Subjective Performance Evaluation

Jasmijn C. Bol
IESE Business School
University of Navarra
Barcelona, Spain 08034
e-mail: docjcbol@iese.edu

November 2005

Subjective Performance Evaluation

Abstract

This paper presents a conceptual framework for analyzing the role of subjectivity in compensation contracting. We start by examining how introducing subjectivity can improve compensation contracting. Both theoretical and empirical papers that model the benefits of subjectivity and that examine the weight placed on subjective performance appraisal are examined. Subsequently, we discuss the drawbacks of subjective performance appraisal. We distinguish between situations in which the principal is also the residual claimant and situations in which he is not. When the principal is the residual claimant, the main problem is enforceability. We discuss specific contract structures that allow the principal to deal with that, as well as their main incentive-related disadvantages. The main difficulty when the principal is not the residual claimant is supervisor bias. We examine supervisor incentives, analyze how rating bias influences contract outcomes and inspect several ways to limit this bias. Finally, we present an outline of the issues that invite additional work.

I. SUBJECTIVITY IN COMPENSATION CONTRACTING

This paper presents a conceptual framework to analyze the role of subjectivity in compensation contracting. For the majority of employees it is hard to capture individual contributions to total firm value accurately using only objective performance measures. Nevertheless, only recently have researchers dedicated attention to the role of subjectivity in compensation contracting. Not surprisingly, these studies show that subjectivity is an important characteristic of most incentive contracts (Prendergast, 1999; Gibbs, Merchant, Van der Stede and Vargus, 2004). Murphy and Oyer (2003), using a cross-sectional dataset of executive bonus plans, illustrate several ways in which firms exercise discretion in awarding annual bonuses[1]. Subjectivity is not limited exclusively to incentive contracts for executives and managers: the Bureau of National Affairs (1981), as cited by Kahn and Sherer (1990) reports that, for all types of jobs, compensation contracts involving subjective appraisal are more common than those involving objective measurement.

The main focus in the compensation contracting literature has, however, been on the explicit compensation contracts of workers whose individual contribution is relatively easy observed. These studies have provided important insights in incentive provision and contract design (see e.g. Prendergast, 1999). Although many of these insights remain valid in compensation contracts that include subjectivity, subjective performance appraisal also introduces many additional concerns. The trade-off between risk and incentive, for example, is probably less predictive of contract design when supervisors do not truly differentiate between good and bad performers.

[1] Other examples of studies that find indications of extensive use of subjectivity in compensation contracts are, Bushman, Indjejikian and Smith (1996), Hayes and Schaefer (2000), Gibbs, Merchant, Van Der Stede and Vargus (2004) and Moers (2005).

This paper offers an extensive discussion of how compensation contracting policies change when employee performance is (partly) assessed subjectively. We raise several points that have not been made explicit in previous work. We examine, for example, the incentives of the principal in subjective performance appraisal, distinguishing principals who are residual claimants of the employees' contributions explicitly from those who are not. We also analyze how subjective performance appraisal biases performance-rating distributions and discuss how these biases affect employee behavior. Both theoretical and empirical perspectives are presented. Moreover, our discussion is not limited to the analysis of economics based research, as several studies, e.g. Ittner, Larcker and Meyer (2003), have indicated that for studying subjective performance appraisal psychology based explanations may be equally or more predictive than economics based explanations.

The paper proceeds, in the next section, by discussing how subjective performance measures can improve incentive contracting. We examine whether the weight placed on subjective performance appraisal is indeed higher when incentive contracting improvement is expected. In section III, we analyze enforceability, the main problem of introducing subjectivity into a compensation contract when the principal is residual claimant. Our focus is on the contract structures that make compensation contracts that include subjectivity enforceable, as well as their main disadvantages for incentive contracting purposes. Section IV also deals with problems encountered when subjectivity is part of the compensation contract, but in this section, we focus on the incentive problems of principals who are not residual claimants. We discuss several reasons why supervisors might bias subjective performance ratings. Subsequently, we analyze how bias influences contract outcomes and inspect several ways to limit rating bias. Finally, we conclude and point out numerous avenues for future research.

II. SUBJECTIVITY AND INCENTIVES

Contracting theory tells us that an ideal performance measure for incentive purposes would be a measure that reflects an employee's true contribution to firm value. A measure that is accurate, informative, timely, and that takes into account the effects of current actions on the future profitability of the firm, all this without transferring too much risk to the agent (Holmstrom, 1979; Baker, Gibbons and Murphy, 1994; Gibbs et al., 2004). Unfortunately, in practice, objective performance measures rarely satisfy this description: they are in most cases extremely noisy and noncongruent (Feltham and Xie, 1994; Murphy and Oyer, 2003). If objective performance measures that do not capture the true contribution to firm value are used, incentives will be distorted and negative side effects will result[2].

While an employee's exact contribution can usually not be captured perfectly by objective performance measures, often it can be subjectively assessed (Murphy and Oyer, 2003). This provides firms with the possibility to supplement or even completely replace objective performance measurement.

Subjectivity can be introduced into a compensation contract in three ways; (i) by allowing for (ex post) flexibility in the weighting of objective performance measures, (ii) by using subjective performance measures, and (iii) by allowing for ex post discretional adjustments based on factors other than the performance measures specified ex ante (Gibbs et al., 2004). Although the way subjectivity is introduced is

[2] The problem of providing distorted incentives has long been recognized by researchers. Kerr argued as early as 1975 that value-maximizing behavior is often too complicated to be measured accurately by straightforward objective measures.

likely to be of influence, most theoretical and empirical investigations into subjectivity do not make a distinction (Ittner et al., 2003). Research that would illustrate the distinct characteristics of these different forms of subjectivity and would analyze how this affects their role in compensation contracting would be very valuable. In the following three subsections, we discuss how the introduction of subjectivity into compensation contracts can improve incentive contracting.

Mitigation of Incentive Distortions

Ideally, compensation contracts should incorporate all information about all different job dimensions, weighted properly, so that incentives are appropriately balanced across these different dimensions (Holmstrom, 1979). Objective performance measures often inadequately account for, or entirely ignore, some dimensions of the employee's job (Holmstrom and Milgrom, 1991; Feltham and Xie, 1994). Baker, Gibbons, and Murphy (1994) show that complementing objective performance measures with subjective ones can improve incentive contracting as value-enhancing efforts that are not easily quantified can be included. They explore the use of a contract that is simultaneously based on an objective performance measure and a subjective assessment of total contribution, and show that in certain circumstances the combination of objective measurement and subjective appraisal can yield profit while neither separately can. The inclusion of the subjective assessment of total contribution in the compensation contract reduces distortionary incentives as it prevents the agent from only directing his effort towards the objectively measured tasks. Confirming the prediction that objective performance measures will often not adequately account for all job dimensions, Brown (1990) and MacLeod and Parent

(1999) find that jobs with many tasks are characterized by subjective performance assessment.

Accounting measures, frequently used objective performance measures, are by nature backward-looking, and do not accurately reflect the effects of employees' actions or decisions on future firm value (Kaplan and Norton, 1996). Subjectivity can mitigate an excessively short-term focus in situations where objective measures are unable to create the right long-term focus. Empirical findings suggest that that the weight on subjective performance measures indeed increases with the importance of actions that affect both present and future firm value (Bushman et al., 1996) and that subjectivity is used to encourage employees to adopt a long-term focus (Gibbs et al., 2004).

When using an explicit incentive contract with objective performance measures it is hard to include information not foreseen ex ante. The only way to prevent distorted incentives is to include the possibility of renegotiating the contract. Renegotiating and changing formal bonus contracts can be expensive (Baker, Jensen and Murphy, 1988). A way to overcome this limitation is to use an incentive contract with some subjectivity in order to allow the principal to exploit any additional relevant information that surfaces during the contracting period. Hence, subjectivity can improve incentive contracting by enabling the principal to communicate and evaluate performance according to the changed priorities of the organization (Davila, 2003).

Risk Reduction

Objective performance measures are often noisy, that is, they provide some information about the agent's effort but are contaminated by uncontrollable random events. This introduces uncertainty, and the principal will have to compensate the agent for bearing this risk. Subjectivity allows the principal to filter out uncontrollable event. This reduces risk and allows for stronger incentives (Holmstrom, 1979; Banker and Datar, 1989; Feltham and Xie, 1994). Hayes and Schaefer (2000) find empirical evidence consistent with principals substituting away from noisier objective performance measures to possibly subjective privately observed measures.

Adjusting for uncontrollable events can also help deal with organizational interdependencies. Objective performance measures are often too broad or too narrow to capture an employee's contribution accurately. Too narrow performance measures will lead to incentive distortion, while too broad performance measures will put too much risk on the employee. By allowing for some discretion, the principal can control for this problem. Both Murphy and Oyer (2003) and Gibbs, Merchant, van der Stede and Vargus (2004) present empirical evidence indicating that subjectivity is used to deal with interdependencies.

Incentive contracts typically include performance targets, which are set ex ante at a level that will motivate the agent to provide effort (Murphy, 2000). However, when uncontrollable events change the business environment, the optimal performance targets are likely to shift. If the principal has the possibility to make ex post adjustments he can make sure the agents remain properly motivated. The principal, however, might be reluctant to adjust for 'good luck' as this might be perceived as increasing the performance standard when performance is better than expected (Gibbons, 1987). The use of subjectivity to adjust performance targets is therefore only likely to occur when 'bad luck' makes targets unreachable. Gibbs, Merchant, van der Stede and Vargus (2004) show that subjectivity is positively related to the difficulty of meeting the performance target when failure to achieve the target is nontrivial for compensation.

Limitation of Vulnerability to Manipulation

Objective performance measures are normally defined in clear numerical terms, which makes them susceptible to manipulation (Holmstrom and Milgrom, 1991). Employees often have a better understanding of which (not necessarily desirable) actions lead to increased measured performance. They can use this information to “game the system,” to take actions that improve their compensation but go against the company’s interest. Subjectivity can reduce the vulnerability to gaming as performance is appraised ex post instead of measured according to measures that are set ex ante. This provides the principal with the possibility to punish the agent when manipulation or intentional noncongruent behavior is detected. The threat of being punished for knowingly taking actions that go against the company’s interests will diminish the employee’s incentives to manipulate the system.

Advantages and Drawbacks

In this section, we have examined how introducing subjectivity can improve incentive contracting. We have discussed analytical models illustrating the benefits of subjectivity and empirical papers that confirm that the weight placed on subjectivity is higher in situations where subjectivity is expected to be more beneficial. Studies that actually empirically investigate the outcomes of introducing subjectivity into the compensation contract are, however, almost entirely lacking. The study by Ittner, Larcker and Meyer (2003) is one of the few studies that address this issue: they investigate the introduction of a 'balanced scorecard' bonus plan with subjective weightings and, contrary to expectations, they find that the system was considered disappointing and eventually abandoned. This investigation shows that, although subjective performance assessment might improve incentive contracting in several ways, the introduction of subjectivity also has negative consequences. Subjective assessment influences the behavior of both principal and agent, and these behavioral changes can be so strong that they destroy firm value. We turn to the drawbacks of subjectivity in the following two sections. First, in the next section, we will discuss subjective performance appraisal in situations where the principal is also the residual claimant. In the subsequent section, we will discuss how matters changes when he is not.

III. SUBJECTIVE APPRAISAL BY RESIDUAL CLAIMANTS

The most pronounced problem with subjectivity is that it provides supervisors with the possibility to assess performance untruthfully. The mere fact that performance evaluation is subject to the supervisor's discretion makes it impossible for a court of law to enforce the contract. The lack of third party enforcement provides both parties with the possibility to renege on their pledges. If the supervisor is the residual claimant, additional compensation rewarded to the agent will decrease his own wealth, which gives him incentives to underreport performance to keep costs down. The agent, on the other hand, will not offer the promised effort if he does not believe that his effort will be rewarded appropriately. When performance is assessed subjectively, it is not sufficient for the size of the promised bonus to be large enough to outweigh the personal costs of effort. For the compensation contract not to lose its incentive effects the agent must trust the principal. This trust is brought about if the agent believes that the principal has nothing to gain or something to lose by reneging. Hence, in order to ensure that the contract does not break down it must be self-enforcing, i.e., the contract must not create incentives for either party to renege (MacLeod and Malcomson, 1989). In the following subsections, we will discuss several ways to obtain self-enforceability.

Reputational Concerns

The firm's concern for its reputation may induce it to honor its compensation promises, even when it has contemporaneous incentives to cheat (Carmichael, 1989). A firm's reputation is valuable in a compensation contracting setting as it influences the agent's and potential future agents' expectations of the firm's future behavior. A firm's reputation provides the agents with an idea of how the firm has treated its agents in the past. Agents use this information to build their expectations and will act accordingly in the next contracting period. A firm with a bad labor market reputation will therefore have difficulties in hiring new agents and in using compensation contracts that include subjectivity to make its current agents offer the desired level of effort. A recent survey among 92 companies indicates that a bad labor market reputation is indeed a concern; nearly two-thirds of the surveyed companies say that their management's lack of credibility has been an obstacle in their relationship with employees (The Conference Board, 1997 as cited by Portales, Ricart and Rosanas, 1998).

The key requirement to sustain a compensation pledge based on reputational concerns is that the reputational costs of breach offset the immediate gains of breach. For this to be probable, the firm must value its reputation sufficiently. The value of a firm's labor market reputation depends on the extra profit it allows the firm to earn and on how strongly these profits are discounted. The strength of the effect of reputation on profits depends on how rapidly and accurately information about agreement compliance or breaching flows to other agents in the firm and to potential agents in the labor market. This information is not always transmitted flawlessly, especially not to external labor markets (Bull, 1983). The main problem is that after a breach both parties have an incentive to claim that the other party broke the agreement. Since there are no observable contract terms outsiders cannot verify the correctness of these claims. Moreover, for outsiders it is hard to distinguish between deliberate opportunistic breaches and honest confusions about the content of the agreements.

Bertrand (2004) empirically considers whether firms are indeed less likely to renege on their pledges when future profits are expected to be high, than when firms expect to gain little in the future from their reputation. She shows that firms are more likely to breach their compensation promises when more severe competition lowers earnings and increases default possibilities.

The conditions under which reputational effects can induce the principal to honor his compensation pledges are rather stringent but, as illustrated by, e.g., Bull (1983) and Macleod and Malcomson (1989), they can generate efficient outcomes that would not arise in a static setting. Baker, Gibbons and Murphy (1994), in their examination of the simultaneous use of objective and subjective performance assessment, show that the combination makes these conditions less stringent, as the objective part of the compensation contract can increase the value of the ongoing relationship between the firm and the employee and thereby decrease the firm's temptation to renege on the bonus promise.

Bonus Pools

Another potential way to eliminate a firm's financial incentive to renege on its pledges is by setting up fixed bonus pools. In this case, the principal objectively determines the total amount of bonus to be paid out to the employees covered by the bonus pool, without committing to the manner in which the money is distributed among them. Since the total bonus is determined by contractible objective performance measures, individual performance evaluation can be carried out both objectively and subjectively, as the residual claimant will not longer have ex post incentives to provide deflated performance ratings. Baiman and Rajan (1995) investigate the characteristics of bonus pools and show that as long as the subjective assessment is informative about at least one agent, bonus pool arrangements are strict Pareto improvements.

Despite the advantages of fixed bonus pools, clear examples of companies that truly pre-commit to a wage bill are rare (Prendergast and Topel, 1993). This might be explained by the fact that bonus pools are zero-sum games. Each agent's compensation depends on the performance of other agents covered by the arrangement. These interdependencies make performance pay more risky (less controllable) as the amount of money received by the agent is no longer solely determined by his own performance but also by the performance of other agents (Baiman and Rajan, 1995). Moreover, cooperation among agents is discouraged as performance cannot only be improved by increasing one's own effort but also by sabotaging the others (Lazear, 1989).

Protecting his residual claim is obviously not the principal's sole potential motive to assess performance incorrectly. Other potential motives for both residual and non-residual claimants are discussed in the next section.

IV. SUBJECTIVITY IN MULTI-LAYERED AGENCY RELATIONSHIPS

Even though the principal is assumed to be the residual claimant of the employee's output in most agency models (e.g. Bull, 1983; Baker et al., 1994), this is not always the case. In most principal-agent relationships the principal will not be the residual claimant, as companies are almost always multi-layered (Prendergast, 1999). The fact that most supervisors are not residual claimants severely changes the incentives in the performance appraisal process; the financial incentive to renege on earlier made pledges is limited or non-existent. Nevertheless, it is still rational for employees not to trust their performance appraisals, as discretion still leaves room for inaccuracy, preferences, and bias (Prendergast and Topel, 1993). Supervisors are themselves agents, who have incentives to evaluate workers in ways not desired by the principal (Prendergast, 1999).

Subjectivity and Biased Ratings

The literature in economics and especially in human resources management has presented considerable evidence on biased subjective performance ratings. For example, Bretz, Milkovich and Read (1992), using a combination of several databases[3], find that performance appraisal systems based on subjective assessment typically have five performance levels to differentiate between employees but that usually only three of these five levels are used. The practice of offering all employees ratings that differ little from the norm is referred to as "centrality bias" and has been well documented in the human resource literature, e.g., Landy and Farr (1980), Murphy and Cleveland (1991).

[3] They combine the 1989 survey by the Wyatt Company, the 1990 survey by The Conference Board, and their own survey of the Fortune Industrial 100. For more information see Bretz, Milkovich and Read (1992).

Subjective performance ratings have also been found to be skewed towards the top end of the scale: this is referred to as “leniency bias”. Bretz, Milkovich and Read (1992) find that it is common for 60 to 70% of the firm’s employees to be rated in the top two performance levels. These skewed performance distributions imply that supervisors overstate performance. An alternative explanation could be an outstanding workforce. This explanation is not very credible as the phenomenon is documented consistently across organizations (Jawahar and Williams, 1997). It is unlikely that all organizations have mainly outstanding employees. Moreover, if an outstanding workforce caused the skewed distributions, the pattern should be found in both objective and subjective performance ratings. Moers (2005) compares the performance ratings patterns of objective and subjective performance measures and shows that objective performance measures do not show the same extent of compression and skewedness.

The following subsections discuss several plausible reasons for a supervisor to bias performance evaluations.

Information Gathering

In the agency literature, the principal is often assumed to receive signals on the agents’ actions constantly and costlessly. This is not very realistic, as in most cases monitoring is costly and only performed sporadically (Prendergast, 2002). If supervisors are not explicitly rewarded for accurate performance rating, it is not likely that they will be motivated to invest time in gathering information since they bear all information gathering costs (e.g. monitoring costs), while they receive little of the benefit from conducting more accurate evaluations (Baker et al., 1988). If information is not gathered thoroughly and deliberately, supervisors will have a stronger tendency to base their ratings on general impressions. This tendency to attend to a global impression rather than to carefully distinguish between different performance dimensions, the so called "halo effect" (Fox, Bizman and Herrmann, 1983), is more common when performance rating is done very fast (Jennings, Palmer and Thomas, 2004). The lack of information is also likely to result in too heavy a reliance on knowledge of prior performance or on expectations. When the prior opinion is inconsistent with current performance this reliance will naturally lead to inaccurate and biased ratings (Murphy, Balzer, Lockhart and Eisenman, 1985; Mount and Thompson, 1987).

Despite the fact that a lack of incentives to dedicate time to information gathering and processing might lead to biased ratings, it remains uncommon for managers to be evaluated on how they manage the appraisal process. Napier and Latham (1986) examine manager's expected rewards for rating accuracy and find that supervisors do not perceive any direct private consequences for conducting thorough performance appraisals. This is confirmed by Bretz, Milkovich and Read (1992), who show that only about one quarter of the investigated firms holds their supervisors accountable for how they manage the performance appraisal process.

Rewards

The lack of direct awards for accurate ratings might be explained by the fact that companies do not necessarily seek accurate ratings. Companies are interested in the effectiveness of performance-based compensation contracts in maintaining or increasing the employees' future level of performance, not necessarily in the accuracy of the performance appraisal as such. Longenecker, Sims, and Gioia (1987) investigate the different objectives of performance appraisal and found that motivating and rewarding subordinates is considered to be the most important one by practicing executives.

In order to motivate the supervisor to manage the performance appraisal process in an efficient way the firm can link his compensation and promotion possibilities to his unit's performance (Harris, 1994). These incentives will motivate the supervisor to manage the performance appraisal process in a productivity maximizing way. Telling an employee that his performance has been poor can easily result in discouragement and decreased effort levels. In order to keep employees motivated the supervisor might feel it is necessary to overstate performance ratings, especially when performance is only temporarily suffering. The supervisor might also compress ratings, as compressed ratings are expected to reduce disharmony between employees and to create a cooperative work environment. Concern that knowledge about bad performing employees may reflect poorly on the supervisor's own capabilities might also lead to lenient ratings.

Negative Consequences

Supervisors must live with their employees in a day-to-day relationship and providing employees with negative feedback might therefore be simply unpleasant (Harris, 1994). Offering employees harsh but accurate ratings is likely to damage personal relationships and to lead to discussions and criticism (Bernardin and Buckley, 1981). Supervisors avoid the real and psychic cost of communicating poor evaluations by providing compressed and lenient ratings.

This defensive behavior is found to be stronger when the expected consequences of providing negative feedback are more pronounced. For example, Varma, Denisi and Peters (1996) find that rating bias is positively related to positive interpersonal affect. The ultimate use of the performance ratings is also likely to affect the extent of bias. Empirical studies show that ratings used to make administrative decisions such as merit pay or promotions are more lenient, and have less variance, than ratings obtained for training or employee development purposes (Jawahar and Williams, 1997).

Favoritism

Supervisors are utility-maximizing agents who take personal considerations into account when appraising employees. For instance, supervisors might inflate ratings to reward appreciated employees, to encourage loyalty, or to promote their personal agendas. On the other hand, supervisors may deflate ratings to punish rebellious employees (Longenecker et al., 1987). Empirical evidence indicates that supervisors do not necessarily assign performance ratings based on "real" performance but rather to serve their self-interest (Ferris and Judge, 1991).

Employees anticipate that supervisors will act on personal preferences and react by engaging in "rent-seeking activities", actions that agents carry out in order to increase ratings at the expense of actions that are truly productive. For instance, an employee could exert excessive effort on visible tasks, or he could work 'too hard' in order to signal worker quality. Furthermore, the employee may intend to influence the supervisor's decision, for example by ingratiation[4] (Higgins et al., 2003) or by currying favors (Milgrom, 1988). Several studies in the personnel literature have investigated the effectiveness of this kind of behavior and found that especially ingratiation led to highly positive performance ratings (Ferris and Judge, 1991).

[4] Ingratiation is behavior designed to increase the supervisor's liking of oneself or to make oneself appear friendly in order to get what one wants (Higgins, Judge and Ferris, 2003).

These rent-seeking activities by employees are harmful to the organization as agents devote time and energy trying to influence the supervisor's decision that would have been better spent on productive activities. Moreover, the value of performance ratings for personnel decisions is greatly reduced when it is not clear whether good performance ratings result from favoritism or from genuine good performance. It increases the probability that employees are promoted beyond their capacities (Prendergast, 1999). However, favoritism can also be beneficial to the company, as supervisors derive utility from exercising bias. This allows the firm to charge supervisors a price for exercising their preferences (Prendergast and Topel, 1996).

Cognitive Limitations

Divergence from true ratings is not necessarily intentional[5]; bias might also be a consequence of the supervisor's cognitive limitations. Cognitive limitations may prevent supervisors from fully exploiting the information found in a diverse set of performance measures. Unconsciously, supervisors focus more on some performance dimensions than on others (without them necessarily being more important). Lipe and Salterio (2000), for example, show that when superiors have discretion in weighing different performance measures, they appear to disregard unique measures and overemphasize common measures ex post. Ittner, Larcker, and Meyer (2003), provide empirical evidence on the 'outcome effect': they show that supervisors put greater weight on financial outcome measures than on the drivers of financial results[6].

---

[5] The "halo effect" and excessive reliance on prior performance or expectations are also forms of bias that are done unconsciously, however since they are more profound when the supervisor does not gather information thoroughly and deliberately we have discussed them in the information gathering section.

[6] Bias caused by cognitive limitations is not unique to subjective weights on objective performance measures. Cognitive limitations are likely to influence performance ratings in all situations where supervisors need to consider the importance of different actions / aspects of behavior.

When supervisors put more emphasis on certain dimensions of performance ex post, employees will put more emphasis on these dimensions ex ante. Hence, although subjective weights on objective performance measures can be used to 'back out' unintended dysfunctional behavior created by objective measurement (Baker et al., 1994), supervisors' cognitive limitations might still prevent the compensation contract from striking the right balance between the different performance dimensions. In order to understand the implications of introducing subjectivity into compensation contracts we must take into consideration supervisors' incentives to bias ratings but also the supervisors' inability to avoid doing so.

Subjectivity, Bias and Outcomes

Although there is considerable empirical evidence on the existence of biased ratings, studies that consider the consequences of rating bias on compensation contracting empirically are rare. Nevertheless, we would expect the lack of undistorted ratings to influence incentive contracting in several ways[7]. First, bias is a form of randomness that adds risk to the monitoring process. Thus, though allowing for a holistic view of the agent's true value added eliminates some risk, it introduces a different source of risk.

In the same way, subjectivity might not be as effective in dealing with multi-tasking problems as assumed. Subjectivity might prevent employees from solely spending time on measurable job dimensions; it does not keep employees from focusing on more visible job dimensions or from spending time on rent-seeking activities. Hence, introducing subjectivity into compensation contracting does not guarantee improved effort allocation.

[7] Outside of the realm of strictly compensation contracting related issues, biased ratings are also of dubious value for personnel decisions (Jawahar and Williams, 1997). The wrong employees may be promoted and employees with real training needs might not be identified.

Probably the most important drawback of biased performance ratings is that employees might feel treated so randomly and unfairly that they lose all trust in the performance appraisal system. Behavioral research by e.g., Tremblay and Roussel (2001) has shown that employees do not only care about the amount of compensation received (distributive justice), but also about the fairness of the procedures that determine the amount of rewards (procedural justice). Employees do not necessarily consider objective performance measurement systems to be fair, especially not when true value added is not captured. Still, objective performance assessment has the important advantage of setting clear measurement criteria. This provides employees with a certain amount of 'control' as they know which actions will lead to increased compensation. With subjective performance assessment, it is not exactly clear which actions determine the performance ratings. This uncertainly prevents employees from checking whether performance was assessed accurately. In this situation, employees might expect that bias and favoritism played an important role, even when they did not. This negative feeling affects morale, perhaps even among workers who are unduly favored. Moreover, employees that feel discriminated against may resign, which will result in high turnover costs and loss of human capital. Ittner, Larcker and Meyer (2003) analyze the implementation of a compensation plan with subjectively determined weights and find that the subjectivity led many employees to complain about favoritism and uncertainty in the criteria used to determine rewards. Gibbs, Merchant, van der Stede and Vargus (2004) examine the role of trust when compensation contracts include subjective assessment. Their empirical evidence indicates that the relation between subjectivity and pay satisfaction and performance is more positive the greater the manager's tenure, their proxy for trust.

There are, however, also reasons to believe that bias could improve incentive provision. One of the benefits claimed by supervisors is increased employee motivation (Longenecker et al., 1987). Empirical evidence suggests that those who receive good evaluations become encouraged to deliver greater effort while those receiving bad evaluations become discouraged. By not confronting bad performers with their true ratings, the supervisor intends to keep bad performing employees motivated. Especially when promotion possibilities are based on the performance of several years it might be better not to signal to the employee early in the period that he has no probability of being promoted (Prendergast, 1999).

For the same reason, it might be beneficial to inflate the ratings of all employees. Individuals have a tendency to rate themselves higher than their supervisors do (Shore and Thornton, 1986; Harris and Schaubroeck, 1988). If supervisors provide workers with their true ratings, which are bound to be lower than expected, employees are likely to feel unappreciated and underpaid. When employees receive less compensation than they think they deserve, they are expected to lower their effort to restore their feeling of equity (Kahn and Sherer, 1990). Pearce and Porter (1986) have investigated how employees react to performance ratings and show that feedback describing an employee as satisfactory (as compared to above average or outstanding) leads to reduced organizational commitment. Finally, as argued by personnel managers and labor unions, compressed ratings may improve motivation as they decrease disharmony among workers and help create a cooperative work environment.

Although compressed and lenient ratings are likely to improve the motivation of part of the employees, it is far from obvious that they are morale improving for top performers. Better workers are likely to feel disenchanted when employees that perform worse are rewarded almost equally. Hence, whether lenient and compressed ratings improve overall productivity will depend on whether the motivation of poor performers is sufficiently increased to offset the decreased motivation of top performers. Unfortunately, we do not know enough about the relative (empirical and theoretical) significance of the different effects.

Preventing Biased Ratings

There are several ways to keep ratings from becoming biased. Forced rating schemes are one such way. If the supervisor is required to categorize workers into groups, he can no longer bias the distribution of the performance ratings. Although it appears to be a good way to avoid biased ratings, Bretz and Mikovich (1989)[8] find that only about 20 percent of the firms in their sample used forced rankings in their performance appraisal systems. Forced ratings, just as bonus pools, put employees in direct competition with each other, which may explain why they are not frequently used. Moreover, if supervisors really do not want to distinguish between employees, they can rotate high and low ratings over time, leaving average ratings still compressed (Prendergast and Topel, 1993). Forced ratings also do not guarantee fair ratings: although they impose the desired distribution, they do not preempt favoritism.

Another way to control for bias is to punish its occurrence (Prendergast and Topel, 1993). This is difficult, as supervisors typically have an informational advantage. In addition, when supervisors know that their ratings are being monitored they will probably use their discretion over other matters that influence employee productivity, e.g. job assignment, to justify their ratings (Prendergast and Topel, 1993). Monitoring is therefore not likely to be effective, while it is likely to be costly.

[8] As cited by Prendergast and Topel (1993).

A less expensive alternative would be to allow employees to appeal the supervisor's rating decision. Unfortunately, appeal procedures have been found to be rather ineffective. On the one hand, employees often do not report unfair treatment due to fear of retaliation. On the other hand, management is often reluctant to reverse decisions made by supervisors, as overruling them might unable them to induce the same effort levels in the future.

In the personnel literature, rating training is often mentioned as a way to improve the accuracy of performance ratings. Rating training might make supervisors more aware of their unconscious biases and cognitive limitations but it does not take away their incentives to bias ratings deliberately.

Performance ratings bias is extremely difficult to constrain. None of the methods discussed above seems satisfactory. Consequently, firms may design compensation contracts that are not contingent on performance. Instead of relying on performance measurement, the firm may resort to non-corruptible methods, such as seniority and other bureaucratic rules, to determine pay and promotions. This has clear disadvantages, as incentives and selection criteria will not be set optimally.

V. CONCLUSIONS AND FUTURE RESEARCH OPPORTUNITIES

In section II, we discussed analytical models indicating that the introduction of subjectivity can improve incentive contracting (Baker et al., 1994; Baiman and Rajan, 1995). We also examined empirical research that investigates whether the weight placed on subjective versus objective performance assessment is indeed higher in compensation contracting situations where the advantages of subjectivity are expected to be larger. Although these studies provide us with some valuable insights, research that explains the uses and effects of subjective bonuses is still in an early stage of development. Subjectivity is a complex concept with multiple facets (Ittner et al., 2003); it affects the behavior of both principal and agent in various ways. Nevertheless, most studies only consider a single characteristic of subjectivity, e.g. how subjectivity increases flexibility in appraisal or how it intensifies uncertainty. Moreover, although several ways of introducing subjectivity into compensation contract exist, in general, no distinction is made. Our understanding of the role of subjectivity in incentive contracting would be greatly deepened if the different characteristics of subjectivity and their potential simultaneous presence in contracting were analyzed in greater detail.

Another area of interest is the relationship between subjectivity and other elements of the compensation contract. Usually only a part of the total compensation is subjectively determined (other parts may be either fixed or objectively determined), but the different parts of the compensation contract are likely to be interrelated. Research should therefore not be limited to studying the characteristics of subjective performance assessment in isolation. Understanding whether subjectivity acts as a complement or as a substitute in incentive provision, how the design of the total compensation contract changes when subjectivity is added and which external and internal factors drive these results, would be an important step in the right direction.

We also lack understanding of how the role of subjectivity in compensation contracting differs from the role of subjectivity in promotion and job assignments decisions. Understanding the link between discretion in determining pay and in determining responsibility would greatly increase our comprehension of organizational design and the functioning of control mechanisms.

In section III, we examined the literature on compensation contracts that assumes that the principal acts as the residual claimant. We illustrated the necessary conditions to make compensation pledges credible and discussed several self-enforceable contract structures. These studies analyze how protection of the residual claim drives the principal's behavior but ignore alternative incentives that the principal might have. It is likely that, for example, personal preferences influence the principal's behavior. Incorporating these alternative motives in the economic models would improve our understanding of the complexity of social behavior in compensation contracting.

In section IV, we discussed mainly psychology-based research concerned with subjective performance appraisal. These studies provide consistent evidence indicating that subjective ratings tend to be inflated and compressed, especially when these ratings are used for administrative purposes (Jawahar and Williams, 1997). Although it is important to have an understanding of ratings distributions, it does not help us improve incentive contracting if we do not understand how bias influences the employees that are subject to it. We know little about the perceptions of employees at different positions in the rating distribution. How do better performers who suffer from compressed ratings respond, will they become so disillusioned that they will reduce their effort or leave the firm? How do bad performers that benefit from the bias react, will they decrease effort as compensation seems to be ensured? To understand the implications of subjective performance appraisal we must go beyond studying the effect of subjectivity on ratings to investigating the effects on employee behaviour.

The studies that have analyzed rating bias have predominately focused on purely subjective performance assessment. Not much is known about how the supervisors' behaviour might change when part of the performance is objectively measured. Supervisors might feel more capable to confront employees and hence introduce less bias when their subjective assessment is confirmed by the outcome of objective performance measures.

In conclusion, subjective performance appraisal is an important element of many compensation contracts. Nevertheless, subjectivity has not received a lot of attention in the incentives literature. Subjective performance appraisals by non-residual claimants and compensation contracts that include subjective performance appraisal at lower organizational levels have hardly been investigated. In order to develop a good understanding of the role incentive contracts play in management control systems we must also focus on the subjective elements of compensation contracts. These elements are, however, not easily studied using traditional datasets; management control research should therefore find less traditional ways to tackle these issues (Gibbs et al., 2004).

REFERENCES

Baiman, S. and M. V. Rajan (1995) 'The Informational Advantages of Discretionary Bonus Schemes', *The Accounting Review,* 70(4): 557-579.

Baker, G. P., R. Gibbons and K. J. Murphy (1994) 'Subjective Performance Measures in Optimal Incentive Contracts', *Quarterly Journal of Economics*.

Baker, G. P., M. C. Jensen and K. J. Murphy (1988) 'Compensation and Incentives: Practice vs. Theory', *Journal of Finance,*(3): 593-616.

Banker, R. D. and S. M. Datar (1989) 'Sensitivity, Precision, and Linear Aggregation of Signals for Performance Evaluation', *Journal of Accounting Research,* 27(1): 21-39.

Bernardin, J. H. and R. M. Buckley (1981) 'Strategies in Rater Training', *Academy of Management Review,* 6(2): 205-212.

Bertrand, M. (2004) 'From the Invisible Handshake to the Invisible Hand? How Import Competition Changes the Employment Relationship', *Journal of Labor Economics,* 22(4): 723- 765.

Bretz, R. D., G. T. Milkovich and W. Read (1992) 'The Current State of Performance Appraisal Research and Practice: Concerns, Directions, and Implications', *Journal of Management,* 18(2): 321-352.

Brown, C. (1990) 'Firm's Choice of Method of Pay', *Industrial and Labor Relations Review,* 43: 165-182.

Bull, C. (1983) 'Implicit Contracts in the Absence of Enforcement and Risk Aversion', *The American Economic Review,* 73(4): 658-671.

Bushman, R. M., R. J. Indjejikian and A. Smith (1996) 'CEO compensation: The Role of Individual Performance Evaluation', *Journal of Accounting and Economics,* 21(2): 161-193.

Carmichael, H. L. (1989) 'Self-Enforcing Contracts, Shirking, and Life Cycle Incentives', *Journal of Economic Perspectives,* 3(4): 65-83.

Davila, T. (2003) 'Short-term Economic Incentive in New Product Development', *Research Policy,* 32: 1397-1420.

Feltham, G. A. and J. Xie (1994) 'Performance Measure Congruity and Diversity in Multi-Task Principal/Agent Relations', *The Accounting Review,* 69(3): 429-453.

Ferris, G. R. and T. A. Judge (1991) 'Personnel/Human Resources Management: A Political Influence Perspective', *Journal of Management,* 17(2): 447-488.

Fox, S., A. Bizman and E. Herrmann (1983) 'The Halo Effect: Is it a Unitary Concept?' *Journal of Occupational Psychology,* 56(4): 289-296.

Gibbons, R. (1987) 'Piece-Rate Incentive Schemes', *Journal of Labor Economics,* 5(4): 413-429.

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(2): 409-436.

Harris, M. M. (1994) 'Rater Motivation in the Performance Appraisal Context: A Theoretical Framework', *Journal of Management,* 20(4): 737-756.

Harris, M. M. and J. Schaubroeck (1988) 'A Meta-Analysis of Self-Supervisor, Self-Peer, and Peer-Supervisor Ratings', *Personnel Psychology,* 41: 43 - 62.

Hayes, R. M. and S. Schaefer (2000) 'Implicit Contracts and the Explanatory Power of Top Executive Compensation for Future Performance', *Rand Journal of Economics,* 31(2): 273-293.

Higgins, C. A., T. A. Judge and G. R. Ferris (2003) 'Influence Tactics and Work Outcomes: A Meta-Analysis', *Journal of Organizational Behavior,* 24: 89-106.

Holmstrom, B. (1979) 'Moral Hazard and Observability', *Bell Journal of Economics,* 10(1): 74-91.

Holmstrom, B. and P. Milgrom (1991) 'Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design', *Journal of Law, Economics and Organization,* 7(2): 24-53.

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(3): 725-758.

Jawahar, I. M. and C. R. Williams (1997) 'Where All the Children are Above Average: The Performance Appraisal Purpose Effect', *Personnel Psychology,* 50: 905-926.

Jennings, T., J. K. Palmer and A. Thomas (2004) 'Effects of Performance Context on Processing Speed and Performance Ratings', *Journal of Business and Psychology,* 18(4): 453-463.

Kahn, L. M. and P. D. Sherer (1990) 'Contingent Pay and Managerial Performance', *Industrial and Labor Relations Review,* 43(Special Issue): 107-S - 120-S.

Kaplan, R. S. and D. P. Norton (1996) *The Balanced Scorecard: Translating Strategy into Action*, Boston, Massachusetts: Harvard Business School Press.

Kerr, S. (1975) 'On the Folly of Rewarding A, While Hoping for B', *Academy of Management Journal,* 18(4): 769-783.

Landy, F. and J. Farr (1980) 'Performance Ratings', *Psychological Bulletin,* 87(72-107): 72-107.

Lazear, E. P. (1989) 'Pay Equality and Industrial Politics', *Journal of Political Economy,* 97(3): 561-580.

Lipe, M. G. and S. E. Salterio (2000) 'The Balanced Scorecard: Judgmental Effects of Common and Unique Performance Measures', *The Accounting Review,* 75(3): 283-298.

Longenecker, C. O., H. P. Sims and D. A. Gioia (1987) 'Behind the Mask: The Politic of Employee Appraisal', *The Academy of Management Executive,* 1(3): 183-193.

MacLeod, B. W. and J. M. Malcomson (1989) 'Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment', *Econometrica,* 57(2): 447-480.

MacLeod, B. W. and D. Parent (1999) 'Job Characteristics and the Form of Compensation', *Research in Labor Economics,* 18: 177 - 242.

Milgrom, P. (1988) 'Employment Contracts, Influence Activities, and Efficient Organization Design', *Journal of Political Economy,* 96(1): 42-60.

Moers, F. (2005) 'Discretion and Bias in Performance Evaluation: The Impact of Diversity and Subjectivity', *Accounting, Organizations and Society,* 30(1): 67-80.

Mount, M. K. and D. E. Thompson (1987) 'Cognitive Categorization and Quality of Performance Ratings', *Journal of Applied Psychology,* 72(2): 240-246.

Murphy, K. J. (2000) 'Performance Standards in Incentive Contracts', *Journal of Accounting and Economics,* 30(3): 245-278.

Murphy, K. J. and P. Oyer (2003) 'Discretion in Executive Incentive Contracts: Theory and Evidence', *Working Paper, University of Southern California and Stanford University*.

Murphy, K. R., W. K. Balzer, M. C. Lockhart and E. J. Eisenman (1985) 'Effects of Previous Performance on Evaluations of Present Performance', *Journal of Applied Psychology,* 70(1): 72-84.

Murphy, K. R. and J. Cleveland (1991) *Performance Appraisal: An Organizational Perspective*, Boston: Allyn and Bacon.

Napier, N. K. and G. P. Latham (1986) 'Outcome Expectancies of People who Conduct Performance Appraisals', *Personnel Psychology,* 39(4): 827-837.

Pearce, J. L. and L. W. Porter (1986) 'Employee Responses to Formal Performance Appraisal Feedback', *Journal of Applied Psychology,* 71(2): 211-218.

Portales, C., J. E. Ricart and J. M. Rosanas (1998) 'Understanding Trust to Build Strong Relationships in Organizations', in Hitt, M. A., J. E. Ricart and R. D. Nixons, *Managing Strategically in an Interconnected World*: John Wiley & Sons Ltd.

Prendergast, C. (1999) 'The Provision of Incentives in Firms', *Journal of Economic Literature,* 37(1): 7-63.

Prendergast, C. (2002) 'Uncertainty and Incentives', *Journal of Labor Economics,* 20(2): 115-137.

Prendergast, C. and R. H. Topel (1993) 'Discretion and Bias in Performance Evaluation', *European Economic Review,* 37(2-3): 355-365.

Prendergast, C. and R. H. Topel (1996) 'Favoritism in Organizations', *Journal of Political Economy,* 104(5): 958-978.

Shore, L. M. and G. C. Thornton (1986) 'Effects of Gender on Self-and Supervisory Ratings', *Academy of Management Journal,* 29(1): 115 - 129.

Tremblay, M. and P. Roussel (2001) 'Modelling the Role of Organizational Justice: Effects on Satisfaction and Unionization Propensity of Canadian Managers', *Journal of Human Resource Management,* 12(5): 717-737.

Varma, A., A. S. Denisi and L. H. Peters (1996) 'Interpersonal Affect and Performance Appraisals: A Field Study', *Personnel Psychology,* 49: 341-360.