇進といった過去の情報が反映する被評価者の能力に関する情報内容と、当期成果など他の情報が反映する被評価者の能力に関する情報内容には、当然重なりが生じる。したがって、仮説 1.1 および仮説 1.2 の検証には、前期成果および前期昇進が被評価者の能力を反映している情報内容のうち、当期成果などの他の情報が反映している情報内容をコントロールする必要がある。この点について、本研究では以下の変数をコントロールする。

まず、勤続年数 (*career*) をコントロールする。労働経済学における一連の研究では、キャリアを重ねることで従業員の能力が向上するとされている (Becker1964; Lazear and Oyer 2013)。そこで、本研究では、従業員の労働年数の代理変数として A 社での勤続年数 (*career*) を採用する。ただし、A 社では中途採用も行われており、中途採用された従業員については労働年数と勤続年数が対応しない。そのため、これらを区別するため新卒採用である場合 1 をとるダミー変数 (*new*) をコントロール変数に追加する。新卒採用に関するダミー変数をコントロールする理由は、新卒採用者と中途採用者を労働年数の差をコントロールするためである。

次に、仮説の検証に使用する前期成果 ($performance_{i,t-1}$) にくわえて、当期成果 ($performance_{i,t}$) をコントロール変数としてモデルに組み込む。当期成果は以下の二つの効果を考慮して、コントロール変数として採用している。なお、一つ目の効果は主観評価に正の影響を与える一方で、二つ目の効果は主観評価に負の影響を与える。一つ目の理由は、当期成果は被評価者の能力の関数になっていると考えられており (ex. Gibbons and Waldman 1999; Lazear 2009)、被評価者の能力を反映していると考えられるためである。ただし、前述したように *subjective2* は操作化の段階ですでにコントロールしているため、この影響は受けない。二つ目の理由は、賞与評価 (当期成果) の違いによる主観評価の裁量の程度の違いを取り除くためである。A 社での評価制度を考えれば、賞与評価 (当期成果) ごとに主観評価の裁量の程度が異なると考えられる。A 社の評価制度では、7 段階の賞与評価を決定後に 6 段階の昇給評価をつけるというものであった。昇給評価の評価方針では、賞与評価を明確に考慮することが義務づけられているため、賞与評価が高いほど昇給評価は、(評価方針を前提とすれば) 自動的に高くなる。しかし、昇給評価の尺度は 6 段階で一定であるため、賞与評価が高いほど上司の裁量によって決まる主観評価の幅が狭くなる。このような評価者の裁量の程度をコントロールするためにも、賞与評価 (当期成果) をコントロール変数として採用する。

また、従業員の能力に影響を与える変数として、従業員の職能資格 (*grade*) をコントロールする。内部労働市場に関する理論的研究によると、能力の高い従業員は高い階層に割り当てられることが効率的である (Gibbons and Waldman 1999; 2006)。内部労働市

場の古典的な経験的研究である Medoff and Abraham (1980) でも等級と成果には正の相関が確認されている。

上記以外に、営業職および技術職といった職種の違いや、事業部の違いをコントロールするため職種・事業部ごとにダミー変数を設定する。さらに、年度に関するダミー変数も分析に含める。年度に関するダミー変数を含める理由は、全社の業績や従業員全体の年度を通じた能力向上など年度間の差異をコントロールするためである。さらに、従業員ごとの平均的な能力をコントロールするため、従業員に関する固定効果を含めたモデルでも検証を行う。

4.3 分析モデル

4.3.1 過去の情報との関連

仮説 1.1 および仮説 1.2 については、以下の線形モデルをもちいて仮説の検証を行う。

$$subjective_{i,t} = \beta_0 + \beta_1 performance_{i,t-1} + \beta_2 promotion_{i,t-1} + controls + \epsilon_{i,t}$$

上式の被説明変数である *subjective1* および *subjective2* は、それぞれ評価割合および尺度を調整した賞与評価、賞与評価ごとの昇給評価の平均値を、昇給評価から除算したものであった。したがって、いずれの操作化においても *subjective* の尺度は昇給評価と同じであるため、主観評価 *subjective* の単位が 1 高いということは、昇給評価が 1 高いことを意味している。したがって、上式の各係数は、昇給評価を基準とした解釈が可能となる。例えば、上式の β_1 が 1 であった場合、「前期成果（前期の賞与評価）が 1 高ければ、当期の昇給評価が 1 高くなる」と解釈できる。したがって、上式のモデルでは標準誤差や有意水準に基づく有意性の判断だけでなく、係数の大きさを昇給評価の尺度に基づいて解釈できるという利点がある。

4.3.2 主観評価の識別力

仮説 2.1 について検証を行うためには、サンプル内における能力の高い従業員と低い従業員を特定する必要がある。この点について、本研究では以下のような方法を採用した。ま

ず、従業員ごとに主観評価 (*subjective1* および *subjective2*) の平均値を算出し、従業員が受けた平均的な主観評価の値を求めた。その後、この従業員ごとの主観評価の平均値について、上位 10 パーセントを能力の高い従業員として「*High*」、下位 10 パーセントを「*Low*」としてグルーピングした。ただし、このグルーピング操作を行ううえで、分析対象期間である 9 年間のうち主観評価の変数 (*subjective1* および *subjective2*) にデータの欠損があった場合、グループに所属する従業員の数が年度により異なってしまうため適切な比較ができない。そのため、仮説 2.1 以降の分析については、分析対象期間である 9 年間においてデータの欠損が生じていないサンプル、すなわち 2004 年度以前に入社し 2013 年度までリサーチ・サイトに在籍していたサンプルを分析対象とする。なお、このようにサンプルを限定することで、生存バイアスの問題が生じる可能性が考えられる。この生存バイアスは、*High* グループには特段の影響を与えない一方で、*Low* グループには無視できない影響を与えると考えられる。なぜなら、低い主観評価を受け続けた従業員は離職してしまい、分析対象期間の最終年度である 2013 年度まで在籍していない可能性が高いからだ。そのため、*Low* グループの結果について解釈する際には生存バイアスを考慮する必要があるが、この点については後述する。

仮説 2.1 については、次の線形モデルをもちいて仮説の検証を行う^{*13}。過去の情報を評価に取り込むことで、時間とともに主観評価が被評価者の能力の値へと推移していくなれば、*High* と *Low* の主観評価の差は一定ではなく、年度が進むごとにその差は開いていくと考えられる。そのため、以下の線形モデルにおいて β_3 が有意に正であるならば、仮説 2.1 が支持されることになる。

$$mean_{t,j} = \beta_0 + \beta_1 time_{t,j} + \beta_2 High_{t,j} + \beta_3 time_{t,j} \times High_{t,j} + \epsilon_{t,j}$$

- $mean_{t,j}$: 年度 t のグループ j (= *High*, *Low*) における主観評価の平均値

^{*13} 仮説 2.1 の検証について、生存バイアスが与える影響は分析対象サンプルに離職してしまった従業員が含まれないため、*Low* グループの平均値が本来よりも過大になってしまうというものである。すなわち、生存バイアスの影響により *Low* グループと *High* グループの差が本来よりも小さく推定されるため、仮説 2.1 の検証に使用する β_3 に関して有意性検定の検定力を低下させてしまう。しかし、本研究の結果では、この低下した検定力のもとでさえ β_3 が有意であった。そのため、生存バイアスの存在は、仮説 2.1 の分析結果における本研究の結論に影響を与えるものではない。

- $time_{t,j}$: サンプルの年度-2004*¹⁴
- $High_{t,j}$: サンプルのグループが *High* であった場合、1 をとるダミー変数

仮説 2.2 については、以下の線形モデルをもちいて検証を行う。過去の情報を蓄積することで、時間とともに主観評価が正確になっているならば、年度が進むごとにグループ内の標準偏差は小さくなると考えられる。そこで、本研究では、Holmstrom(1999) や Banker and Datar (1989) といった理論モデルを前提に、主観評価の正確さをグループ内の標準偏差の逆数で操作化する。

仮説 2.2 のもとでは、時間が経過することで、*High* および *Low* のどちらのグループも等しく主観評価が正確になると考えられる。それゆえ、本来であれば時間経過の単独効果のみをモデルに含めれば仮説の検証には十分である。それにもかかわらず、 $time$ と $High$ の交互作用項をモデルに含めているのは、生存バイアスの影響を考慮するためである。すなわち、時間が経過することで低い評価が確定してしまった従業員は離職してしまい観測できず、*Low* グループでは時間経過が主観評価の正確さに与える影響が (*High* グループに比べて) 過少に推定されることが考えられる。すなわち、*Low* グループでは生存バイアスのため、主観評価の精度が高まるという効果が観測されない可能性がある。ここでは、交互作用項をモデルに含めることで、生存バイアスの影響を受けにくい *High* グループにおける時間経過の影響のみを独立して取り出すことを想定している。すなわち、生存バイアスにより *Low* グループにおける主観評価の精度が高まるという効果が観測されなかった場合にも、*High* グループにおいてこの効果が観測されるならば、 β_3 が有意に正となり、生存バイアスの影響を受けにくい *High* グループについてのみ仮説の検証ができる。上記の議論より、 β_1 が有意に正であった場合は両グループにおいて仮説 2.2 が支持、 β_3 のみ有意に正であった場合は生存バイアスの影響を受けにくい *High* グループにおいてのみ仮説 2.2 が支持されたと考える。

*¹⁴ データセットには 2004 年度からのサンプルが含まれているものの、本研究では 1 年前の変数 (前期成果 ($performance_{i,t-1}$) や前期昇進 ($promotion_{i,t-1}$) など) を分析に使用するため、最初期の 2004 年度は分析対象サンプルから欠落している。それゆえ、分析対象サンプル内での最初期年度は 2005 年度からである。

$$1/sd_{t,j} = \beta_0 + \beta_1 time_{t,j} + \beta_2 High_{t,j} + \beta_3 time_{t,j} \times High_{t,j} + \epsilon_{t,j}$$

- $sd_{t,j}$: 年度 t のグループ $j (= High, Low)$ における主観評価の標準偏差
- $time$: 年度-2004
- $High$: グループが High であった場合、1 をとるダミー変数

仮説 2.3 の検証については、以下の線形モデルをもちいて検証を行う。モデルの解釈には、仮説 2.2 の結果が前提となる。仮説 2.2 において両グループの主観評価の精度が高まるという効果が観測された（仮説 2.2 の検証モデルにおいて β_1 が有意に正であった場合）場合、以下のモデルでは評価の更新により β_3 および評価の精度向上により β_1 が有意に正になると考えられる。一方、仮説 2.2 において $High$ グループの主観評価の精度のみが高まるという効果が観測された（仮説 2.2 の検証モデルにおいて β_3 が有意に正であった場合）場合、 β_3 が有意に正になると考えられる。

$$mean_{t,j} \times 1/sd_{t,j} = \beta_0 + \beta_1 time_{t,j} + \beta_2 High_{t,j} + \beta_3 time_{t,j} \times High_{t,j} + \epsilon_{t,j}$$

5 分析結果

本章では、本研究の分析結果について記述する。まず、分析対象となったサンプルを絞り込むサンプル選択について、その方法およびサンプルサイズについて説明する。次に、分析対象サンプルに関する記述統計を確認する。続いて、主観評価と過去の情報との関連である仮説 1.1 および仮説 1.2 の検証を行う。最後に、主観評価の識別力に関する仮説である仮説 2.1、仮説 2.2 および仮説 2.3 について検証を行う。

5.1 サンプル選択

分析対象となったサンプルは、前章で述べた変数に関して欠損のない全サンプルである。ただし、中途採用に対して一定額の給与を保証しなければならない場合など、賞与評価もしくは昇給評価が別建てでつけられているケースがあった。これらのサンプルにおける評価の変数は実際には「評価」ではないため、分析対象から除外した。また、等級が 1 および 2 のサンプルについても、分析対象から除外した。A 社では、在籍さえしていれば等級 3 までは自動的に昇格する。そのため、等級が 1 および 2 のサンプルについては昇格が従業員の評価と関連しておらず、評価基準も他のサンプルとは異なっている。したがって、評価基準の異なる等級 1 および等級 2 の従業員を分析対象から除外した。ただし、この操作に関して、本研究の主たる結果は影響を受けなかった。これらの手続きの結果、全体で 2980（従業員 × 年度）のサンプルが分析対象となった。

5.2 記述統計量

表 1 に本研究で使用する変数の要約統計量を示した。表 1 によると、 $subjective1_{i,t}$ および $subjective2_{i,t}$ の標準偏差はそれぞれ 0.66 および 0.62 であり、 $performance_{i,t}$ の標準偏差は 0.69 であった。この結果は、A 社において主観評価が賞与評価と同等程度に

表 1: 要約統計量

	平均値	標準偏差	最小値	1Q	2Q	3Q	最大値
$subjective1_{i,t}$	2.49	0.66	0.31	2.29	2.46	2.97	4.46
$subjective2_{i,t}$	0.00	0.62	-1.78	-0.32	-0.08	0.37	2.17
$performance_{i,t}$	4.48	0.69	1.50	4.00	4.50	5.00	6.50
$rise_{i,t}$	4.03	0.76	1.00	4.00	4.00	5.00	6.00
$performance_{i,t-1}$	4.52	0.66	1.50	4.00	4.50	5.00	6.50
$promotion_{i,t-1}$	0.09	0.28	0.00	0.00	0.00	0.00	1.00
$career_{i,t}$	12.97	6.15	2.00	8.00	12.00	17.00	33.00
$grade_{i,t}$	4.36	1.22	3.00	3.00	4.00	5.00	8.00

バラついていることを示している^{*15}。また、A社の人事方針に基づいた主観評価の操作化である $subjective1_{i,t}$ においては、最大値が 4.46、最小値は 0.31、賞与評価の平均値を引いた $subjective2_{i,t}$ においては、最大値が 2.17、最小値は -1.78 であった。これらの結果は、賞与評価が同じであったとしても、昇給評価の最大で 4 近くの差がつくことを示している。前年度の昇進を表す $promotion_{i,t-1}$ の平均値は 0.09 であり、サンプル全体のうち約 10 パーセントが昇進していることがわかる。勤続年数の平均値は 12.97 であり、最小値は 2 である。勤続年数の最小値が 1 でない理由は、前年度のデータがない場合、 $performance_{i,t-1}$ や $promotion_{i,t-1}$ といった前年度のデータを使う変数が作成できないため分析対象サンプルから外れたためである。最後に、等級 $grade$ の変数の最小値が 3 である理由は、サンプル選択の節にて記述したように評価基準の異なる等級が 1 および 2 のサンプルを分析から除外したためである。

表 2 に、本研究で使用する変数の相関係数表を示した^{*16}。表 2 によると、 $subjective1_{i,t}$ と $subjective2_{i,t}$ の相関係数は 0.96 であり、操作化は異なるものの上司の主観評価とい

^{*15} この解釈について、主観評価と賞与評価の尺度の違いには注意が必要である。主観評価を表す変数である $subjective1_{i,t}$ および $subjective2_{i,t}$ は昇給評価にもとづき計算されているため、その尺度は昇給評価と同様に 6 段階評価である。一方、賞与評価は 7 段階で評価されているため、賞与評価の方が標準偏差は大きくなりやすい。この尺度の違いを考慮したとしても、主観評価と賞与評価の標準偏差がほぼ同じ値であるということは、主観評価が賞与評価と少なくとも同等程度にはばらつきやすいと考えられる。

^{*16} 相関係数表において、 $performance_{i,t}$ と $grade_{i,t}$ の相関係数が 0 である理由は、賞与評価 ($performance_{i,t}$) が相対評価であり、評価グループごとに平均が 0 となるように設定されているためであると考えられる。

表 2: 相関係数表

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) $subjective1_{i,t}$	1.00							
(2) $subjective2_{i,t}$	0.96	1.00						
(3) $performance_{i,t}$	0.28	0.01	1.00					
(4) $rise_{i,t}$	0.95	0.83	0.56	1.00				
(5) $performance_{i,t-1}$	0.57	0.43	0.56	0.67	1.00			
(6) $promotion_{i,t-1}$	0.37	0.36	0.09	0.35	0.24	1.00		
(7) $career_{i,t}$	-0.05	-0.02	-0.11	-0.08	-0.11	-0.09	1.00	
(8) $grade_{i,t}$	0.09	0.10	0.00	0.08	0.05	0.10	0.78	1.00

う同じものを測定していることを示している。 $subjective1_{i,t}$ と $performance_{i,t}$ の相関係数は 0.28 であり、A 社の評価方針に基づく主観評価の操作化では主観評価と当期成果に正の相関が見られる。この結果は、主観評価のもつ被評価者の能力に関する情報内容と当期成果のもつ情報内容が一部重複していることを示している。一方、 $subjective2_{i,t}$ と $performance_{i,t}$ の相関係数は 0.01 であり、ほぼ無相関であった。この結果は偶然ではなく、 $subjective2_{i,t}$ の操作化に起因している。前章で説明したように、 $subjective2_{i,t}$ は、昇給評価 ($rise_{i,t}$) を被説明変数、 $performance_{i,t}$ に関するダミー変数を説明変数とした単純線形モデルの残差と対応している。単純線形モデルでは、説明変数と残差は直交する。したがって、説明変数と多重共線関係にある当期成果 ($performance_{i,t}$) と残差である $subjective2_{i,t}$ は理論上無相関になる^{*17}。したがって、 $subjective2_{i,t}$ は、被評価者に関する能力のうち当期成果 ($performance_{i,t}$) には反映されない側面を捉えていると解釈できる。

これらの記述統計から、 $subjective1_{i,t}$ と $subjective2_{i,t}$ のいずれの場合においても、当期成果 ($performance_{i,t}$) と完全に相関するわけではないことが分かった。すなわち、上司による主観評価は当期成果には反映されない被評価者の能力を一部 ($subjective2_{i,t}$ に関しては重複なしに) 捉えているといえる。この点について、さらに明らかにするため昇給評価 ($rise$) と賞与評価 ($performance_{i,t}$) の分布を図 1 に記載した。

^{*17} 表 2 において、 $performance_{i,t}$ と $subjective2_{i,t}$ の相関係数が 0 でない理由は、統計ソフトの小数点の扱いによる誤差だと考えられる。

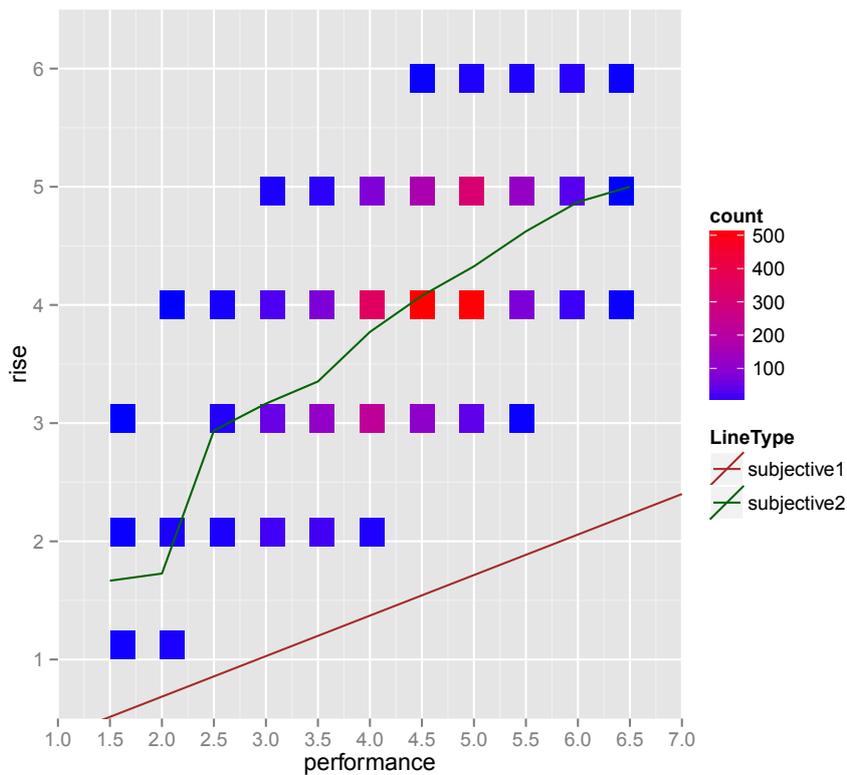


図 1: 昇給評価 (rise) と賞与評価 (performance)

図 1 は昇給評価および賞与評価の分布を示したヒートマップであり、赤色に近いほどサンプル内での観測頻度が高い。図 1 によると、例えば賞与評価が 4.0 であった場合、昇給評価は赤色に近い 4 という評価がつく可能性が最も高い。しかし、5 や 3 という昇給評価がつけられる可能性も低くはなく、場合によっては 2 という昇給評価がついたケースも存在している。すなわち、図 1 における縦側のばらつきは、上司による裁量の程度を示しており、これが本研究の分析対象である主観評価である。本研究における主観評価の変数を可視化するため、 $subjective1_{i,t}$ および $subjective2_{i,t}$ の値がゼロとなる箇所を直線および折れ線で図 1 に示している^{*18}。図 1 では、本研究における主観評価は、これらの直線および折れ線と昇給評価 (rise) との距離により操作化されていることになる^{*19}。

^{*18} $subjective1 = rise - 0.4 * 6/7 * performance_{i,t}$ であったことから、 $subjective1 = 0$ となる集合は $rise = 0.4 * 6/7 * performance_{i,t}$ となる直線になる (図 1 における直線)。

^{*19} 図 1 によると賞与評価 ($performance_{i,t}$) が 4.5 以上のサンプルについては昇給評価が 6 で頭打ちになっ

5.3 過去の情報との関連

表 3 において、仮説 1.1 および仮説 1.2 の検証を行った。分析結果については、通常最小二乗法による結果 (*OLS*) と固定効果モデルによる結果 (*panel linear*) の両方の結果を表示している*²⁰。表 3 によると、前期成果である $performance_{i,t-1}$ の係数は $subjective1_{i,t}$ のモデルで、0.523、0.491、 $subjective2_{i,t}$ のモデルで 0.517、0.483 と、いずれも正に有意 ($p < 0.001$) であり仮説 1.1 が支持されている。前期昇進である $promotion_{i,t-1}$ の係数についても、 $subjective1_{i,t}$ のモデルで、0.551、0.548、 $subjective2_{i,t}$ のモデルで 0.546、0.548 といずれの分析結果も正に有意 ($p < 0.001$) であり、仮説 1.2 が支持されている。また、これらの変数は単に有意であるだけでなく、従業員の受ける昇給評価に十分な影響を与えていることが分かる。主観評価の変数である $subjective1_{i,t}$ および $subjective2_{i,t}$ は、昇給評価をベースに算出されており、その尺度は昇給評価と同じである。そのため、これらの主観評価の変数が 1 単位増加することは、従業員の昇給評価が 1 上昇することを意味する。この解釈を踏まえたうえで、表 3 における前期成果の分析結果を見てみると、前期成果が 1 高ければ当期の昇給評価が平均的に約 0.5 高くなることを意味していることが分かる。前期昇進についても同様に、前年度に昇進している場合には当期の昇給評価が約 0.5 高くなると解釈できる。この昇給評価における 0.5 という数値は、従業員の受ける昇給評価において実質的な影響を与えている数値だと考えられる。図 1 によれば、例えば賞与評価が 4.0 であった場合、期待される昇給評価は 3 もしくは 4 になるが、前期成果が 1 高ければ（もしくは前期に昇進していれば）期待される昇給評価は 4 になり、従業員の昇給評価を決定づける要因になる。これらの結果から、前期成果および前期昇進は主観評価を決定づける重要な要因の一つになっており、評価者である上司がこれらの情報を重視していることがわかる*²¹。

ている。この影響により、仮説 2.1 および仮説 2.2 の検証について、賞与評価 ($performance_{i,t}$) が高い場合に主観評価が高くなりにくくなる。本研究では、前章で記述したように、賞与評価 ($performance_{i,t}$) をコントロール変数に含めることで、この影響について対処している。

*²⁰ 固定効果モデルでは、従業員の固定効果を考慮している。

*²¹ 表 3 の $performance_{i,t-1}$ および $promotion_{i,t-1}$ の係数に関して、通常最小二乗法と固定効果モデルで比較すると、仮説検証に使用した全ての係数について通常最小二乗法の方が大きいことがわかる。この

表 3: 過去の情報との関連

	<i>subjective1_{i,t}</i>		<i>subjective2_{i,t}</i>	
	<i>OLS</i>	<i>panel linear</i>	<i>OLS</i>	<i>panel linear</i>
<i>performance_{i,t-1}</i>	0.523*** (0.018)	0.491*** (0.020)	0.517*** (0.018)	0.483*** (0.020)
<i>promotion_{i,t-1}</i>	0.551*** (0.037)	0.548*** (0.039)	0.546*** (0.036)	0.548*** (0.039)
<i>career_{i,t}</i>	-0.007* (0.003)		-0.008** (0.003)	
<i>performance_{i,t}</i>	-0.039* (0.017)	-0.100*** (0.019)	-0.299*** (0.017)	-0.354*** (0.019)
<i>new_{i,t}</i>	0.045 (0.025)		0.046 (0.025)	
Constant	0.261** (0.088)		-1.042*** (0.088)	
<i>grade</i>	Yes	Yes	Yes	Yes
<i>depart</i>	Yes	Yes	Yes	Yes
<i>job</i>	Yes	Yes	Yes	Yes
<i>year</i>	Yes	Yes	Yes	Yes
Fixed Effect	No	Yes	No	Yes
N	2,980	2,980	2,980	2,980
Adj R ²	0.391	0.257	0.336	0.285

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, () 内の数値は標準誤差を表している。 *grade*, *depart*, *job*, *year* はそれぞれ等級、事業部、職種、年度を表しており、表中の Yes という表記はこれらの変数をダミー変数としてコントロールしていることを意味している。Fixed Effect は従業員の固定効果を表しており、コントロールしている場合は Yes と表記している。なお、固定効果モデル (*panellinear*) に関して、勤続年数 (*career_{i,t}*)、新卒 (*new_{i,t}*) の説明変数を取り除いている理由は完全な多重共線性を避けるためである。固定効果モデルでは従業員の固定効果を含むため、分析期間を通して一定の変数である *new_{i,t}* とは完全な多重共線性の関係にある。また、勤続年数 (*career_{i,t}*) についても従業員の固定効果および年度 (*year*) との完全な多重共線性の関係にある。したがって、勤続年数 (*career_{i,t}*) および新卒 (*new_{i,t}*) の説明変数については、従業員の固定効果および年度 (*year*) によってコントロールされているとみなし、分析から取り除いている。

その他の変数については、まず勤続年数 ($career_{i,t}$) の係数は-0.007 および-0.008 であり、有意 ($p < 0.05$) に負であった*22。これは、等級 ($grade_{i,t}$) が同じである場合には、勤続年数の若い方が成果が高いとする Medoff and Abraham(1980) の分析結果と整合的である。しかし、係数に着目すると、勤続年数の1年の上昇は昇給評価を約0.01上昇させるだけであり、実質的な影響力はないと解釈できる。次に、当期成果 ($performance_{i,t}$) については、いずれの分析結果についても負に有意 ($subjective1_{i,t}$ の通常最小二乗法によるモデルのみ $p < 0.05$ 、他のモデルでは $p < 0.001$) であった。この結果は、前章で記述したように、賞与評価が高くなることで主観評価の裁量が小さくなる効果に起因していると考えられる。この傾向は、先ほどの図1を見れば明らかであろう*23。最後に、新卒 ($new_{i,t}$) の係数は、それぞれ0.045 および0.046 と正ではあったが有意ではなかった。

5.4 主観評価の識別力

続いて、時間が経過することで主観評価の識別力が高まるとする仮説2.1、仮説2.2および仮説2.3について検証を行う。上位・下位グループを抜きだし、その主観評価の時系

結果は通常最小二乗法モデルでは、従業員の固定効果をモデルに含めていないため、各従業員の平均的な能力をコントロールできていないという交絡変数の問題が生じているからである。しかし、両者の分析結果を比べると、 $subjective1_{i,t}$ および $subjective2_{i,t}$ に関して、どちらも通常最小二乗法と固定効果モデルにおける $performance_{i,t-1}$ および $promotion_{i,t-1}$ の係数の小さく、この交絡変数の影響は微々たるものである。この分析結果は従業員の固定効果により従業員の平均的な能力をコントロールしたあとでさえも、前期成果や前期昇進といった過去の情報が主観評価の決定に十分な影響を与えていることを示している。

*22 固定効果モデル ($panellinear$) において、勤続年数 ($career_{i,t}$) および新卒 ($new_{i,t}$) の変数を含めていないのは、完全な多重共線性を避けるためである。

*23 なお、当期成果 ($performance_{i,t}$) の係数について、 $subjective2_{i,t}$ を被説明変数とした結果 (通常 OLS で-0.299、固定効果モデルで-0.354) の方が、 $subjective1_{i,t}$ を被説明変数にした結果 (通常 OLS で-0.039、固定効果モデルで-0.100) よりも、その絶対値が大きい。この結果は、 $subjective1_{i,t}$ の反映する情報内容と $performance_{i,t}$ が反映する情報内容が重複しているためであると考えられる。この効果は、 $subjective1_{i,t}$ を被説明変数としたとき、 $performance_{i,t}$ の係数を正の方向に増加させる。したがって、被評価者の能力に関する情報内容が賞与評価と $subjective1_{i,t}$ で重複しているという正の効果も、賞与評価 $subjective1_{i,t}$ が高まることで主観評価の裁量が小さくなるという負の効果も弱めることになる。一方、 $subjective2_{i,t}$ に関しては、その操作化から $performance_{i,t}$ のもつ情報内容との重複は理論上存在しない。実際に、表2の結果でも $subjective1_{i,t}$ と $performance_{i,t}$ は正の相関があったのに対して、 $subjective2_{i,t}$ と $performance_{i,t}$ はほぼ無相関であった。すなわち、主観評価のもつ情報内容と当期成果 ($performance_{i,t}$) のもつ情報内容の重複が、 $subjective1_{i,t}$ と $subjective2_{i,t}$ で非対称であるため、このような係数の違いが生じているのである。

列における推移を確認する前に、サンプル全体での主観評価の推移について確認するため、以下の回帰モデルの結果を表4に記載した。下式においても仮説1.1および仮説1.2の線形モデルと同様に、前期の主観性が当期の主観性に与える影響を、昇給評価の尺度により解釈することが可能である。なお、ここからは前章で述べたように分析対象サンプルを9年間すべての変数が確保できるサンプルに絞ったため、サンプルサイズが2980から1800に減少している。

$$subjective_{i,t} = \alpha_0 + \alpha_1 subjective_{i,t-1} + \alpha_2 subjective_{i,t-1} \times time_{i,t} + \epsilon_{i,t}$$

表4: 主観評価の系列相関

	<i>subjective1_{i,t}</i>		<i>subjective2_{i,t}</i>	
<i>subjective1_{i,t-1}</i>	0.124*** (0.024)	0.132*** (0.026)		
<i>subjective1_{i,t-1} × time_{i,t}</i>		-0.001 (0.002)		
<i>subjective2_{i,t-1}</i>			-0.033 (0.024)	-0.215*** (0.054)
<i>subjective2_{i,t-1} × time_{i,t}</i>				0.035*** (0.009)
Constant	2.231*** (0.063)	2.230*** (0.063)	0.036* (0.015)	0.037* (0.015)
N	1,800	1,800	1,800	1,800
Adj R ²	0.014	0.014	0.000	0.008

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, ()内の数値は標準誤差を表している。*time_{i,t}*は $time_{i,t} = year_{i,t} - 2004$ と計算されており、分析対象年度を1年目とした年度を表す変数である。推定方法は通常最小二乗法である。

表 4 における $subjective1_{i,t}$ の結果によると、 $subjective1_{i,t}$ には正の系列相関が見られ、前期に高い主観評価を受けた従業員は、当期にも高い主観評価を受ける傾向がうかがえる。仮説 2.1 によれば能力の高い従業員と低い従業員の差は、時間経過により大きくなると予想しているため、この主観評価の系列相関は時間経過により強まると予測される。しかし、 $subjective1_{i,t-1}$ と $time$ の交互作用項は-0.001 と非常に小さく、かつ有意でもないため、この系列相関が時間経過によって強まるという予想は支持されないことがわかる。一方で、 $subjective2$ の結果については、 $subjective2_{i,t-1}$ と $time$ の交互作用項は 0.035 であり、有意に正であり予測が支持される一方で、正の系列相関が見られない。さらに、もっとも絶対値の大きい係数でさえ、 $subjective2_{i,t}$ における $subjective2_{i,t-1}$ の-0.215 であり、従業員の昇給評価に実質的な影響を与えているとは言い難い。これらの一貫しない分析結果は、これらの分析対象がサンプル全体であり、能力が高いわけでもなく低いわけでもないゆえに識別が難しい中位のグループがサンプルの大半を占めるからだと思われる。実際、評価者である上司は説得的な根拠を確保することが難しい場合、評価の差をつけることを回避することが知られている (Bol 2010)。そこで、以下の分析ではサンプルから主観評価の上位もしくは下位 10 パーセントのサンプルのみを抜き出し、両者の評価の推移について比較を行う。上位と下位のみを抜き出し、その識別力について調べることで、最低でも上位と下位を客観評価および主観評価が区別できるのかどうかについて検証を行うことが分析の目的である。なお、前章で説明したように、評価の上位および下位の選定は、従業員ごとの評価の平均値を基準としている。

図 2 に客観評価 ($performance_{i,t}$)*²⁴の、図 3 および図 4 に主観評価の上位および下位グループの評価推移をプロットした。図中の丸印はグループ内における評価の平均値を表しており、棒 (エラーバー) はグループ内における評価の標準偏差を表している。図 2 によると、2005 年度を除いて、上位グループの方が下位グループよりも平均値が高いことが分かる*²⁵。一方、標準偏差に関しては、図 2 の全年度において上位と下位のエラー

*²⁴ ここでは賞与評価 ($performance_{i,t}$) を客観評価の代理変数として扱っている。これは、賞与評価が当期成果を評価対象としており、上司の主観性が入る余地が少ないため、相対的に客観的なものとして解釈できるからである。

*²⁵ 2005 年度に平均値が逆転している理由について、A 社では、2004 年度からこの人事システムを導入したため、導入して年月の浅い 2005 年前後では成果評価が安定していなかったためと考えられる。実際、

バーが重なっている。この結果は、上位と下位の客観評価を比べた場合、確かに上位グループの方が評価は高いものの、その評価の差は標準偏差 1 個分ずれただけで、どちらのグループの評価なのか区別できなくなる程度の差であることが分かる。一方、図 3 および図 4 によると、主観評価では 2009 年以降ほとんどエラーバーが重なっていない。この結果は、分析対象の後半年度においては、例え標準偏差 1 個分評価がずれたとしても、その評価がどちらのグループの評価であるか、主観評価では識別できることを示している。この識別できるという効果は、(1)2007 年度以降も上位と下位のグループの平均値の差が時間とともに拡大する、および (2) 上位グループの標準偏差が時間とともに小さくなる、という二つの影響に起因しており、これらの傾向は本研究で設定した仮説と整合的である。

エラーバーが重ならないというこの基準は、どの程度の識別力があるのかを考えてみる。各グループの評価が正規分布にしたがうと仮定した場合、*High* グループの評価が平均値から標準偏差 1 個分下振れする確率は、0.159 である (正規分布は左右対称なので、*Low* グループの評価が標準偏差 1 個分上振れする確率も同様)。標準偏差 1 個分という基準をもちいて、サンプルを *High* グループと *Low* グループに分類することを考えた場合*²⁶、全サンプルのうち $0.841(= 1/2 \times (1 - 0.159) + 1/2 \times (1 - 0.159))$ は正しいグループに分類できることになる。しかし、この分類手順は *High* にも、*Low* にも分類されないサンプルが出てきてしまうため、次のように解釈を変えよう。先ほどと同様にサンプルを *High* グループと *Low* グループに分類するという手続きを考え、今回はグループの分類基準となる閾値 b を設定するという状況を考える。このとき、「*High* グループの平均 - *High* グループの標準偏差 $> b > Low$ グループの平均 + *Low* グループの標準偏差」という不等式*²⁷が成立するように b を設定できれば、84.1% 以上の確率で正しく分類できることになる。

人事システムを導入して 3 年たった 2007 年度以降は上位グループの評価は 3、下位グループの評価は 2 で安定している。

*²⁶ 具体的には、そのサンプルの評価が「*High* グループの平均 - *High* グループの標準偏差」よりも大きい場合には *High* グループに分類する、「*Low* グループの平均 + *Low* グループの標準偏差」よりも小さい場合には *Low* グループに分類するという手続きを指している

*²⁷ 図 3 および図 4 より、後半期間における主観評価の各グループのエラーバーが重なっていない期間に関しては、「*High* グループの平均 - *High* グループの標準偏差 $> Low$ グループの平均 + *Low* グループの標準偏差」という不等式が成立する。

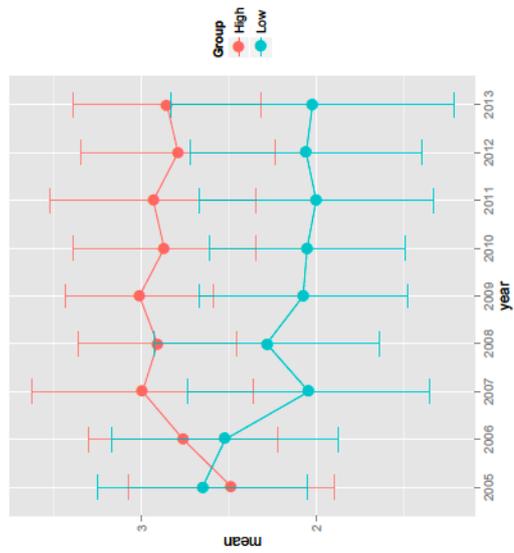


図 2: 客観評価 (*performance*) の推移

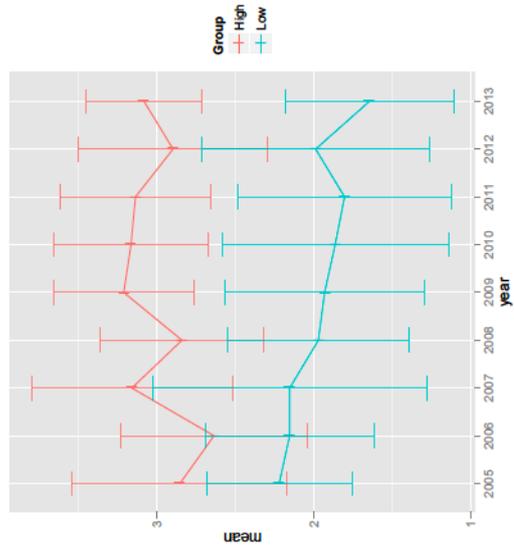


図 3: 主観評価 (*subjective1*) の推移

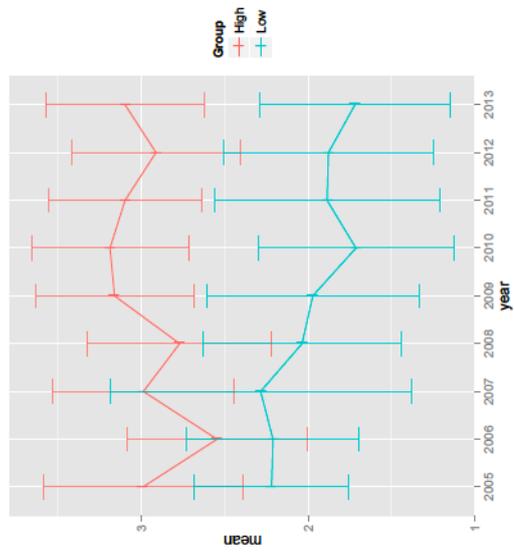


図 4: 主観評価 (*subjective2*) の推移

表 5: 主観評価の識別力

	$mean_{t,j}$			$1/sd_{t,j}$			$mean \times 1/sd_{t,j}$		
	subjective1	subjective2	performance	subjective1	subjective2	performance	subjective1	subjective2	performance
$time_{t,j}$	-0.060** (0.019)	-0.069** (0.020)	-0.070** (0.020)	-0.039 (0.039)	-0.031 (0.029)	-0.025 (0.028)	-0.186 (0.110)	-0.182* (0.077)	-0.165 (0.091)
$Height_{t,j}$	0.563** (0.152)	0.458* (0.161)	0.186 (0.160)	-0.386 (0.312)	-0.118 (0.232)	0.167 (0.222)	-0.133 (0.877)	0.501 (0.613)	0.824 (0.722)
$time_{t,j} \times Height_{t,j}$	0.093** (0.027)	0.105** (0.029)	0.094** (0.028)	0.139* (0.055)	0.084 (0.041)	0.033 (0.039)	0.551** (0.156)	0.409** (0.109)	0.229 (0.128)
Constant	2.270*** (0.108)	2.334*** (0.114)	2.541*** (0.113)	1.811*** (0.221)	1.819*** (0.164)	1.674*** (0.157)	4.122*** (0.620)	4.220*** (0.433)	4.223*** (0.511)
N	18	18	18	18	18	18	18	18	18
Adjusted R ²	0.930	0.917	0.842	0.353	0.357	0.340	0.758	0.846	0.680

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, () 内の数値は標準誤差を表している。変数の添え字については、 t がサンプルの年度、 j がサンプルの所属グループを表している。したがって、サンプルサイズは 2 グループ (上位および下位) \times 9 年 (2005 年から 2013 年) の 18 である。上位 (および下位) グループの抽出方法は分析方法で記したように、従業員ごとに 9 年間の評価を平均し、平均評価の高い上位 (もしくは下位) 10% を抽出するという方法をとっている。そのため、 $subjective1$ の分析では、 $subjective1$ の高い上位 (もしくは低い下位) 10% の従業員が、 $subjective2$ の分析では、 $subjective2$ の高い上位 (もしくは低い下位) 10% の従業員が抽出されており、各分析で抽出サンプルが異なる。なお、ここでの $performance$ は賞与評価 ($performance$) を対象として、同様の分析を行った結果である。

しかし、客観評価では、いずれの期間においてもエラーバーが重なっており、不等式が成立するように、閾値 b を設定することはできない。それならば、客観評価では標準偏差いくつ分ならば両グループのエラーバーが重ならないのか。最終年度における *High* グループの平均値は 2.85、標準偏差は 0.533 であり、*Low* グループの平均値は 2.021、標準偏差は 0.809 なので、標準偏差 0.62 個分ならば、最終年度において両グループのエラーバーは重ならない*28。正規分布を仮定した場合、*High* グループの客観評価が平均値から標準偏差 0.62 個分下振れする確率は、0.268 である。この標準偏差 0.62 個分という基準のもとで、正しいグループに分類されるサンプルの割合を計算すると 73.2% となる。主観評価における分類では、最低でも 84.1% のサンプルが正しく分類されることが保証されていたことを考えると、この値は決して高いとはいえないだろう。

表 5 において、仮説 2.1、仮説 2.2 および仮説 2.3 の検証を行った。まず、仮説 2.1 は、主観評価の推移に関する仮説であり、時間が経過することで、能力の高い従業員と低い従業員の主観評価の差が開いていくというものであった。仮説 2.1 の検証結果について、 $mean_{t,j}$ を被説明変数としたときの $time_{t,j}$ と $High_{t,j}$ の交互作用項はそれぞれ、*subjective1* において 0.093、*subjective2* において 0.105 であり、いずれも正かつ有意 ($p < 0.01$) であり、仮説 2.1 が支持されている。ただし、客観評価である *performance* についても同様に $time_{t,j}$ と $High_{t,j}$ の交互作用項は有意 ($p < 0.01$) に正であった。この結果は 2005 年度において、前述したように上位と下位の平均値が逆転していることに起因している可能性がある。次に、仮説 2.2 は主観評価の精度向上に関する仮説であり、時間が経過することで、主観評価の精度が高まるというものであった。仮説 2.2 の検証結果については、 $1/sd_{t,j}$ を被説明変数とする回帰モデルにおいて、 $time_{t,j}$ と $High_{t,j}$ の交互作用項は、*subjective1* において 0.139 ($p < 0.05$)、*subjective2* において 0.084 であり、*subjective2* に関しては有意ではなかったものの、いずれも正かつ *performance* における交互作用項の係数である 0.033 よりも大きかった。この結果は、*High* グループにおいては仮説 2.2 が支持されたものの、*Low* グループについては仮説 2.2 が支持されな

*28 標準偏差 a 個分ならば両グループを識別できるとしたとき、以下の不等式が成立しなければならない。
 $High$ グループの平均 $- a \times High$ グループの標準偏差 $\geq Low$ グループの平均 $+ a \times Low$ グループの標準偏差

かったことを意味する*²⁹。したがって、仮説 2.2 は *subjective1* の *High* グループのみにおいて部分的に支持されたといえる。最後に、識別力に関する仮説である、仮説 2.3 の結果については、 $mean_{t,j} \times 1/sd_{t,j}$ を被説明変数としたときの $time_{t,j}$ と $High_{t,j}$ の交互作用項はそれぞれ、*subjective1* において 0.551、*subjective2* において 0.409 であり、いずれも正かつ有意 ($p < 0.01$) であった。また、*performance* における結果については、0.229 と正であったものの有意ではなかった。この主観評価と客観評価における分析結果の違いは、両者の識別力の違いを示していると考えられる。すなわち、客観評価は過去の情報を取り込まず現在の成果を客観的に測定する。そのため、2007 年度以降については各グループにおいて評価の平均値も標準偏差も一定になり、時間経過の影響を受けない。一方で、主観評価は過去の情報を取り入れていくために、人事システムが安定する 2007 年度以降においても従業員に対する評価は更新され両グループの平均値の差は拡大、かつ評価が正確になることでグループ内の標準偏差は小さくなっていく*³⁰。本研究のこれらの結果は、長期的インセンティブにおいて重要となる従業員を識別するという目的においては、過去の情報を取り込み従業員の能力を識別できる主観評価の方が客観評価よりも優れているという主張を支持するものである。

*²⁹ *Low* グループにおいて、仮説 2.2 が支持されなかった理由は生存バイアスの影響によるものだと考えられる。本研究の分析では、*High* グループと *Low* グループを比較するという分析の目的上、サンプルを分析期間の 9 年間すべてにおいて変数がそろったサンプルに限定した。この操作は *High* グループの分析には影響を与えない一方で、*Low* グループの標準偏差を本来よりも過大に推定するという影響を与える。なぜならば、*Low* グループにおいて評価が標準偏差が小さくなったサンプル、すなわちほとんどの期間で低い評価を受けている従業員は離職してしまうため、分析対象から外れてしまうからである。

*³⁰ 仮説 2.2 における *subjective2* の結果では、*High* グループの標準偏差が小さくなるという傾向は見られたものの、その結果は有意ではなかった。それゆえ、仮説 2.3 における *subjective2* の有意性検定は、仮説 2.1 の影響のみを考慮すれば成立するのではないかという疑問が生じるが、この解釈は正しくない。実際、仮説 2.1 および仮説 2.2 の検証モデルにおいて、客観評価である *performance* の有意性検定の結果は *subjective2* の結果と、ほぼ同様であるにもかかわらず、仮説 2.3 の検証モデルにおいて、*performance* の $time_{t,j}$ と $High_{t,j}$ の交互作用項は有意ではない。この結果は、図 2 および表 5 から明らかなように、客観評価 (*performance*) は時間が経過することで標準偏差が小さくなるという効果が、主観評価に比べて小さいからである。一方、*subjective2* の標準偏差は、図 3 から分かるように小さくなっていることが確認できる。実際、仮説 2.2 の検証モデルにおける $time_{t,j}$ と $High_{t,j}$ の係数は、*subjective2* において 0.041 であり、*performance* において 0.23 であった。*subjective2* における交互作用項の係数は、有意ではないものの、客観評価 (*performance*) に比べて大きい。こうした仮説 2.2 の検証モデルにおける *subjective2* と *performance* の結果の違いが、仮説 2.3 における非対称な検証結果をもたらしている。したがって、仮説 2.3 の有意性検定を考えるうえでは、あくまで仮説 2.1 における効果と仮説 2.2 における効果の双方を考慮する必要がある。

表 6: 上位および下位グループの当期成果の傾向

	<i>performance</i> _{<i>i,t</i>}	
	[<i>subjective1</i> _{<i>i,t</i>}]	[<i>subjective2</i> _{<i>i,t</i>}]
<i>time</i> _{<i>i,t</i>}	-0.002 (0.018)	-0.011 (0.018)
<i>High</i> _{<i>i,t</i>}	0.907*** (0.144)	0.613*** (0.141)
<i>time</i> _{<i>i,t</i>} × <i>High</i> _{<i>i,t</i>}	0.015 (0.026)	0.018 (0.025)
Constant	4.022*** (0.103)	4.167*** (0.100)
N	369	360
Adj R ²	0.373	0.244

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, () 内の数値は標準誤差を表している。1 行目は *subjective1*_{*i,t*} が上位もしくは下位 10% のグループ、2 行目は *subjective2*_{*i,t*} が上位もしくは下位 10% のグループであり、被説明変数はどちらも *performance*_{*i,t*} である。なお、どちらのグループについても上位もしくは下位 10% を抽出したにもかかわらず、1 行目と 2 行目でサンプルサイズが異なるのは、*subjective1*_{*i,t*} のグループに *subjective1*_{*i,t*} の値が同じサンプルが複数含まれるためである。本節の分析では、このように同じ値をとるサンプルを除外していない。なぜなら、グループ間でサンプルサイズを合わせるため、サンプルを恣意的に除外すればグループ内平均の値が歪められてしまう恐れがあるからである。また、本節の分析では、グループ内平均および標準偏差に関する時系列通じた変化に焦点を当てているため、このようなサンプルサイズの違いは結論に重大な影響を与えない。

本章の最後に、主観評価の識別力に関する分析結果が交絡変数の影響によるものではないことを示す。本研究では、サンプルから上位グループおよび下位グループを抽出し、各グループの時系列推移をプロットし、その後、統計的検定を行うという方法を採用した。この分析方法は、評価の識別力を可視化し、結果を解釈しやすくするという点では優れた

方法である。しかし、単純な線形モデルによる検証のように、変数を追加することで、交絡変数の影響を統制することはできない。そこで、全グループの記述統計を確認し、本研究の分析結果が交絡変数の影響によるものであるという積極的な証拠がないことを提示する。

まず、上位・下位グループそれぞれの要約統計量および相関係数表を付録に記載した。相関係数表において、前年成果および前年昇進と主観評価の相関係数が下位グループの方が強いという傾向が見られた。これは低い評価をつけるためには説得的な根拠が必要であるという Bol (2010) の主張と整合的である。しかし、上位グループと下位グループで相関係数が著しく逆転するなど、両グループにおいて顕著な相関構造の差は見られなかった。したがって、両グループの識別力に注目した本研究の分析結果が交絡変数の影響を受けているという積極的な証拠は、少なくとも分析に使用した変数の相関構造には見られなかった。

一方で、要約統計量の結果については記述する必要がある箇所が見られた。それは、両グループの当期成果の差である。当然ではあるが、上位グループの当期成果の平均値は下位グループのそれよりも高い。この傾向は、上位グループにおける主観評価の標準偏差が小さくなるという傾向が、主観評価の精度向上ではなく、上位グループに所属する従業員の学習もしくは成長によるものかもしれないという懸念を生じさせる。すなわち、上位グループの成果がもともと高いにもかかわらず、仮に従業員が学習・成長により能力を向上させていた場合、上位グループの賞与評価が高まることで、昇給評価における上司の裁量の余地の程度が低くなってしまう。このような傾向が見られる場合、本研究における分析では、時間が経過するにつれ主観評価の標準偏差が小さくなるという傾向が見られることになる。この主張は、本研究の主張する主観評価の精度向上の影響とは異なる。そのため、上位グループの標準偏差が時間とともに小さくなるという傾向が、従業員の学習・成長による成果の向上によるものではないことを示す必要がある。そこで、当期成果 ($performance_{i,t}$) を被説明変数、時間 ($time_{i,t}$) と *High* グループに関するダミー変数 ($High_{i,t}$) の交互作用項を説明変数とした回帰分析を行い、結果を表 6 に記載した。表 6 によると、上位グループおよび下位グループの双方で時間の経過とともに当期成果が上

がっている（下がっている）というトレンドは見られなかった。この結果は、上位グループに所属する従業員が学習・成長することにより、能力を向上させているという主張を支持しない。以上の議論および分析結果から、本研究の識別力に関する分析結果が交絡変数の影響によるものだという積極的な証拠を得ることはできなかった。

6 追加分析

本章では、これまでの分析結果を踏まえて、追加分析を行う。まず、追加分析の必要性について説明する。その後、追加分析の分析方法・手順について記述し、最後に、追加分析の結果について記載する。

6.1 追加分析の必要性

これまでの分析から、(1) 主観評価が過去の評価と相関していること、(2) 時間が経つことで、上位グループと下位グループの主観評価が識別しやすくなることが明らかになった。しかし、前章までの分析では、本研究の結果が、上司のえこひいきやバイアスによるものであることを否定できない。

例えば、主観評価が過去の評価と相関しているという分析結果は、上司が適切に過去の情報を取り入れたというよりも、むしろ評価に悩んだ上司が過度に過去の評価に依存した可能性もある。先行研究では、主観評価が前期成果や前期昇進といったベンチマークを参照する理由として、(1) シグナル、(2) 評価に伴う負担、(3) ハロー効果の三つの理由があげられている。まず、評価者が前年度の成果や勤続年数といった基準を参照する理由の一つは、これらのベンチマークが従業員の能力を反映しているからである。前年度の成果は従業員の能力についての関数になっていると考えられ、勤続年数が長く経験があるほど能力が高いであろうと推察される。実際、多くの分析的研究では、そのような仮定が設定されている (ex. Gibbons and Waldman 1999)。すなわち、評価者は能力の推定精度を改善するために、能力を反映しているこれらのベンチマークを参照すると考えられる。本研究における主張は、この観点に立脚している。次に、これらのベンチマークとつじつまの合わない評価をつけた場合、根拠もなく低い評価をつけたことで従業員からの不満が高まる。そのため、こうした評価に不満をもつ従業員を説得するための時間的の浪費、もしくは従業員が不満をもつことによる心理的ストレスなどを評価者は負担しなければならなくなる (Bol et al. 2016)。こうした評価に伴う負担を避けるため、評価者これらのベンチ

マークと「つじつまが合う」ように評価を行うと考えられる。この観点によれば、本来、上司は自身が決定した主観評価よりも正確に従業員の能力を識別できるにもかかわらず、これらの評価コストの存在のため自分のもっている情報を不十分にしか活用できていないことになる。最後に、評価者は認知的なバイアスによりこれらのベンチマークに過度に依存していることも考えられる。例えば、ハロー効果やアンカリング効果のため、評価者は参照すべき、もしくは参照しやすい複数のベンチマークを過度に利用する可能性がある (Ittner et al 2003; Thorsteinson et al 2008)。主観評価が過去の情報に依存しているという仮説 1.1 および仮説 1.2 の分析結果に関して、評価に伴う負担、およびハロー効果の影響は、それが従業員の能力を反映している限りにおいて本研究の結論に影響を与えない。しかし、これらの影響のため、従業員の能力を過大（もしくは過少）に評価してしまうことで、上司の主観評価が従業員の能力を反映しているとは言い難いものになってしまっている可能性がある。

また、時間が経つことで上位・下位の識別がしやすくなるという仮説 2.1、仮説 2.2 および仮説 2.3 の結果についても、従業員の能力というよりも、単に上司の従業員に対する好みが増進していく様子を反映している可能性がある。例えば、えこひいきに関する代表的な理論モデルである Prendergast(1996) では、上司の従業員に対する好みに不確実性の仮定を置いている。本研究の採用したアーカイバル分析では、時間経過により主観評価の識別力が増すという傾向がえこひいきによるものなのか特定することはできない。しかし、この点においても、先ほどと同様に、本研究の結論において重要なのは、えこひいきの影響が上司の主観評価と従業員の能力の対応関係を大幅に損ねるほど強いのかという点である。

これらの点に関して、本研究の結論に関する頑健性を高めるため、主観評価が従業員の能力を反映しているかについて追加分析を行う。追加分析では、高い主観評価を受けた従業員が、その後に高い成果を挙げたかどうかを検証することで、上司の決定した主観評価が適切であったかどうかをデータから事後的に判断するという方法を採用する。

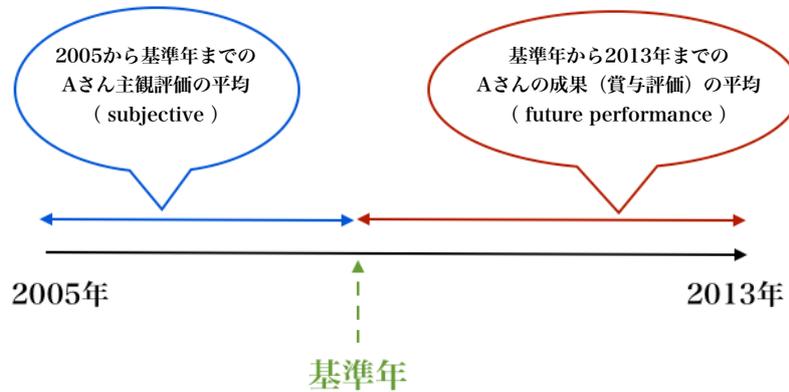


図 5: 追加分析の方法

6.2 分析方法

次の手順で追加分析を行う^{*31}。前半期間に高い主観評価を受けた従業員が後半期間に高い成果を挙げるかどうかを検証することが、追加分析の目的である。被説明変数である成果 (*future performance*) の代理変数には、賞与評価を採用した。そのため、手順 5 における回帰分析では、将来の平均賞与評価に与える影響として係数を解釈できる。

1. 2004-2013 年の間に、基準年を設定
2. 基準年以前の前半期間サンプルと、基準年より後の後半期間サンプルにサンプルを分割
3. 前半期間のサンプルで、主観性 (*subjective*) を従業員ごとに平均値を算出
4. 後半期間のサンプルで、成果 (*future performance*) を従業員ごとに平均値を算出
5. 手順 3 の前半期間における平均主観評価を説明変数、手順 4 の後半期間における平均成果を被説明変数として回帰分析を行う。

^{*31} 分析の意図を分かりやすくするため、図 5 に分析のイメージ図を示した。

なお、当然であるが分析期間の間に昇格している従業員が存在する。これらの従業員については、昇格後の後半期間においては、昇格したあとの等級をベースに評価されるため、評価基準が厳しくなる可能性がある。そこで、この等級ごとの違いの影響をコントロールするため、後半期間の平均等級をコントロールした。^{*32}なお、以下で述べる分析結果は等級のコントロール変数を取り除く、前半期間の当期成果平均をコントロール変数として追加するという操作に対して頑健であった。

6.3 分析結果

表 7: 主観評価の将来成果の予測

	<i>future performance_i</i>			
	2008	2009	2010	2011
<i>subjective_i</i>	0.617*** (0.114)	0.707*** (0.114)	0.684*** (0.118)	0.746*** (0.126)
<i>grade_i</i>	-0.029 (0.032)	-0.038 (0.032)	-0.035 (0.033)	-0.029 (0.037)
Constant	3.059*** (0.288)	2.872*** (0.285)	2.928*** (0.297)	2.723*** (0.320)
N	200	200	200	200
Adj R ²	0.124	0.158	0.139	0.147

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, () 内の数値は標準誤差を表している。*future performance_i* は基準年から 2013 年までの *performance_{i,t}* について従業員ごとの平均値をとったものである。同様に *past subjective* は、2005 年から基準年までの *subjective_{1,i,t}* に関する従業員ごとの平均値である。*grade* はコントロール変数であり、基準年から 2013 年までの *grade_{i,t}* に関する従業員ごとの平均値をとっている。

^{*32} 前章での分析では、等級 (*grade*) をダミー変数でコントロールしたが、ここでは等級 (*grade*) を一つの連続変数としてコントロールする。本章での分析方法では、分析に使用する変数は期間中の平均値をとっているため、変数自体が複数期間に渡っている。そのため、等級も期間を通じて一定ではないため、他の変数と同様に期間中の平均値をとっている。

表 8: 主観評価の将来成果の予測

	<i>future performance_i</i>			
	2008	2009	2010	2011
<i>subjective_i</i>	0.371** (0.138)	0.487** (0.146)	0.516** (0.156)	0.683*** (0.167)
<i>grade_i</i>	-0.001 (0.034)	-0.008 (0.034)	-0.010 (0.035)	-0.012 (0.038)
Constant	4.492*** (0.168)	4.520*** (0.170)	4.536*** (0.175)	4.518*** (0.192)
N	200	200	200	200
Adjusted R ²	0.029	0.048	0.046	0.074

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, () 内の数値は標準誤差を表している。*future performance_i* は基準年から 2013 年までの *performance_{i,t}* について従業員ごとの平均値をとったものである。同様に *past subjective* は、2005 年から基準年までの *subjective2_{i,t}* に関する従業員ごとの平均値である。*grade* はコントロール変数であり、基準年から 2013 年までの *grade_{i,t}* に関する従業員ごとの平均値をとっている。

表 6 に主観評価の変数に *subjective1* を採用したときの追加分析、表 7 に主観評価の変数に *subjective2* を採用したとき追加分析の結果をそれぞれ記した。まず、表 6 によると、前半期間における主観評価の平均値である *subjective_i* の係数は、それぞれ 0.617, 0.707, 0.684, 0.746 であった。これらの係数はいずれも有意 ($p < 0.001$) であり、基準年度によらず前半期間に高い主観評価を受けた従業員は、後半期間においても平均的に高い成果をあげていることがわかる。また、係数の大きさに着目すると、いずれの年度に依っても係数は約 0.7 前後である。表 7 の結果でも、*subjective_i* の係数はそれぞれ 0.371, 0.487, 0.516, 0.683 と小さくはなるものの、いずれの基準年度の分析においても有意 ($p < 0.01$) であった。どちらの分析についても、基準年度が遅い分析では、主観評価に使用できる情報が多く、予測する期間が短いため係数が大きくなる傾向が見られる。しかし、いずれの分析結果においても前半期間の主観評価が 1 高ければ、後半期間の成果はそ

それぞれの分析で 0.7、0.5 程度高くなるという傾向が見られる。成果 ($performance_i$) が年二回の 7 段階評価であったことを考えると、0.7 や 0.5 程度という影響は十分に大きいといえるだろう。なお、等級 ($grade_i$) の係数については予測通りすべて負であった。しかし、等級 ($grade_i$) が 9 段階であることを考慮すれば、いずれの係数も大きいとはいえず、また有意でもなかった。

7 結論

本論文の締めにあたる本章では、これまでの議論を踏まえて、本研究の要約、貢献および限界について記述する。まず、本研究における分析結果の要約を行い、分析結果から導かれる本研究の主張について再確認を行う。続いて、本研究の貢献について明示したのちに、本研究の限界について記述を行う。

7.1 分析結果の要約

本研究の目的は、長期的インセンティブにおける主観評価の役割を明らかにすることであった。本研究では、主観評価が過去の情報を蓄積することで、時間とともに能力の高い従業員と低い従業員を識別しやすくなるという性質に注目し、日本の上場企業一社のデータをもちいて分析を行った。分析の結果、(1) 前期成果および前期昇進が高ければ、当年度の主観評価が高くなること、(2) 時間が経つことにより、主観評価の識別力が増加することの2点が明らかになった。また、これらの分析結果が認知バイアスやえこひいきにより不適切に歪められた主観評価である可能性を排除するため追加分析を行った。追加分析の結果、分析の前半期間における主観評価の平均値が高ければ、分析の後半期間における成果の平均値も高くなっていた。この結果は、A社における主観評価が、少なくとも将来成果を予測する程度には適切であることを担保している。

本研究のこれらの分析結果については、単に「有意」であるだけではなく、係数の大きさに注目した解釈が可能である。いずれの分析結果においても、本研究で注目した説明変数は、解釈のうえでは実務上ある程度意義のある影響力を被説明変数に対してもっていた。したがって、係数の有意性検定だけではなく、係数の解釈から導かれる説明変数のもつ影響力の大きさという観点からも本研究の分析結果は担保されている。

7.2 本研究の貢献

まず、本研究は主観評価に関する一連の研究に貢献がある。主観評価に関する研究は、えこひいき (Prendergast 1994)、認知バイアス (Lipe and Salterio 2000)、特定の指標のみ重視する傾向 (Ittner et al 2003)、前年度の客観評価に左右される傾向 (Woods2009)、ハロー効果 (Thorsteinson 2004)、寛大化・中心化バイアス (Moers2005; Bol 2009) など、その負の側面を捉える研究が多いものの、主観評価が果たす役割について取り扱った経験的研究は多くない。しかし、現実の企業では様々な場面で主観評価が多用されており、主観評価に関する一連の研究は、なぜ主観評価がこれほどまでに現実で採用されているのか、その役割について十分な経験的証拠を提示できているとは言い難い。それゆえ、過去の情報を取り入れることで、能力の高い従業員と低い従業員を識別できるという主観評価の役割について明らかにできた点は本研究の貢献の一つである。

次に、本研究は主観評価の時系列上の特徴について明らかにした点についても貢献がある。主観評価の時系列上の特徴について明らかにするためには、実験室実験や質問票調査といった方法では時間的制約のため困難である。そのため、企業のアーカイバルデータを入手する必要があるものの、主観評価などの業績評価に関する研究は、企業内のデータが必要となるため入手が難しい (Moers 2007; Lazear and Oyer 2013)。それゆえ、主観評価が時間経過によりどのように推移するのか、その推移のパターンは企業において何らかの役割を果たすのかについて十分に明らかになってはいなかった。本研究では、10年間という比較的長期のデータを入手できたこと、および従業員の能力を識別するという主観評価の役割に注目したことにより、主観評価が企業においてどのように推移しているのかを可視化できた。時間が経過することで従業員に対する識別力が向上するという主観評価の特徴は、客観評価には存在しない時系列上の性質であり、長期的評価には主観評価が適しているという一連の研究 (ex. Gibbs 2003) に対して貢献がある。

最後に、本研究は長期的インセンティブと業績評価に関する一連の研究に対しても貢献がある。長期的インセンティブと業績評価に関する研究について、現実には上司の査定などの主観評価が昇給や昇進に取り入れられているにもかかわらず、これまで主観評価が長

期的インセンティブに果たす役割について十分に明らかにできていなかった。これらの研究の中でも、Grabner and Moers(2013) は昇進における職務配置という側面に注目し、昇進に職務内容の変更が伴う場合は主観評価が適していることを明らかにした。しかし、Grabner and Moers(2013) の結果だけでは、昇給や昇格といった職務内容の変更を伴わない長期的インセンティブにおいて、上司の主観評価が利用されるのはなぜかという疑問には答えることができなかった。本研究は、報酬の支払いが長期に渡る長期的インセンティブにおいては優秀な従業員が誰であるかを判断することが重要であり、この点において過去の情報を取り入れ、時間が経過することで従業員の能力に対する識別力が向上する主観評価が適していると主張する。本研究のこの結果および主張は、これまでの先行研究では説明できなかった業績評価と長期的インセンティブの関係を説明しているという点において貢献がある。

7.3 本研究の限界

まず、本研究の結果から、主観評価が過去の評価を取り入れることで識別力が向上するため、主観評価が長期的インセンティブに適していることが明らかになった。しかし、過去の情報を取り入れるだけならば、主観評価である必要はなく、過去の客観評価をも取り入れた公式的な評価に基づき昇給や昇格に使用してもよい。実際、本研究の結果はHolmstrom (1999) や Gibbons and Waldman(1999) といった理論モデルにおける従業員の能力推定に関する結果と整合的であるが、彼らの理論モデルでは過去の客観評価より公式的に評価を算出する。現実の企業がこのような過去の評価を取り入れた公式的な評価ではなく、主観評価を採用している理由として評価の柔軟性を残すなど様々な理由が考えられるものの、本研究の結果からはいずれの理由についても明らかにできない。

次に、本研究の結果は業績評価によるモチベーションや、それに伴う成長の影響を十分に考慮できていない。例えば、実際には当時能力が低かったにもかかわらず、高い主観評価をつけられたために、従業員が努力し実際にその後能力が高くなるという場合である。本研究の採用した従業員ごとに主観評価を平均するという方法では、「当時は能力の低かった」従業員を選別することができないため、このような影響を統制することは

できていない。また、従業員の学習や成長といった側面も積極的には考慮していない。Campbell (2011) によれば、タイトなモニタリングはコントロールを強める反面、従業員が現場で「実験」し、学習する機会を無くしてしまう。これらの影響が十分に大きいならば、客観評価の差が時間とともに増加していく傾向はこの学習や成長およびモチベーションに関する従業員ごとの違いを反映している可能性がある^{*33}。しかし、学習やモチベーションの影響のみでは、上位グループの主観評価の標準偏差が時間とともに減少している、という分析結果を説明できない。ゆえに、上位グループと下位グループの主観評価が識別しやすくなるという本研究の分析結果が、学習やモチベーションの影響を含んでいる可能性は否定できないものの、主観評価の識別力向上という概念なしに本研究結果の全てを説明することはできない。

また、本研究の分析では評価者と被評価者との関係性が主観評価の識別力に与える影響については検証できてはいない。Bol (2011) によれば、職場を共有しているなど評価者と被評価者の関係が近く情報収集のコストが低ければ、主観評価を甘めにつける、もしくは極端に差をつけなくするといった傾向^{*34}が弱まる。このように、評価者と被評価者の関係性が、主観評価の識別力に影響を与えることは十分考えられる。しかし、データ上の制約から本研究ではこのような評価者と被評価者の関係が識別力に与える影響については検証できておらず、将来の研究課題である。本研究の分析結果から主張できることは、あくまで組織全体として平均的に主観評価の識別力が向上していくということである。これは、被評価者の能力に対する評判が組織内で共有されていくために生じると考えられる。ただし、繰り返しになるが、本研究のこの分析結果は評価者と被評価者の関係が主観評価に与える影響を含んでいる点には留意が必要である。

最後に、本研究の結果を解釈する際には、生存バイアスについて考慮する必要がある。本研究では上位グループと下位グループの比較を適切に行うため、後半の分析を分析対象期間のすべてについてデータが存在するサンプルを分析対象にした。そのため、解雇や離職した従業員を分析対象にできていない。分析結果について、下位グループの標準誤差が

^{*33} time の変数をコントロールしているため、従業員全体の平均的な学習の影響はコントロールしている。コントロールできていないのは、従業員ごとに学習能力に違いがある場合である。

^{*34} いわゆる、寛大化・中心化バイアスである (see Saal et al. 1980)。

時間経過とともに低下しなかったことは、この生存バイアスに起因する可能性がある。

参考文献

- Abernethy, M. A., J. Bouwens, and L. Van Lent, 2004, Determinants of Control System Design in Divisionalized Firms, *The Accounting Review* 79, 545–570.
- Anthony, R. N., and V. Govindarajan, 2007, *Management Control Systems*, 12th edition (McGraw-Hill).
- Baik, B., Evans J. H., Kim K., and Yanadori Y., 2016, White collar incentives, *Accounting, Organizations and Society* 53, 34–49.
- Baker, G., M. Gibbs, and B. Holmstrom, 1994a, The Internal Economics of the Firm—Evidence From Personnel Data, *The Quarterly Journal of Economics* 95, 881–919.
- Baker, G. P., R. Gibbons, and K. J. Murphy, 1994b, Subjective Performance Measures in Optimal Incentive Contracts, *The Quarterly Journal of Economics* 109, 1125–1156.
- Banker, R. D., 2000, An Empirical Investigation of an Incentive Plan that Includes Nonfinancial Performance Measures, *The Accounting Review* 75, 65–92.
- Banker, R. D., and S. M. Datar, 1989, Sensitivity, precision, and linear aggregation of signals for performance evaluation, *Journal of Accounting Research* 27, 21–39.
- Banker, R. D., S. Y. Lee, and G. Potter, 1996, A field study of the impact of a performance-based incentive plan, *Journal of Accounting and Economics* 21, 195–226.
- Becker, S., 1962, Investment in Human Capital: A Theoretical Analysis, *The Journal of Political Economy* 70, 9–49.
- Bertrand, M., and A. Schoar, 2003, Managing with Style: The Effect of Managers on Firm Policies, *The Quarterly Journal of Economics* 118, 1069–1208.
- Bol, J., Kramer S., and Maas V. S., 2016, How control system design affects performance evaluation compression: The role of information accuracy and outcome transparency, *Accounting Organizations and Society* 51, 64–73.

- Bol, J. C., 2008, Subjectivity in Compensation Contracting, *Journal of Accounting Literature* 27, 1–24.
- Bol, J. C., 2011, The Determinants and Performance Effects of Managers' Performance Evaluation Biases, *The Accounting Review* 86, 1549–1575.
- Campbell, D., 2008, Nonfinancial Performance Measures and Promotion-Based Incentives, *Journal of Accounting Research* 46, 297–332.
- Campbell, D, M. J. Epstein, and F. A. Martinez-Jerez, 2011, The Learning Effects of Monitoring, *Accounting Review* 86, 1909–1934.
- Chow, C. W., 1983, The Effects of Job Standard Tightness and Compensation Scheme on Performance: An Exploration of Linkages, *The Accounting Review* 58, 667–685.
- Cichello, M. S., C. E. Fee, C. J. Hadlock, and R. Sonti, 2009, Promotions, Turnover, and Performance Evaluation: Evidence from the Careers of Division Managers, *The Accounting Review* 84, 1119–1143.
- DeVaro, J., and M. Waldman, 2012, The Signaling Role of Promotions: Further Theory and Empirical Evidence, *Journal of Labor Economics* 20, 91–147.
- Dohmen, T. J., 2004, Performance, seniority, and wages: Formal salary systems and individual earnings profiles, *Labour Economics* 11, 741–763.
- Ederhof, M., 2011, Incentive Compensation and Promotion-Based Incentives of Mid-Level Managers: Evidence from a Multinational Corporation, *The Accounting Review* 86, 131–153.
- Feltham, G. A, and J. Xie, 1994, Performance Measure Congruity and Diversity in Multi-Task Principal/Agent Relations, *The Accounting Review* 69, 429–453.
- Frederickson, J. R., 1992, Relative Performance Information: The Effects of Common Uncertainty and Contract Type on Agent Effort, *The Accounting Review* 67, 647–669.
- Gibbons, R., and K. J. Murphy, 1992, Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence, *Journal of Political Economy* 100, 468–

505.

- Gibbons, R., and M. Waldman, 1999, A Theory of Wage and Promotion Dynamics Inside Firms, *The Quarterly Journal of Economics* 114, 1321–1358.
- Gibbons, Robert, and M. Waldman, 2006, Enriching a Theory of Wage and Promotion Dynamics inside Firms, *Journal of Labor Economics* 24, 59–107.
- Gibbs, M., K. A. Merchant, W. A. Van der Stede, and M. E. Vargus, 2004, Determinants and Effects of Subjectivity in Incentives, *The Accounting Review* 79, 409–436.
- Grabner, I., 2014, Incentive system design in creativity-dependent firms, *The Accounting Review* 89, 1729–1750.
- Grabner, I., and F. Moers, 2013, Managers' Choices of Performance Measures in Promotion Decisions: An Analysis of Alternative Job Assignments, *Journal of Accounting Research* 51, 1187–1220.
- Holmström, B., 1979, Moral hazard and observability, *The Bell Journal of Economics* 10, 74–91.
- Holmström, B., 1982, Moral hazard in teams, *The Bell Journal of Economics* 11, 74–91.
- Holmström, B., 1999, Managerial Incentive Problems—A Dynamic Perspective, *Review of Economic Studies* 66, 169–182.
- Holmstrom, B., and P. Milgrom, 1991, Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design, *Journal of Law, Economics, and Organization* 6, 24–52.
- Höppe, F., and F. Moers, 2011, The choice of different types of subjectivity in CEO annual bonus contracts, *The Accounting Review* 86, 2023–2046.
- Ittner, C. D., D. F. Larcker, and M. W. Meyer, 2003, Subjectivity and the Weighting of Performance Measures: Evidence from a Balanced Scorecard, *The Accounting Review* 78, 725–758.

- Jensen, M. C., and W. H. Meckling, 1976, Theory of the firm: Managerial behavior, agency costs and ownership structure, *Journal of Financial Economics* 3, 305–360.
- Kahneman, D., P. Slovic, and A. Tversky, 1974, Judgment under uncertainty: heuristics and biases, *Science* 185, 1124–1131.
- Lambert, R. A., 2001, Contracting theory and accounting, *Journal of Accounting and Economics* 32, 3–87.
- Lazear, E. P., 2004, The Peter Principle: A Theory of Decline, *Journal of Political Economy* 112, S141–S163.
- Lazear, E. P., 2009, Firm - Specific Human Capital : A Skill - Weights Approach, *Journal of Political Economy* 117, 914–940.
- Lazear, E. P., and P. Oyer, 2013, Personnel Economics, In R. Gibbons and P. Milgrom(Eds.), *Handbook of Organizational Economics* 479–519.
- Lazear, E. P., and S. Rosen, 1981, Rank-Order Tournaments as Optimum Labor Contracts, *Journal of Political Economy* 89, 841–864.
- Lipe, M. G., and S. E. Salterio, 2000, The balanced scorecard: Judgmental effects of common and unique performance measures, *The Accounting Review* 75, 283–298.
- Matsumura, E. M., and J. Y. Shin, 2006, An Empirical Analysis of an Incentive Plan with Relative Performance Measures: Evidence from a Postal Service, *The Accounting Review* 81, 533–566.
- Medoff, J. L., and K. G. Abraham, 1980, Experience, Performance, and Earnings, *The Quarterly Journal of Economics* 95, 703–736.
- Merchant, K. A., and J. Manzoni, 1989, The Achievability of Budget Targets in Profit Centers A Field Study, *The Accounting Review* 64, 539–558.
- Milgrom, P., and J. Roberts, 1992, *Economics, Organization, and Management* (Prentice-Hall, Inc.).
- Moers, F., 2005, Discretion and bias in performance evaluation: the impact of diversity and subjectivity, *Accounting, Organizations and Society* 30, 67–80.

- Moers, F., 2006, Performance Measure Properties and Delegation, *The Accounting Review* 81, 897–924.
- Moers, F., 2007, Doing Archival Research in Management Accounting, In C. S. Chapman, A. G. Hopwood and M. D. Shields (Eds.), *Handbook of Management Accounting Research* 1, 399–413.
- Nagar, V., 2002, Delegation and Incentive Compensation, *The Accounting Review* 77, 379–395.
- Prendergast, C., 1992, Career Development and Specific Human Capital Collection, *Journal of the Japanese and International Economies* 6, 207–227.
- Prendergast, C., 1999, The Provision of Incentives in Firms, *Journal of Economic Literature* 37, 7–63.
- Prendergast, C., and R. H. Topel, 1993, Discretion and bias in performance evaluation, *European Economic Review* 37, 355–365.
- Prendergast, C., and R. H. Topel, 1996, Favoritism in Organizations, *Journal of Political Economy* 104, 958–978.
- Saal, F. E., R. G. Downey, and M. A. Lahey, 1980, Rating the ratings: Assessing the psychometric quality of rating data., *Psychological Bulletin* 88, 413–428.
- Sprinkle, G. B., and M. G. Williamson, 2007, Experimental Research in Management Accounting, In C. S. Chapman, A. G. Hopwood and M. D. Shields (Eds.), *Handbook of Management Accounting Research* 1, 415–444.
- Thorsteinson, T. J., Breier J., Atwell A., Hamilton C., and Privette M., 2008, Anchoring effects on performance judgments , *Organizational Behavior and Human Decision Processes* 107, 29–40.
- Woods, A., 2012, Subjective adjustments to objective performance measures: The influence of prior performance, *Accounting Organizations and Society* 37, 403–425.

付録：上位・下位グループごとの記述統計

A. 上位・下位グループごとの要約統計量

表 A1: $subjective1_{i,t}$ の上位 10% グループにおける要約統計量

	平均値	標準偏差	最小値	1Q	2Q	3Q	最大値
$subjective1_{i,t}$	3.00	0.56	1.29	2.46	3.11	3.29	4.46
$subjective2_{i,t}$	0.38	0.57	-1.32	-0.08	0.37	0.68	1.92
$performance_{i,t}$	4.99	0.58	3.00	4.50	5.00	5.50	6.50
$rise_{i,t}$	4.71	0.61	3.00	4.00	5.00	5.00	6.00
$performance_{i,t-1}$	5.00	0.60	3.00	4.50	5.00	5.50	6.50
$promotion_{i,t-1}$	0.17	0.38	0.00	0.00	0.00	0.00	1.00
$career_{i,t}$	13.86	5.08	3.00	10.00	13.00	17.00	30.00
$grade_{i,t}$	4.97	1.19	3.00	4.00	5.00	6.00	8.00

表 A2: $subjective1_{i,t}$ の下位 10% グループにおける要約統計量

	平均値	標準偏差	最小値	1Q	2Q	3Q	最大値
$subjective1_{i,t}$	1.97	0.66	0.31	1.63	1.80	2.46	3.80
$subjective2_{i,t}$	-0.40	0.64	-1.78	-0.78	-0.36	-0.08	1.64
$performance_{i,t}$	4.01	0.68	2.00	3.50	4.00	4.50	5.50
$rise_{i,t}$	3.34	0.75	1.00	3.00	3.00	4.00	5.00
$performance_{i,t-1}$	4.03	0.65	2.00	3.50	4.00	4.50	5.50
$promotion_{i,t-1}$	0.04	0.21	0.00	0.00	0.00	0.00	1.00
$career_{i,t}$	15.60	5.41	3.00	11.75	16.00	20.00	27.00
$grade_{i,t}$	4.18	0.97	3.00	3.00	4.00	5.00	7.00

表 A3: $subjective2_{i,t}$ の上位 10% グループにおける要約統計量

	平均値	標準偏差	最小値	1Q	2Q	3Q	最大値
$subjective1_{i,t}$	2.97	0.54	1.29	2.46	3.11	3.29	4.46
$subjective2_{i,t}$	0.39	0.55	-1.32	-0.08	0.37	0.68	1.92
$performance_{i,t}$	4.81	0.55	3.00	4.50	5.00	5.00	6.50
$rise_{i,t}$	4.62	0.57	3.00	4.00	5.00	5.00	6.00
$performance_{i,t-1}$	4.82	0.55	3.00	4.50	5.00	5.00	6.50
$promotion_{i,t-1}$	0.15	0.36	0.00	0.00	0.00	0.00	1.00
$career_{i,t}$	14.35	5.28	3.00	11.00	14.00	18.00	30.00
$grade_{i,t}$	4.87	1.09	3.00	4.00	5.00	5.00	8.00

表 A4: $subjective2_{i,t}$ の下位 10% グループにおける要約統計量

	平均値	標準偏差	最小値	1Q	2Q	3Q	最大値
$subjective1_{i,t}$	1.99	0.65	0.31	1.63	1.80	2.46	3.80
$subjective2_{i,t}$	-0.41	0.62	-1.78	-0.78	-0.36	-0.08	1.64
$performance_{i,t}$	4.11	0.67	2.00	4.00	4.00	4.50	5.00
$rise_{i,t}$	3.40	0.74	1.00	3.00	3.00	4.00	5.00
$performance_{i,t-1}$	4.12	0.65	2.00	4.00	4.00	4.50	5.50
$promotion_{i,t-1}$	0.04	0.21	0.00	0.00	0.00	0.00	1.00
$career_{i,t}$	16.55	5.14	5.00	13.00	17.00	20.25	27.00
$grade_{i,t}$	4.36	1.08	3.00	4.00	4.00	5.00	7.00

表 A5: $performance_{i,t}$ の上位 10% グループにおける要約統計量

	平均値	標準偏差	最小値	1Q	2Q	3Q	最大値
$subjective1_{i,t}$	2.84	0.55	1.29	2.29	3.11	3.29	4.29
$subjective2_{i,t}$	0.20	0.55	-1.32	-0.32	0.37	0.68	1.68
$performance_{i,t}$	5.12	0.47	3.00	5.00	5.00	5.50	6.50
$rise_{i,t}$	4.60	0.60	3.00	4.00	5.00	5.00	6.00
$performance_{i,t-1}$	5.11	0.47	3.50	5.00	5.00	5.50	6.50
$promotion_{i,t-1}$	0.14	0.35	0.00	0.00	0.00	0.00	1.00
$career_{i,t}$	13.37	5.04	3.00	10.00	13.00	17.00	27.00
$grade_{i,t}$	4.82	1.19	3.00	4.00	5.00	6.00	7.00

表 A6: $performance_{i,t}$ の下位 10% グループにおける要約統計量

	平均値	標準偏差	最小値	1Q	2Q	3Q	最大値
$subjective1_{i,t}$	2.19	0.68	0.31	1.63	1.97	2.63	3.97
$subjective2_{i,t}$	-0.13	0.67	-1.78	-0.78	-0.17	0.22	1.83
$performance_{i,t}$	3.82	0.65	2.00	3.50	4.00	4.00	5.50
$rise_{i,t}$	3.50	0.75	1.00	3.00	3.00	4.00	5.00
$performance_{i,t-1}$	3.86	0.64	2.00	3.50	4.00	4.50	5.50
$promotion_{i,t-1}$	0.04	0.20	0.00	0.00	0.00	0.00	1.00
$career_{i,t}$	16.65	4.90	5.00	13.00	16.00	20.00	30.00
$grade_{i,t}$	4.79	0.93	3.00	4.00	5.00	5.00	7.00

なお、整数値をとる勤続年数 ($career$) に関して、四分位点などの統計量が小数点表示されているのは、四分位点に該当する値が複数あった場合、これらの値の観測頻度に基づいた加重平均を統計量として採用するからである。

B. 上位・下位グループごとの相関係数表

表 B1: $subjective1_{i,t}$ の上位 10% グループにおける相関係数表

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) $subjective1_{i,t}$	1.00							
(2) $subjective2_{i,t}$	0.98	1.00						
(3) $performance_{i,t}$	0.04	-0.18	1.00					
(4) $rise_{i,t}$	0.94	0.85	0.36	1.00				
(5) $performance_{i,t-1}$	0.31	0.21	0.44	0.44	1.00			
(6) $promotion_{i,t-1}$	0.23	0.24	-0.05	0.20	0.26	1.00		
(7) $career_{i,t}$	0.06	0.07	-0.04	0.04	-0.03	-0.11	1.00	
(8) $grade_{i,t}$	0.14	0.13	0.02	0.14	0.07	0.10	0.80	1.00

表 B2: $subjective1_{i,t}$ の下位 10% グループにおける相関係数表

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) $subjective1_{i,t}$	1.00							
(2) $subjective2_{i,t}$	0.92	1.00						
(3) $performance_{i,t}$	0.24	-0.13	1.00					
(4) $rise_{i,t}$	0.95	0.77	0.52	1.00				
(5) $performance_{i,t-1}$	0.56	0.39	0.48	0.64	1.00			
(6) $promotion_{i,t-1}$	0.50	0.48	0.12	0.48	0.28	1.00		
(7) $career_{i,t}$	-0.12	-0.11	-0.02	-0.11	-0.05	-0.09	1.00	
(8) $grade_{i,t}$	0.08	0.10	-0.03	0.06	0.03	0.10	0.71	1.00

表 B3: $subjective2_{i,t}$ の上位 10% グループにおける相関係数表

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) $subjective1_{i,t}$	1.00							
(2) $subjective2_{i,t}$	0.97	1.00						
(3) $performance_{i,t}$	-0.01	-0.23	1.00					
(4) $rise_{i,t}$	0.94	0.84	0.33	1.00				
(5) $performance_{i,t-1}$	0.37	0.31	0.29	0.45	1.00			
(6) $promotion_{i,t-1}$	0.21	0.22	-0.08	0.17	0.27	1.00		
(7) $career_{i,t}$	0.05	0.08	-0.11	0.01	-0.14	-0.12	1.00	
(8) $grade_{i,t}$	0.10	0.11	-0.04	0.08	-0.05	0.08	0.79	1.00

表 B4: $subjective2_{i,t}$ の下位 10% グループにおける相関係数表

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) $subjective1_{i,t}$	1.00							
(2) $subjective2_{i,t}$	0.93	1.00						
(3) $performance_{i,t}$	0.24	-0.11	1.00					
(4) $rise_{i,t}$	0.95	0.78	0.52	1.00				
(5) $performance_{i,t-1}$	0.55	0.38	0.52	0.65	1.00			
(6) $promotion_{i,t-1}$	0.50	0.48	0.11	0.47	0.17	1.00		
(7) $career_{i,t}$	-0.09	-0.11	0.06	-0.06	0.05	-0.15	1.00	
(8) $grade_{i,t}$	0.14	0.10	0.13	0.16	0.18	0.05	0.68	1.00

表 B5: $performance_{i,t}$ の上位 10% グループにおける相関係数表

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) $subjective1_{i,t}$	1.00							
(2) $subjective2_{i,t}$	0.98	1.00						
(3) $performance_{i,t}$	0.13	-0.05	1.00					
(4) $rise_{i,t}$	0.96	0.90	0.39	1.00				
(5) $performance_{i,t-1}$	0.37	0.30	0.36	0.44	1.00			
(6) $promotion_{i,t-1}$	0.26	0.27	-0.05	0.23	0.20	1.00		
(7) $career_{i,t}$	0.06	0.07	-0.04	0.05	0.05	-0.17	1.00	
(8) $grade_{i,t}$	0.16	0.16	-0.03	0.14	0.10	0.07	0.73	1.00

表 B6: $performance_{i,t}$ の下位 10% グループにおける相関係数表

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) $subjective1_{i,t}$	1.00							
(2) $subjective2_{i,t}$	0.94	1.00						
(3) $performance_{i,t}$	0.15	-0.16	1.00					
(4) $rise_{i,t}$	0.96	0.81	0.44	1.00				
(5) $performance_{i,t-1}$	0.56	0.44	0.36	0.61	1.00			
(6) $promotion_{i,t-1}$	0.35	0.31	0.17	0.36	0.18	1.00		
(7) $career_{i,t}$	-0.06	-0.07	0.02	-0.05	-0.05	-0.17	1.00	
(8) $grade_{i,t}$	0.24	0.19	0.15	0.26	0.16	0.07	0.70	1.00

謝辞

本論文は筆者が神戸大学大学院経営学研究科博士後期課程に在籍中の研究成果である。同研究科教授の三矢裕先生には指導教官として、その遂行にあたり終始、ご指導を戴いた。ここに深謝の意を表す。同研究科教授 松尾貴巳先生、並びに、同研究科教授 梶原武久先生には副査としてご助言を戴くとともに本論文の細部にわたりご指導を戴いた。ここに深謝の意を表す。

本研究の遂行には、リサーチ・サイトとなった企業の方々の協力が不可欠であった。本研究の遂行に真摯に協力して頂いた A 社の方々に、この場を借りてお礼申し上げたい。

本論文の執筆にあたって、松山大学経営学部講師 佐久間智広先生、並びに鹿児島大学稲盛アカデミー講師 劉美玲先生には本論文に直接目を通して戴き、丁寧なご助言並びにご指導を戴いた。ここに感謝の意を表す

本研究科 三矢裕 研究室の大学院生には、日頃より有益なご討論、ご助言を戴いた。この場を借りてお礼申し上げたい。

本研究は、神戸大学管理会計研究会での研究会報告の機会を戴き、いくつもの貴重なコメントを得ることができた。頂戴したコメントは本研究の随所に反映されている。本研究に真摯に向き合ってくださいました神戸大学管理会計研究会の皆様、改めてお礼申し上げたい。