Performance Evaluation Inflation

Russell Golman and Sudeep Bhatia

March 9, 2011

Department of Social and Decision Sciences, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA. E-mail: rgolman@andrew.cmu.edu; sudeepb@andrew.cmu.edu

Abstract

We show that a surprisingly simple mechanism can account for subjective performance evaluation inflation. When a manager observes noisy signals of employee performance and the manager strives to produce accurate ratings but feels worse about unfavorable errors than about favorable errors, the manager's selfishly optimal ratings will be biased upwards. Both the uncertainty about performance and the asymmetry in the manager's utility are necessary conditions for performance evaluation inflation. Neither factor alone generates any inflation. Moreover, the extent of the bias is increasing in the variance of the performance signal and in the asymmetry in aversion to unfair ratings. This result suggests that performance appraisals based on well-defined unambiguous criteria will have less bias. Additionally, we demonstrate that employer and employee can account for biased performance evaluations when they agree to a contract, and they can use relative performance contracts to do so even when they are unaware of the extent of the bias.

KEYWORDS: inflation, noisy signals, performance appraisal, ratings, relative performance evaluation, subjective performance evaluation

1 Introduction

Subjective performance evaluations play an important role in labor markets. Often times the difficulty of specifying all contingencies in labor contracts requires the additional monitoring of worker performance. In these scenarios performance evaluation incentivizes worker effort [26, 32] and serves as an informational tool for the firm [21]. It is estimated that between 74 to 89% of firms have a formal performance appraisal system [28]. Subjective performance evaluations also play an important role in worker recruitment, with roughly half of all workers finding jobs through external references [24].

However, many researchers have found that subjective performance evaluations suffer from leniency and rating-compression effects [32, 16, 14, 22]. In particular, performance appraisal ratings are shown to display an extreme upward bias, with 60 to 70% of those being assessed rated in the top two categories of rating scales [4]. In a tournament with employees ranked by third-party "referees", this has generated a Lake Wobegon Effect with almost everyone rated above average [25]. Performance evaluations have also been shown to display a centrality bias with supervisors offering ratings that differ little from the norm [32, 23]. Such compression effects seem to be especially prevalent when ratings are a primary determinant of wages. Besides reducing the informational value of performance evaluations, such biases distort wages and can impact worker effort and firm productivity.

Given the prevalence of leniency bias, it has been suggested that managers consciously inflate ratings in order to help workers, improve relations, or avoid conflict [20]. But, we should not necessarily conclude that managers knowingly and intentionally give workers higher ratings than they deserve. We suggest an alternative, more subtle explanation. Job performance measures inevitably suffer from some measurement error [16, 27]. Social and contextual factors influence a manager's perceptions of employee performance [14, 19]. To construct a performance evaluation measure, managers aggregate noisy signals of job performance with prior information about employee competence [2]. Managers may prefer to issue accurate, unbiased ratings, but may introduce bias in the face of imprecise performance measures because of secondary preferences such as aversion to unfairly low ratings.

We show that this surprisingly simple mechanism can account for subjective performance evaluation bias. Bias emerges when the manager is uncertain about worker performance and has asymmetric fairness preferences. Previous work has demonstrated the relevance of fairness preferences held by the employer or employee in the design and implementation of contracts [13, 8, 5]. These preferences impact worker effort in trust and gift-exchange settings [6], and support the use of unenforceable bonus or trust contracts, or incomplete contracts, instead of standard incentive contracts [8, 7]. Here, we hypothesize that managers' fairness considerations support a preference for accurate evaluations and that other considerations, such as employee sympathy or workplace harmony [29, 12], are secondary. Nevertheless, if the worker performance signal contains noise, then managers who feel differently about unfairly positive evaluations and unfairly negative ones will provide biased ratings. If a manager feels worse about unfavorable errors than about favorable errors, then the managers selfishly optimal evaluations will display a leniency bias. This bias is generated despite the fact that the manager strives to produce fair ratings. Moreover, the degree of bias correlates with the noisiness of the performance signal and the asymmetry in the manager's fairness preferences.

After proposing our account for the leniency and compression effects in subjective performance evaluations, we consider the effect of this bias on incentive contracts involving sophisticated principals and agents. In particular, employers and employees are shown to be able to correct for a performance evaluation bias if they are aware of the extent of the bias. In addition, relative performance evaluations can allow the parties to correct for the bias even without information about its extent.

The rest of the paper is organized as follows: Section 2 describes the model of subjective performance evaluation. Section 3 applies this model to derive performance evaluation inflation and compression effects. Section 4 examines the impact of biased evaluations on incentive contracts. Section 5 concludes.

2 A Mathematical Model of Performance Evaluation

A manager is tasked with evaluating a heterogeneous distribution of employees who vary in their levels of competence. For simplicity, assume an employee (arbitrarily, employee i) has true competence $x_i \in \mathbb{R}$ (expressible as a real number). Obviously,

x_i is unknown to the manager, but the manager does know that $x_i \sim N(\bar{x}, \theta^2)$, i.e., that competence is normally distributed in the population with mean $\bar{x}$ and variance θ^2. This knowledge serves as the manager's prior. The manager then observes a signal of the employee's performance $y_i \sim N(x_i, \sigma^2)$. The signal depends of course on the employee's true competence, but has unbiased noise with variance σ^2. Thus, σ captures the uncertainty in the performance measure.

After observing the signal of employee performance, the manager issues a rating $z_i \in \mathbb{R}$ to employee i. We assume the utility of the manager takes the form

$$U(z_i) = \begin{cases} -\lambda(x_i - z_i) & \text{if } z_i < x_i \\ -(z_i - x_i) & \text{otherwise,} \end{cases} \tag{1}$$

with $\lambda > 1$. This reflects a scenario in which the manager would like to issue a rating equal to the employee's true competence, but the manager feels worse about issuing a rating that is undeservedly low than about issuing a rating undeservedly high. Presumably, the manager's primary goal is to assign ratings fairly, but the manager may also sympathize somewhat with the employee or may not want to discourage the employee, or may wish to ingratiate him or herself with the employee. These secondary goals produce an asymmetry in the manager's utility function that is captured by the factor λ. Note that if there was no uncertainty in the performance measure, i.e., if $\sigma = 0$, the manager would know the employee's true competence x_i and would issue a perfectly accurate and unbiased rating $z_i = y_i = x_i$.

If the performance evaluation is used to determine employee compensation in a tournament or in some other incentive contract, then the manager's utility function described by Equation 1 with $\lambda > 1$ is appropriate when the manager is not the residual claimant of the employee's value added. Usually this is the case [33]. Of course, it is just as straightforward to consider a manager who is personally responsible for the compensation package determined by the evaluation. In that case, we might well have $\lambda < 0$ describing a myopic manager who wants to hold down compensation or, if the manager has a reputation to protect or a strong fairness norm to abide by, perhaps $0 < \lambda < 1$. When $\lambda < 1$, the bias we find runs in the opposite direction.

3 Rating Inflation

A manager's rating strategy is a function $\zeta : \mathbb{R} \to \mathbb{R}$ where $z_i = \zeta(y_i)$. The manager updates her belief about the employee's true competence using Bayesian inference. She then chooses a rating contingent on this inference. Her selfishly optimal rating maximizes her utility function.

Theorem 1 *The rule for determining the rating z_i is*

$$\zeta(y_i) = \bar{x} + (y_i - \bar{x})\frac{\theta^2}{\sigma^2 + \theta^2} + \sqrt{2\frac{\sigma^2\theta^2}{\sigma^2 + \theta^2}}\,\mathrm{erf}^{-1}\left(\frac{\lambda - 1}{\lambda + 1}\right). \tag{2}$$

A straightforward proof is in the appendix. Note that $\mathrm{erf}(\cdot)$ is the error function, which is necessary to express the cumulative distribution function of a normal distribution.

The second term in equation (2) depends on the performance signal that the manager observes. The normalization factor of $\frac{\theta^2}{\sigma^2+\theta^2}$ appears because the performance signal is inherently noisy and the manager should discount the magnitude of the signal to account for this noise. The manager knows that the variance of performance signals in the population is $\text{var}(y_i) = \sigma^2 + \theta^2$ whereas the variance of competence is only $\text{var}(x_i) = \theta^2$. Thus, to balance dispersion caused by the noisy signal, an employee's expected rating conditional on his true competence is compressed towards the mean.

The last term in equation (2) is a bias. The manager's best estimate of the employee's true competence after seeing the performance signal is $\bar{x} + (y_i - \bar{x})\frac{\theta^2}{\sigma^2+\theta^2}$, but the manager adds into the rating this additional term that is positive for $\lambda > 1$. (In cases where the manager feels worse about undeserved high ratings than undue low ones, we have $\lambda < 1$, and we predict rating deflation. This might occur if ratings determine compensation paid directly out of the manager's pocket and the manager's preferences reflect a strong fairness norm along with some degree of self-interest.) The size of this bias is increasing in σ, in θ, and in λ.

We thus obtain the following predictions (for $\lambda > 1$):

Corollary 1 *An employee's expected rating, conditional on his true competence x_i, is increasing linearly in his competence, with compression towards the mean $\bar{x}$ and leniency bias.*

Corollary 2 *The average rating in the population exceeds average competence and is increasing in signal noisiness σ, in employee heterogeneity θ, and in preference* *asymmetry λ.*

Corollary 3 *The fraction of the population rated above average is greater than one half and is increasing in signal noisiness σ and in preference asymmetry λ, but decreasing in employee heterogeneity θ.*

4 Incentive Contracts

When offering an incentive contract with compensation based on a manager's subjective evaluation, a sophisticated employer accounts for the manager's biased ratings. The employer (the principal) utilizes an incentive contract to align the incentives of an employee (the agent) facing moral hazard in deciding how much effort to exert on the job. When effort is not observable objectively, it is not directly contractible, and an additional agent (the manager) may be tasked with evaluating the employee's performance. Employees may vary in their ability, and it may be impossible to distinguish ability from effort generally.

Now, suppose employee (i's) competence $x_i \in \mathbb{R}$ is the (weighted) sum of ability and effort, $x_i = a_i + \rho e_i$. Ability is normally distributed in the population with mean $\bar{a}$ and variance θ^2, i.e., $a_i \sim N(\bar{a}, \theta^2)$. Effort $e_i \in \mathbb{R}$ is a choice variable for the employee. As in Section 2, the manager observes a noisy signal of the employee's performance $y_i \sim N(x_i, \sigma^2)$ and issues a rating $z_i \in \mathbb{R}$ to maximize utility given by Equation (1).

The rating z_i is contractible performance measure, whereas neither competence nor effort is contractible, and their effect on total firm value is too diffuse to be a useful measure. A contract between the employer and the employee will specify the wage as a function of the manager's rating, $w_i = f(z_i)$. For illustration, we suppose the value to the firm of the employee's contributions is $V(x_i) = e^{kx_i}$. We adopt this exponential functional form because it seems reasonable that value created is convex in competence, as the marginal productivity of effort should be increasing in ability, and because it guarantees that the employee's value is positive. We suppose employee utility is an additively separable function of wealth and effort exertion. We consider risk averse employees with Bernoulli utility for wealth $\ln(w_i)$ satisfying constant relative risk aversion. The cost of effort $c(e)$ is increasing and convex, with $lim_{e \to e_{\min}} c'(e) = 0$, $lim_{e \to e_{\max}} c'(e) = \infty$, and $c''(e) > 0$ for all e. Thus, $U_E(w_i, e_i) = \ln(w_i) - c(e_i)$. The employer's profit is $\Pi(x_i, w_i) = V(x_i) - w_i$.

We assume the employee (and obviously the employer as well) does not know his own ability when agreeing to a contract (as in [35]). The employee then discovers his type after agreeing to a contract, but before deciding how much effort to exert on the job. If agents knew their type before agreeing to a contract, there would be adverse selection in choosing from a menu of contracts, with low-ability types trying to imitate high-ability types and high-ability types trying to distinguish themselves (c.f. [3, 18, 31, 34]). It would be efficient for employees to sort themselves and avoid exposure to the uncertainty surrounding their true ability, and we expect employees would obtain credentials to signal their ability [17]. As we are interested in retaining heterogeneity in employee performance, we consider the case in which sorting contracts by ability is impossible. While we would generally assume that agents know their own type when one's type determines one's preferences, in models in which one's type refers to one's quality, it is quite reasonable to assume a lack of self-knowledge [15].

We consider contracts of the form $w_i = \alpha e^{\beta z_i}$ with $\alpha \geq 0$ and $\beta \geq 0$. We suppose that the labor market is competitive and there is just a single employer. In equilibrium employees are indifferent between accepting the contract or taking an outside option with utility normalized to 0. The employer offers the contract that maximizes its profit subject to this constraint.

Theorem 2 *In equilibrium, restricting to contracts of the form $w_i = \alpha e^{\beta z_i}$, the employer offers (and the employee accepts) a contract with*

$$\beta = \arg\max_{\hat{\beta} \geq 0} \left\{ e^{k\left(\bar{a} + \frac{\theta^2}{2} + \rho (c')^{-1}\left(\hat{\beta}\rho\frac{\theta^2}{\sigma^2+\theta^2}\right)\right)} - e^{c\left((c')^{-1}\left(\hat{\beta}\rho\frac{\theta^2}{\sigma^2+\theta^2}\right)\right) + \hat{\beta}\frac{\theta^2}{2}} \right\} \tag{3}$$

and

$$\alpha = \exp\left(c(e^*) - \beta \left[\bar{a} + \rho e^* + \sqrt{2\frac{\sigma^2\theta^2}{\sigma^2+\theta^2}} \operatorname{erf}^{-1}\left(\frac{\lambda - 1}{\lambda + 1}\right) \right] \right). \tag{4}$$

All employees exert effort $e^* = (c')^{-1}\left(\beta\rho\frac{\theta^2}{\sigma^2+\theta^2}\right)$. The optimal effort level is independent of ability. Employee competence is normally distributed, $x_i \sim N(\bar{x}, \theta^2)$, with mean $\bar{x} = \bar{a} + \rho e^*$. The manager's rating function for determining z_i is given by Equation (2).

The proof, which relies on backward induction, is in the appendix.

The employer and employee can account for bias in the manager's performance rating when they agree to a contract, but doing so requires a priori knowledge of the extent of this bias. In particular, the contract parameter α depends on λ, which specifies the manager's sympathies for the employee. In general, λ may not be observable to the employer or the employee at the time of contracting. In such situations, the parties can use a relative performance contract to get around this lack of information.

Observe that the contract $w_i = \tilde{\alpha} e^{\beta(z_i - \bar{z})}$ is equivalent to $w_i = \alpha e^{\beta z_i}$ for $\tilde{\alpha} = \alpha e^{\beta \bar{z}}$. The average rating given by the manager is $\bar{z} = \bar{x} + \sqrt{2\frac{\sigma^2\theta^2}{\sigma^2+\theta^2}}\,\mathrm{erf}^{-1}\left(\frac{\lambda-1}{\lambda+1}\right)$. Thus, we obtain the following corollary:

Corollary 4 *In equilibrium, restricting to contracts of the form $w_i = \tilde{\alpha} e^{\beta(z_i - \bar{z})}$, the employer offers (and the employee accepts) a contract with β as in Equation (3) and $\tilde{\alpha} = e^{c(e^*)}$. The manager's ratings, employee effort, and compensation are the same as in Theorem 2.*

Note that the parameters of the relative performance contract described in Corollary 4 do not depend on λ. Thus, if λ is unknown, the employer can just as well offer a relative performance contract with $\tilde{\alpha} = e^{c(e^*)}$ and β as in Equation (3).

5 Conclusion

Noise in the performance signal and a stronger aversion to unfairly low ratings than to overly high ones together bring a manager to inflate performance evaluation ratings. We need not assume the manager desires an inflated profile of ratings. The manager may well wish to have accurate, unbiased ratings, but if noisy signals are inevitable, the manager may still introduce an upward bias to counteract the inherent imprecision of the performance signal. Both the amount of noise in measuring performance and the degree of asymmetry in preferences over ratings error contribute to the size of the bias in the manager's selfishly optimal rating. This suggests that extreme leniency in performance evaluation can be mitigated by defining more concrete, unambiguous evaluation criteria.

Relative performance evaluation may be part of an optimal contract for determining employee compensation as it provides incentives for risk-averse employees to perform well even when absolute performance signals are noisy [11, 10, 17, 30]. Sometimes a performance signal directly determines compensation, but often a subjective judgment of employee performance must be made [1]. In such cases, a manager's incentives will influence the rating given to the employee [33]. An incentive compatible mechanism for the manager to report unbiased performance signals must go beyond simply creating a preference for accurate ratings. A manager with any other-regarding preferences will still introduce bias into the performance evaluation to the extent the performance signal is imprecise and noisy. Use of subjective performance evaluation in determining compensation thus depends on identifying clear and concrete performance standards. The bias introduced into a subjective performance evaluation can be accounted for by a sophisticated employer or employee when agreeing to a contract determining compensation. When the manager's sympathies cannot be precisely identified a priori, a cardinal tournament can mimic a traditional contract without requiring this information. Thus, relative subjective performance evaluation appears to be more robust to incomplete information than traditional contracts based only on subjective evaluation of a single agent's performance.

Appendix

Proof of Theorem 1

Straightforward application of Bayes' Law yields the posterior density $p(x_i|y_i) \sim N\big(\bar{x} + (y_i - \bar{x})\frac{\theta^2}{\sigma^2+\theta^2}, \frac{\sigma^2\theta^2}{\sigma^2+\theta^2}\big)$ (see [9], pg. 46). For a given y_i, the expected utility from rating z_i is

$$E[U(z_i)] = \int_{-\infty}^{z_i} -(z_i - x_i)\, p(x_i|y_i)\, dx_i \;+\; \int_{z_i}^{\infty} -\lambda(x_i - z_i)\, p(x_i|y_i)\, dx_i.$$

The first order condition for a selfishly optimal rating is then

$$\begin{aligned} \frac{d}{dz_i}E[U(z_i)] \;&=\; -\int_{-\infty}^{z_i} p(x_i|y_i)\, dx_i \;+\; \lambda \int_{z_i}^{\infty} p(x_i|y_i)\, dx_i \\ &=\; \lambda - (\lambda+1)\,\frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{z_i - \bar{x} - (y_i - \bar{x})\frac{\theta^2}{\sigma^2+\theta^2}}{\sqrt{2\frac{\sigma^2\theta^2}{\sigma^2+\theta^2}}}\right)\right] \;=\; 0. \end{aligned}$$

Equation (2) is found by inverting to solve for z_i. ∎

Proof of Corollary 1

Integrate over the performance signal to find that an employee's expected rating conditional on his true competence x_i is

$$E\left[\zeta(y_i)\,|\,x_i\right] = \;\bar{x} + (x_i - \bar{x})\frac{\theta^2}{\sigma^2+\theta^2} \;+\; \sqrt{2\frac{\sigma^2\theta^2}{\sigma^2+\theta^2}}\;\mathrm{erf}^{-1}\left(\frac{\lambda-1}{\lambda+1}\right). \qquad \blacksquare$$

Proof of Corollary 2

Integrating over the performance signal and the competence level reveals that the average rating in the population is $\bar{x} + \sqrt{2\frac{\sigma^2\theta^2}{\sigma^2+\theta^2}}\;\mathrm{erf}^{-1}\left(\frac{\lambda-1}{\lambda+1}\right)$. ∎

Proof of Corollary 3

As there is a normal distribution of performance signals across the population, the fraction of employees rated above average is

$$\Pr(z_i > 0) \;=\; \Phi\left(\sqrt{2}\,\frac{\sigma}{\theta}\;\operatorname{erf}^{-1}\left(\frac{\lambda-1}{\lambda+1}\right)\right) \;=\; \frac{1}{2}\left[1+\operatorname{erf}\left(\frac{\sigma}{\theta}\;\operatorname{erf}^{-1}\left(\frac{\lambda-1}{\lambda+1}\right)\right)\right]. \quad \blacksquare$$

Proof of Theorem 2

Given that $x_i \sim N(\bar{x}, \theta^2)$, we obtain the manager's rating funtion in Theorem 1.

Given the contract parameters and the manager's rating function, an employee with ability a_i chooses effort e_i (and thus competence $x_i = a_i + \rho e_i$) to maximize $E\left[\ln\left(\alpha e^{\beta\zeta(y_i)}\right)\right] - c(e_i)$ where y_i is of course a stochastic function with mean x_i. The first order condition is then $\beta\rho\frac{\theta^2}{\sigma^2+\theta^2} = c'(e_i)$. The convexity of $c(\cdot)$ implies this is indeed a maximum of expected utility, and the range of $c'(\cdot)$ from 0 to ∞ guarantees that there is a solution: $e_i = (c')^{-1}\left(\beta\rho\frac{\theta^2}{\sigma^2+\theta^2}\right)$ for all i. We denote this equilibrium effort level e^*. Because ability is normally distributed across employees and all employees choose the same effort level regardless of ability, competence is then also normally distributed.

An employee's expected utility if he agrees to the contract (not knowing his own ability) is

$$\overline{U}_E = \ln\left(\alpha e^{\beta\left(\bar{x}+\sqrt{2\frac{\sigma^2\theta^2}{\sigma^2+\theta^2}}\;\operatorname{erf}^{-1}\left(\frac{\lambda-1}{\lambda+1}\right)\right)}\right) - c(e^*),$$

after averaging over his ability and the signal the manager receives. In a competitive labor market with an outside option yielding 0 utility, $\overline{U}_E = 0$. Thus,

$$\ln(\alpha) + \beta \left(\bar{x} + \sqrt{2\frac{\sigma^2\theta^2}{\sigma^2+\theta^2}}\ \mathrm{erf}^{-1}\left(\frac{\lambda-1}{\lambda+1}\right)\right) - c(e^*) = 0.$$

Solving for α yields Equation (4).

The employer determines β (and implicitly α) to maximize expected profit. We integrate with respect to the cumulative distribution functions $P_{y|x}(y_i|x_i)$ and $P_a(a_i)$ for the manager's signal given employee competence and for employee ability to find expected profit:

$$\begin{aligned}
\overline{\Pi}(\beta) = & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{k(a_i+\rho e^*(\beta))} \ - \ \alpha(\beta)\, e^{\beta\left((y_i-\bar{x})\frac{\theta^2}{\sigma^2+\theta^2}+\bar{x}+\sqrt{2\frac{\sigma^2\theta^2}{\sigma^2+\theta^2}}\ \mathrm{erf}^{-1}\left(\frac{\lambda-1}{\lambda+1}\right)\right)}\ dP_{y|x}(y_i|a_i+\rho e^*)\, dP_a(a_i) \\
= & e^{k\left(\bar{a}+\frac{\theta^2}{2}+\rho e^*(\beta)\right)} \ - \ \alpha(\beta)\, e^{\beta\left((\frac{\sigma^2}{2}+\frac{\theta^2}{2})\frac{\theta^2}{\sigma^2+\theta^2}+\bar{x}+\sqrt{2\frac{\sigma^2\theta^2}{\sigma^2+\theta^2}}\ \mathrm{erf}^{-1}\left(\frac{\lambda-1}{\lambda+1}\right)\right)} \\
= & e^{k\left(\bar{a}+\frac{\theta^2}{2}+\rho e^*(\beta)\right)} \ - \ \alpha(\beta)\, e^{\beta\left(\frac{\theta^2}{2}+\bar{x}+\sqrt{2\frac{\sigma^2\theta^2}{\sigma^2+\theta^2}}\ \mathrm{erf}^{-1}\left(\frac{\lambda-1}{\lambda+1}\right)\right)}.
\end{aligned}$$

Plugging in for the functions $e^*(\beta)$ and $\alpha(\beta)$, we have

$$\overline{\Pi}(\beta) = e^{k\left(\bar{a}+\frac{\theta^2}{2}+\rho(c')^{-1}\left(\beta\rho\frac{\theta^2}{\sigma^2+\theta^2}\right)\right)} \ - \ e^{c\left((c')^{-1}\left(\beta\rho\frac{\theta^2}{\sigma^2+\theta^2}\right)\right)+\beta\frac{\theta^2}{2}}.$$

To guarantee $\overline{\Pi}(\beta)$ attains a maximum on $\beta \in [0,\infty)$, we show that $\lim_{\beta\to\infty}\overline{\Pi}'(\beta) < 0$. The derivative is

$$\overline{\Pi}'(\beta) = e^{k\left(\bar{a}+\frac{\theta^2}{2}+\rho e^*(\beta)\right)}\left(k\rho^2\frac{\theta^2}{\sigma^2+\theta^2}\frac{1}{c''(e^*(\beta))}\right) - e^{c(e^*(\beta))+\beta\frac{\theta^2}{2}}\left(\beta\rho^2\left(\frac{\theta^2}{\sigma^2+\theta^2}\right)^2\frac{1}{c''(e^*(\beta))}+\frac{\theta^2}{2}\right).$$

Because $c(\cdot)$ is convex, $c(e^*(\beta))$ grows faster than $e^*(\beta)$ when β is large. The second term, carrying a minus sign, dominates. Thus, Equation (3) is well-defined. ∎

References

[1] Baker, G., Gibbons, R., Murphy, K. (1994). Subjective Performance Measures in Optimal Incentive Contracts. *The Quarterly Journal of Economics* 109 (4), 1125-1156.

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

[3] Bhattacharaya, S., Guasch, J. (1988). Heterogeneity, Tournaments, and Hierarchies. *Journal of Political Economy* 96, 867-881.

[4] Bretz, R., Milkovich, G., Read, W. (1992). The Current State of Performance Appraisal Research and Practice: Concerns, Directions, and Implications. *Journal of Management* 18 (2), 321-352.

[5] Englmaier, F., Wambach, A. (2010). Optimal incentive contracts under inequality aversion. *Games and Economic Behavior* 69 (2), 312-328.

[6] Fehr, E., Kirchler, E., Weichbold, A., Gachter, S. (1998). When Social Norms Overpower Competition: Gift Exchange in Experimental Labor Markets. *Journal of Labor Economics* 16 (2), 324-351.

[7] Fehr, E., Klein, A., Schmidt, K.M. (2007). Fairness and Contract Design. *Econometrica* 75 (1), 121-154.

[8] Fehr, E., Schmidt, K.M. (2007). Adding a stick to the carrot? The interaction of bonuses and fines. *American Economic Review* 97 (2), 177-181.

[9] Gelman, A., Carlin, J., Stern, H., Rubin, D. (2004). *Bayesian Data Analysis.* Second Edition. Boca Raton, FL: Chapman and Hall / CRC Press.

[10] Gibbons, R., Murphy, K. (1990). Relative Performance Evaluation for Chief Executive Officers. *Industrial and Labor Relations Review* 43 (3), 30S-51S.

[11] Green, J., Stokey, N. (1983). A Comparison of Tournaments and Contracts. *Journal of Political Economy* 91, 349-364.

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

[13] Herold, F. (2010). Contractual incompleteness as a signal of trust. *Games and Economic Behavior* 68 (1), 180-191.

[14] Judge, T., Ferris, G. (1993). Social Context of Performance Evaluation Decisions. *Academy of Management Journal* 36 (1), 80-105.

[15] Kruger, J., Dunning, D. (1999). Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments. *Journal of Personality and Social Psychology* 77 (6), 1121-1134.

[16] Landy, F., Farr, J. (1980). Performance Rating. *Psychological Bulletin* 87 (1), 72-107.

[17] Lazear, E., Rosen, S. (1981). Rank-Order Tournaments as Optimum Labor Contracts. *The Journal of Political Economy* 89 (5), 841-864.

[18] Levy, A., Vukina, T. (2002). Optimal Linear Contracts with Heterogeneous Agents. *European Review of Agricultural Economics* 29 (2), 205-217.

[19] Levy, P., Williams, J. (2004). The Social Context of Performance Appraisal: A Review and Framework for the Future. *Journal of Management* 30 (6), 881-905.

[20] Longenecker, C., Sims, H., Gioia, D. (1987). Behind the Mask: The Politics of Employee Appraisal. *Academy of Management Executive* 1 (3), 183-193.

[21] Meyer, J., Allen, N. (1997). *Commitment in the Workplace*. Thousand Oaks, CA: Sage Publications.

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

[23] Mohrman, A., Lawler, E. (1991). *Pay for Performance: Evaluating Performance Appraisal and Merit Pay*. Washington: National Academy Press.

[24] Montgomery, J.D. (1991). Social Networks and Labor Market Outcomes. *American Economic Review* 81 (5), 1407-1418.

[25] Moran, J., Morgan, J. (2003). Employee Recruiting and the Lake Wobegon Effect. *Journal of Economic Behavior & Organization* 50 (2), 165-182.

[26] Murphy, K. (1999). Executive Compensation. *Handbook of Labor Economics* 3, 2485-2563.

[27] Murphy, K. (2008). Explaining the Weak Relationship Between Job Performance and Ratings of Job Performance. *Industrial and Organizational Psychology* 1 (2), 148-160.

[28] Murphy, K., Cleveland, J. (1991). *Performance Appraisal: An Organizational Perspective*. Needham Heights, MA: Allyn & Bacon.

[29] Murphy, K., Cleveland, J., Skattebo, A., Kinney, T. (2004). Raters Who Pursue Different Goals Give Different Ratings. *Journal of Applied Psychology* 89, 158-164.

[30] Nalebuff, B., Stiglitz, J. (1983). Prizes and Incentives: Towards a General Theory of Compensation and Competition. *Bell Journal of Economics* 14, 21-43.

[31] O'Keeffe, M., Viscusi, W., Zeckhauser, R. (1984). Economic Contests: Comparative Rewards Scheme. *Journal of Labor Economics* 2, 27-56.

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

[33] Prendergast, C., Topel, R. (1993). Discretion and Bias in Performance Evaluation. *European Economic Review* 37, 355-365.

[34] Riis, C. (2010). Efficient Contests. *Journal of Economics & Management Strategy* 19 (3), 643-665.

[35] Tsoulouhas, T., Marinakis, K. (2007). Tournaments with Ex Post Heterogeneous Agents. *Economics Bulletin* 4 (41), 1-9.