Collective vs. Relative Performance Evaluation with Career Concerns*

Huseyin Yildirim

Department of Economics
Duke University
Box 90097
Durham, NC 27708

E-mail: hy12@duke.edu

July 25, 2024

Abstract

Contract theory generally prescribes Relative Performance Evaluation (RPE) to filter out similar uncertainties and induce competition among agents. This paper argues that the RPE logic fails when agents are implicitly motivated by career concerns, such as the prospect of future employment and promotions based on perceived talents. This is because RPE can also filter out agents' similar talents they aim to demonstrate. Collective Performance Evaluation can motivate career-concerned agents better by positively tying their reputations. The paper characterizes the optimal performance evaluations and how they vary with agents' prior reputations. The findings can explain the RPE ban in government agencies.

Keywords: Career concerns, performance ratings, relative performance evaluation, collective performance evaluation

JEL Classification: D20, D86, J24, L23, M12

*I thank the participants of the 2024 Bosphorus Workshop on Economic Design and the Duke Theory Lunch for their comments. I also thank Zhiyuan Jia, Xuerong Wang, and Runzhe (Bill) Zhao for research assistance. All errors are mine.

1 Introduction

A key design issue in multi-agent organizations is whether to provide agents with team-based or competitive incentives. A robust prediction of contract theory is that agents are best motivated by competition when pursuing projects subject to similar uncertainties; see Fleckinger, Martimort, and Roux (Forthcoming) for a recent review. In these environments, incentive contracts typically feature Relative Performance Evaluation (RPE), such as a rank-order tournament, that compares one's performance with peers' to filter out systematic productivity shocks beyond the agents' control.

This paper argues that the logic behind RPE largely fails if agents are driven mainly by career concerns à la Holmström (1999). Government employees are prominent examples. As Dewatripont, Jewitt, and Tirole (1999) noted, "... implicit incentives, in the form of career concerns, inside or outside the organization... play a key role in all organizations, private and public, but are particularly strong in the government sector, where formal incentives schemes are often crude and constrained. In this sector, elections, promotions, and future employment in the private sector are major motivations to expend effort in the current job." To illustrate why RPE may not motivate career-concerned agents, consider a manager who hires two new employees with (ex-post) identical yet unknown innate abilities, even to themselves.[1] Each employee is assigned an unrelated task whose outcome is the sum of his ability, effort, and luck. Hence, the unknown ability is the common shock to the employees' performances. The manager offers a preset salary but promises to report the difference between their performances in their future job recommendations as an incentive. Then, the employees would exert *no* costly effort since such RPE would wash out, rather than reflect, their abilities. In contrast, the manager can encourage effort by promising to report their individual or collective performance, each correlating with their abilities. How exactly the manager should aggregate and disclose performances to motivate career-concerned agents is the central question of this paper.

I address this question within a "static" version of Holmström's (1999) classic model. As in the example above, the "principal" or "organization" employs two agents and assigns each a separate project. An agent's performance depends additively on his intrinsic ability or talent, effort, and independent noise. Talents are positively correlated or similar, perhaps because

[1]Another example closer to home would be a professor hiring two master's students as research assistants who plan to attend PhD programs.

of the agents' prior education or the organization's hiring standards. Typical in this context, talents are assumed to be unknown, even to the agents, ruling out any signaling motive. The organization seeks to motivate agents by performance ratings that (linearly) aggregate their outputs. In turn, agents care about what the market will expect of their talents based on disclosed ratings.

To fix ideas, I first consider mutually exclusive markets for agents; perhaps they target different industries, geographies or schools in the next stage of their careers. I show that if the principal commits to disclosing *all* outputs to every market, agents with more similar talents work *less*. This highlights a positive reputational externality that agents ignore when choosing effort. The principal can eliminate this externality by promising to mention each performance only to its relevant market. Such partial disclosure provides incentives equivalent to a single-agent setting.

Yet, the principal can do even better by careful performance ratings. I find that if an agent is sufficiently known by the market in the sense of having a low prior (talent) variance, the optimal rating provides him with team incentives: the agent's rating improves with his peer's performance.[2] To understand, note that an agent with virtually known talent would have little career incentive of his own. To motivate, the principal links his reputation to his similarly talented peer. In general, a positive weight on the peer's performance in one's rating has two opposite incentive effects. While it increases the total exogenous variation in the rating, discouraging the agent, it also increases how much of this variation can be attributed to talents, encouraging him. For a low-variance agent, the attribution effect dominates. By the same logic, the principal provides competitive incentives or RPE to an agent sufficiently unknown by the market: his rating worsens with the peer's performance. In this case, the principal worries that team-based incentives would cause too much variation in the agent's rating. Overall, agents with different prior reputations may be promised different types of performance evaluations: one giving team incentives and the other inducing competition. Regardless, and in sharp contrast to full output disclosure, agents with more correlated talents work *harder* under optimal performance ratings. Hence, if she could, the principal would hire employees with identical rather than diverse backgrounds.

In various applications, however, the principal cannot prevent all her ratings from becoming public. For one, employees may aim for the same future jobs, enabling the recruiters

[2] In practice, the manager can highlight the employees' collective achievements rather than individual performance in job recommendations.

to compare their evaluations and deduce individual outputs. Public ratings thus constrain the principal in improving incentives upon the full output disclosure. I show that the principal must give the agents the *same* rating (or its positive scales) to obscure individual performances. Hence, public ratings that dominate full output disclosure must provide team incentives, ruling out RPE. In the concluding section, I present anecdotal evidence for this finding from both the public and private sectors.

Finally, I also extend the analysis to positively correlated noises that may arise due to similar working conditions. Consistent with contract theory, noise correlation favors competitive incentives since it would otherwise amplify the exogenous variation in the rating, reducing talent's role in it. Therefore, when incentives are implicit, it is crucial for the principal to filter out the right common shock (talent vs. noise) in her performance evaluations.

Related Literature. Holmström (1979) seminally proved the following sufficient statistic result: additional performance measures should be exploited in the incentive contract if and only if they contain valuable information about the agent's effort. In multi-agent situations, this result implies competitive incentives or RPE if agents' outputs are technologically independent but prone to common productivity shocks, as shown by Lazear and Rosen (1981), Holmström (1982), Green and Stokey (1983), Nalebuff and Stiglitz (1983), Mookherjee (1984), and Gromb and Martimort (2007), among others. For empirical evidence on RPE, see, for instance, Gibbons and Murphy (1990), Bloomfield, Marvao, and Spagnolo (2023) and the references therein.

In richer environments, however, researchers have also established the optimality of team-based contracts, most notably when: (1) agents can monitor peers' efforts (e.g., Ma, 1988; Miller, 1997; Che and Yoo, 2001), (2) they perform complementary tasks (e.g., Itoh, 1991; Legros and Matthews, 1993), or (3) their performances correlate more strongly with higher efforts (Fleckinger, 2012). Fleckinger et al. (Forthcoming) offer an enlightening recent review of multi-agent contracts under moral hazard.[3] The above-mentioned features favoring team-based incentives are absent in my setup. More importantly, I do not allow for explicit incentive contracts.

Instead, following Holmström (1999), agents in my model are implicitly driven by career concerns (see also Dewatripont et al. 1999). Within this framework, several authors, including Meyer (1994), Auriol, Friebel, Pechlivanos (2002), Ortega (2003), Arya and Mittendorf

[3] These authors also demonstrate that the rationale behind RPE does not require risk-averse agents but the presence of an agency problem.

(2011), Chalioti (2016), Meagher and Prasad (2016) and Yildirim (2024), have examined team incentives where, unlike here, what the market can observe about agents' performances, often the total output, is exogenously fixed. Nevertheless, these authors have also noticed that career-concerned agents may work harder in teams despite blurred individual performance.[4]

This paper also relates to those on the optimal performance rating systems, the closest being Rodina (2017), Hörner and Lambert (2021), and Yildirim (2024), as they employ standard career concerns models. Hörner and Lambert find that a career-concerned agent is sometimes best motivated by negatively weighing his past performance in the current rating. Their observation is akin to that in Section 4, except that one's performance is benchmarked against the peer's in my "static" multi-agent setting. Rodina considers a more general information design problem and shows that more uncertainty about the agent's ability may create stronger incentives. With multiple agents, uncertainty in ability is amplified by disclosing an aggregation of outputs. In an extension, Yildirim (2024) delves into the disclosure issue addressed in Section 4 but assumes symmetric agents and offers no comparison to output disclosure or public ratings, which are central to this investigation.[5]

Finally, assuming symmetric agents, Meyer and Vickers (1997) and Fleckinger et al. (Forthcoming, Section 7.1) pose a parallel question to this paper and conclude that career concerns favor individual performance evaluation when talents are positively correlated. Their observation corresponds to my output disclosure benchmarks (see Lemma 2). However, I show that the organization can generally do better by disclosing an aggregation of outputs, i.e., performance rating, to the market. In turn, a positive correlation between agents' talents *helps*, rather than hurts, incentives.

The remainder of the paper is organized as follows. Section 2 introduces the model. Section 3 performs a benchmark analysis. Sections 4 and 5 characterize optimal performance ratings when confidential and public, respectively. Section 5 discusses anecdotal evidence and concludes.

[4]See also Bar-Isaac (2007) and Neeman, Öry, and Yu (2019) for a similar insight.

[5]More tangentially, this paper also relates to those that argue how limited transparency in the trade history of a privately informed agent may encourage reputation building (e.g., Ekmekci, 2011; Liu and Skrzypacz, 2014; Pei, 2023).

2 Model

A "principal" or "organization" employs two agents indexed by $i \in \{1,2\}$ to work on independent projects. Agent i exerts unobservable effort $x_i \geq 0$ at a personal cost $c(x_i)$, satisfying

$$c', c'' > 0, \text{ with } c(0) = c'(0) = 0 \text{ and } c'(\infty) > 1. \tag{C1}$$

(C1) will guarantee interior equilibrium efforts. Later, I will also add $c''' \geq 0$ for this purpose. As in Holmström (1999), agent i's output is given by:

$$y_i = \eta_i + x_i + \varepsilon_i, \tag{1}$$

where η_i and ε_i denote his talent and exogenous noise, respectively. The additive technology is particularly useful here to ensure that the correlation between agents' outputs does not depend on their efforts, which is often another reason to link their incentives.

Agents choose their efforts simultaneously and only once. Standard in the career-concerns models, they do so uncertain of their talents and noise terms, ruling out any signaling motive. Also standard is that all players believe them to be jointly normal with the marginal distributions:

$$\eta_i \sim N(\mu_i, \sigma_i^2) \text{ and } \varepsilon_i \sim N(0, \sigma_\varepsilon^2),$$

where $\sigma_i, \sigma_\varepsilon > 0$, and **(1)** η_1 and η_2 are positively correlated with $Corr(\eta_1, \eta_2) = \rho \in (0,1]$,[6] **(2)** η_i is independent of ε_1 and ε_2, and **(3)** ε_1 and ε_2 are independent. Agents may have correlated talents because of their education or the principal's selection process.[7] For ease of reference, I will call an agent *less known* if he has a *higher talent variance*, and vice versa.

Once efforts are chosen, the principal observes the output vector $\mathbf{y} = (y_1, y_2)$ but cannot verifiably contract on it. Instead, she promises agent i a fixed wage normalized to zero and disclosure of the following performance rating to the market:

$$r_i(\mathbf{y}) = y_i + \lambda_i y_{-i}, \tag{2}$$

where $\lambda_i \in \mathbb{R}$, and the unit coefficient for y_i is without loss. Like linear agency contracts, linear aggregation rules lend significant tractability under normal priors.[8] Using the classifi-

[6] I assume $\rho > 0$ to streamline the exposition, but the results would symmetrically hold for $\rho < 0$.

[7] Agents may also have correlated noises due to similar working conditions, which I consider in Section 6.

[8] Recall that a vector of random variables is jointly normal if and only if any linear combination is normally distributed.

cation by Fleckinger et al. (Forthcoming), I say that agent i faces

$$\begin{cases} \text{Collective Performance Evaluation (CPE)} & \text{if} \quad \lambda_i > 0, \\ \text{Relative Performance Evaluation (RPE)} & \text{if} \quad \lambda_i < 0, \\ \text{Independent Performance Evaluation (IPE)} & \text{if} \quad \lambda_i = 0. \end{cases}$$

In words, agent i's rating improves with the peer's performance under CPE but worsens under RPE. The principal publicly commits to (λ_1, λ_2) at the outset and that she cannot misrepresent the agents' outputs in her ratings.[9] Until Section 5, agents are assumed to enter separate markets after their work for the principal. For instance, the employees may pursue different industries, and the research assistants may apply to distinct PhD programs in the next phase of their careers. Therefore, the principal can disclose to agent i's exclusive market only his rating r_i or both r_i and r_{-i}.

In terms of the payoffs, the principal cares about the total output, $y_1 + y_2$. Each agent cares about the market's expectation of his talent conditional on disclosed rating(s) and the effort cost.

Remark 1 *Technically, the current setup is a two-period version of Holmström (1999). Because agents would exert no effort in the last period, their future expected outputs in the first period, which become their competitive market offers, would reduce to their expected talents. I omit this detail for brevity.*

Remark 2 *The principal placing some weight on the market's ex-post beliefs about the agents would have no qualitative effect on my results: thanks to the law of iterated expectations, the expectation of those beliefs would always equal the prior means, μ_i, in equilibrium, regardless of the principal's choices.*

3 Benchmarks

I first present three benchmarks to fix ideas for optimal disclosure. For notational convenience below, let

$$\psi \equiv c'^{-1}, \tag{3}$$

denote the inverse of marginal cost where $\psi' > 0$, $\psi(0) = 0$ and $\psi(1) < \infty$ by (C1).

[9]Organizations may achieve such commitment by delegating performance ratings to their HR departments with clear guidelines; see, for instance, the 2017 handbook published by the United States Office of Personnel Management (https://www.opm.gov/policy-data-oversight/performance-management/measuring/employee_performance_handbook.pdf). Public commitment to rating methodology is less evident for educators writing recommendation letters. However, given the little tension between the principal and the market, it will play little role here.

3.1 First-best

Suppose a planner could dictate the agents' effort levels to maximize the total expected output net of the effort costs. Then, the planner would solve

$$\max_{x_1,x_2} E[y_1 + y_2] - c(x_1) - c(x_2),$$

which, by using (1), reveals

$$x_i^{FB} = \psi(1). \tag{4}$$

Absent incentive issues, the first-best effort is independent of prior distributions and the same for both agents as it equates the unit marginal return of effort to its marginal cost for each.

3.2 Full output disclosure

Unlike the planner, the principal cannot dictate the effort levels, which are unobservable. But suppose she commits to disclosing the entire output vector $\mathbf{y}$ to every market. That is, $\lambda_1 = \lambda_2 = 0$, and agent i's market observes $\mathbf{y}$. Note that unable to monitor each other's action, the agents play a simultaneous-move game at the effort stage. Let $\mathbf{x}^{FO} = (x_1^{FO}, x_2^{FO})$ be a (Nash) equilibrium of this game. Then, observing $\mathbf{y}$ and conjecturing $\mathbf{x}^{FO}$, the (Bayesian) market infers agent i's expected talent to be:

$$E[\eta_i|\mathbf{y}, \mathbf{x}^{FO}] = \mu_i + \alpha_i \left(y_i - E[y_i^{FO}]\right) + \beta_{-i} \left(y_{-i} - E[y_{-i}^{FO}]\right), \tag{5}$$

where $y_i^{FO} \equiv y_i | x_i^{FO}$ and

$$\alpha_i = \frac{\sigma_i^2(\sigma_{-i}^2 + \sigma_\varepsilon^2) - (\rho\sigma_i\sigma_{-i})^2}{(\sigma_i^2 + \sigma_\varepsilon^2)\,(\sigma_{-i}^2 + \sigma_\varepsilon^2) - (\rho\sigma_i\sigma_{-i})^2} \text{ and } \beta_{-i} = \frac{\rho\sigma_i\sigma_{-i}(\sigma_i^2 + \sigma_\varepsilon^2) - (\rho\sigma_i\sigma_{-i})^2}{(\sigma_i^2 + \sigma_\varepsilon^2)\,(\sigma_{-i}^2 + \sigma_\varepsilon^2) - (\rho\sigma_i\sigma_{-i})^2}. \tag{6}$$

Eq.(5) says that the market runs a linear regression to estimate agent i's mean talent.[10] Intuitively, the market revises its belief about agent i's talent to the extent that it is surprised by his output and, given the correlation, his peer's.

Anticipating (5) as his ex-post market offer (see Remark 1), agent i maximizes his ex-ante payoff that best responds to the peer's equilibrium effort:

$$\max_{x_i} E\left[E[\eta_i|\mathbf{y}, \mathbf{x}^{FO}]|x_i, x_{-i}^{FO}\right] - c(x_i), \tag{7}$$

[10]Given the normal priors, the random vector $(\widetilde{\eta}_i, \widetilde{y}_i, \widetilde{y}_{-i})$ is a trivariate normal. From here, (5) follows; see, for instance, Tong (2012, Th. 3.3.4).

where the expected market offer simplifies to: $E\left[E[.]|.\right] = \mu_i + \alpha_i\left(x_i - x_i^{FO}\right)$. Plugging this into (7), the equilibrium effort that solves (7) is immediate:

$$x_i^{FO} = \psi(\alpha_i). \tag{8}$$

Note that taking his action in expectation of the peer's, agent i's equilibrium effort does not depend on β_{-i}. As such, whether the market provides the agents with CPE or RPE has no direct incentive effect. But the market's use of y_{-i} has an indirect incentive effect via talent correlation in α_i. In particular, it is readily observed from (6) that α_i is strictly decreasing in ρ, implying these bounds (for $\rho = 1$ and $\to 0$):

$$\frac{\sigma_i^2}{\sigma_i^2 + \sigma_{-i}^2 + \sigma_\varepsilon^2} \leq \alpha_i < \frac{\sigma_i^2}{\sigma_i^2 + \sigma_\varepsilon^2}. \tag{9}$$

Therefore, we have

Lemma 1 *Under full output disclosure, agents with more correlated talents work less, i.e.,* $\frac{\partial x_i^{FO}}{\partial \rho} < 0$.

Proof. Immediate from (8) given that α_i is strictly decreasing in ρ, and $\psi' > 0$. ■

Under full output disclosure, having similar talents de-motivates agents by introducing a positive reputation externality: when an agent works harder to surprise the market with his output, it also helps his peer's reputation due to correlated talents. The principal can eliminate this externality by partial disclosure.

3.3 Partial output disclosure

Suppose the principal discloses only y_i to agent i's (separate) market. Then, the market cannot use y_{-i} to glean information about i's talent. Breaking the correlation between the agents' outputs, such partial output disclosure or IPE is incentive equivalent to full output disclosure with $\rho \to 0$. Hence, using (8) and (9), agent i's equilibrium effort reduces to that in a single-agent setting:

$$x_i^{PO} = \psi\left(\frac{\sigma_i^2}{\sigma_i^2 + \sigma_\varepsilon^2}\right). \tag{10}$$

This leads us to

Lemma 2 $x_i^{FO} < x_i^{PO} < x_i^{FB}$.

Proof. Immediate from (4), (8) and (9). ■

In words, a principal committed to disclosing the agents' outputs provides the strongest incentive when she tailors her disclosure to each market. The equilibrium effort remains below the first-best under partial disclosure due to the exogenous noise, which dampens the career motive. Lemma 2 generalizes Meyer and Vickers (1997) and Fleckinger et al. (Forthcoming, Section 7.1) to asymmetric agents. However, the principal can do even better by carefully rating performance, as I explore next.

4 Optimal performance rating

In Lemma 2, the principal prefers partial output disclosure to manage the market's data aggregation process in (5) and better motivate the agents. The principal can further manage that process by first aggregating the data into performance ratings.

Suppose the agents learn the principal's rating system in (2) and simultaneously choose their efforts.[11] Let $\mathbf{x}^* = (x_1^*, x_2^*)$ be their equilibrium efforts. Based on his disclosed rating, r_i, and the conjectured efforts, $\mathbf{x}^*$, agent i's market estimates his talent to be:[12]

$$E[\eta_i|r_i(\mathbf{y}), \mathbf{x}^*] = \mu_i + \frac{Cov(\eta_i, r_i(\mathbf{y}^*))}{Var(r_i(\mathbf{y}^*))} \left(r_i(\mathbf{y}) - E[r_i(\mathbf{y}^*)]\right). \tag{11}$$

As with (5), the market revises agent i's mean talent to the extent that it is surprised by his rating and can explain this variation by his talent. Given (11), agent i best responds to his peer's effort x_{-i}^*:

$$\max_{x_i} E\left[E[\eta_i|r_i(\mathbf{y}), \mathbf{x}^*]|x_i, x_{-i}^*\right] - c(x_i), \tag{12}$$

where, by canceling terms, $E\left[E[.]|.\right] = \mu_i + \frac{Cov(\eta_i, r_i(\mathbf{y}^*))}{Var(r_i(\mathbf{y}^*))}(x_i - x_i^*)$. Hence, in equilibrium, x_i^* must satisfy the first-order condition for (12):

$$\frac{Cov(\eta_i, r_i(\mathbf{y}^*))}{Var(r_i(\mathbf{y}^*))} = c'(x_i^*), \tag{13}$$

or more explicitly,

$$\frac{\sigma_i^2 + \lambda_i \rho \sigma_i \sigma_{-i}}{\left(\sigma_i^2 + \sigma_\varepsilon^2\right) + 2\lambda_i \rho \sigma_i \sigma_{-i} + \lambda_i^2 \left(\sigma_{-i}^2 + \sigma_\varepsilon^2\right)} = c'(x_i^*). \tag{14}$$

[11] As is evident from (14), it suffices for agent i to know only his rating, λ_i.

[12] Recall that agent i's market learns only his rating, r_i, in this part. Also, the random vector $(\widetilde{\eta}_i, r_i(\widetilde{\mathbf{y}}))$ is bivariate normal.

To maximize the expected total output, the principal maximizes the total equilibrium effort: $x_1^* + x_2^*$. Given $c'' > 0$, this amounts to maximizing the LHS of (14) for each agent by choosing λ_i.

Proposition 1 *The unique optimal rating for agent i has*

$$\lambda_i^* = \frac{\sigma_i}{\rho\sigma_{-i}} \left[\sqrt{1 + \rho^2 \frac{\sigma_{-i}^2}{\sigma_{-i}^2 + \sigma_\varepsilon^2} \left(\frac{\sigma_\varepsilon^2}{\sigma_i^2} - 1 \right)} - 1 \right]. \tag{15}$$

In particular,

- **(a)** $sgn\,[\lambda_i^*] = sgn \left[\frac{\partial \lambda_i^*}{\partial \rho} \right] = sgn\,[\sigma_\varepsilon - \sigma_i]$,
- **(b)** $\frac{\partial x_i^*}{\partial \rho} \geq 0$, *with strict inequality for* $\sigma_i \neq \sigma_\varepsilon$,
- **(c)** $\sigma_i > \sigma_{-i}$ *implies* $\lambda_i^* < \lambda_{-i}^* < 1$ *and* $x_i^* > x_{-i}^*$,
- **(d)** $x_i^{PO} \leq x_i^* < x_i^{FB}$, *with* $x_i^{PO} = x_i^*$ *whenever* $\sigma_i = \sigma_\varepsilon$.

Proof. Note from (14) that the principal can always elicit positive effort by setting $\lambda_i = 0$. Therefore, its numerator, the covariance term, must be positive for the effort-maximizing λ_i; namely,

$$\sigma_i^2 + \lambda_i \rho \sigma_i \sigma_{-i} > 0. \tag{16}$$

Next, denoting its LHS by $\Omega(\lambda_i; .)$, (14) reads,

$$\Omega(\lambda_i; .) = c'(x_i^*). \tag{17}$$

Since $c'' > 0$, maximizing x_i^* with respect to λ_i requires that $\Omega'(\lambda_i; .) = 0$, which yields (15). Evidently, λ_i^* is always real. To show its uniqueness, note that

$$\Omega''(\lambda_i; .) \text{ at } \Omega'(\lambda_i; .) = 0 \propto -2(\sigma_{-i}^2 + \sigma_\varepsilon^2) \left(\sigma_i^2 + \lambda_i \rho \sigma_i \sigma_{-i} \right) < 0 \text{ by (16).}$$

Hence, $\Omega(\lambda_i; .)$ is strictly quasi-concave in λ_i satisfying (16), which establishes the uniqueness of λ_i^*.

For part (a), eq.(15) readily implies $sgn\,[\lambda_i^*] = sgn\,[\sigma_\varepsilon - \sigma_i]$. Furthermore,

$$\frac{\partial \lambda_i^*}{\partial \rho} = \left(\rho \sqrt{1 + \rho^2 \frac{\sigma_{-i}^2}{\sigma_{-i}^2 + \sigma_\varepsilon^2} \left(\frac{\sigma_\varepsilon^2}{\sigma_i^2} - 1 \right)} \right)^{-1} \lambda_i^*,$$

implying $sgn\left[\frac{\partial\lambda_i^*}{\partial\rho}\right] = sgn\left[\lambda_i^*\right]$.

For part (b), observe from (14) that $c'(x_i^*) = \frac{\sigma_i^2}{\sigma_i^2+\sigma_\varepsilon^2}$, independent of ρ, for $\sigma_i = \sigma_\varepsilon$ since $\lambda_i^* = 0$ by part (a). Now consider $\sigma_i \neq \sigma_\varepsilon$. Plugging (15) into the LHS of (14) reveals

$$\Omega(\lambda_i^*;.) = \frac{1}{2}\frac{rab}{\sqrt{a\,(b+1)\,[a\,(b+1) - r\,(a-1)\,b]} - a\,[(1-r)\,b+1]}, \tag{18}$$

where $a \equiv \sigma_i^2/\sigma_\varepsilon^2$, $b \equiv \sigma_{-i}^2/\sigma_\varepsilon^2$, and $r \equiv \rho^2$. Straightforward algebra shows

$$\begin{aligned}\frac{\partial\Omega(\lambda_i^*;.)}{\partial r} &\propto 2a(b+1) - r\,(a-1)\,b - 2\sqrt{a\,(b+1)\,[a\,(b+1) - r\,(a-1)\,b]} \\ &\equiv J(r).\end{aligned}$$

Moreover,

$$J'(r) \propto (a-1)\left[\sqrt{a\,(b+1)} - \sqrt{a\,(b+1) - rb\,(a-1)}\right] > 0,$$

and $J(0) = 0$. Hence, $J(r) > 0$ for $r > 0$. Given $c'' > 0$, this means $\frac{\partial x_i^*}{\partial\rho} > 0$ by (17).

To prove part (c), suppose $\sigma_i > \sigma_{-i}$. For brevity, let $D(a,b,r) \equiv 2\left(\sqrt{d(a,b,r)} - a\,[(1-r)\,b+1]\right)$ represent the denominator of (18). To show $x_i^* > x_{-i}^*$, it suffices to show $D(a,b,r) < D(b,a,r)$ for $a > b$. Note that

$$\begin{aligned}\frac{D(a,b,r) - D(b,a,r)}{2} &= \sqrt{d(a,b,r)} - \sqrt{d(b,a,r)} - (a-b) \\ &= \frac{d(a,b,r) - d(b,a,r)}{\sqrt{d(a,b,r)} + \sqrt{d(b,a,r)}} - (a-b) \\ &= (a-b)\left(\frac{[a(b+1) - rab] + [b(a+1) - rab]}{\sqrt{d(a,b,r)} + \sqrt{d(b,a,r)}} - 1\right).\end{aligned}$$

It is verified that

$$a(b+1) - rab < \sqrt{d(a,b,r)} \quad \text{and} \quad b(a+1) - rab < \sqrt{d(b,a,r)}.$$

Hence, $D(a,b,r) - D(b,a,r) < 0$ for $a > b$.

To complete part (c), consider first $\sigma_i > \sigma_\varepsilon \geq \sigma_{-i}$. Then, $\lambda_i^* < 0 \leq \lambda_{-i}^*$ by part (a). Next consider $\sigma_\varepsilon \geq \sigma_i > \sigma_{-i}$. Re-arranging (15),

$$\rho\lambda_i^* = \sqrt{\left(\frac{\sigma_i}{\sigma_{-i}}\right)^2 + \rho^2\left(1 - \frac{\sigma_i^2 + \sigma_{-i}^2}{\sigma_{-i}^2 + \sigma_\varepsilon^2}\right)} - \frac{\sigma_i}{\sigma_{-i}}. \tag{19}$$

Since $\sigma_i > \sigma_{-i}$, $\frac{\sigma_i}{\sigma_{-i}} > \frac{\sigma_{-i}}{\sigma_i}$, and $\sqrt{z^2+k} - z$ is strictly decreasing in z for $k > 0$,

$$\rho\lambda_i^* < \sqrt{\left(\frac{\sigma_{-i}}{\sigma_i}\right)^2 + \rho^2\left(1 - \frac{\sigma_i^2 + \sigma_{-i}^2}{\sigma_i^2 + \sigma_\varepsilon^2}\right)} - \frac{\sigma_{-i}}{\sigma_i} = \rho\lambda_{-i}^* \Longrightarrow \lambda_i^* < \lambda_{-i}^*.$$

The case $\sigma_i > \sigma_{-i} \geq \sigma_\varepsilon$ similarly follows. Also, since λ_i^* is strictly increasing in σ_ε by (19),

$$\rho\lambda_i^* < \sqrt{\left(\frac{\sigma_i}{\sigma_{-i}}\right)^2 + \rho^2} - \frac{\sigma_i}{\sigma_{-i}} < \rho \Longrightarrow \lambda_i^* < 1.$$

For part (d), notice from (14) that $x_i^{PO} = x_i^*$ for $\lambda_i^* = 0$. Hence, $x_i^{PO} < x_i^*$ whenever $\lambda_i^* \neq 0$. Finally, it is checked that $\Omega(\lambda_i;.) < 1 \Longleftrightarrow 0 < \left(\sigma_i^2 + \sigma_\varepsilon^2\right) + \lambda_i\rho\sigma_i\sigma_{-i} + \lambda_i^2\left(\sigma_{-i}^2 + \sigma_\varepsilon^2\right)$, which holds for all λ_i because the minimum value of its RHS equals $\sigma_\varepsilon^2 + \sigma_i^2\left(1 - \rho^2\frac{\sigma_{-i}^2}{4(\sigma_{-i}^2 + \sigma_\varepsilon^2)}\right) > 0$. Thus, $c'(x_i^*) < c'(x_i^{FB})$, implying $x_i^* < x_i^{FB}$. ■

Part (a) of Proposition 1 indicates that an agent is given CPE if he is sufficiently known by the market, i.e., $\lambda_i^* > 0$ if $\sigma_i < \sigma_\varepsilon$. Consider, for instance, an agent with a virtually known talent ($\sigma_i \approx 0$). Evidently, he would have little career incentive of his own ($x_i^{PO} \approx 0$). To motivate, the principal positively correlates his rating with the peer's output, tying their reputations. In general, the LHS of (14) suggests two opposite incentive effects of CPE. A higher positive λ_i discourages agent i by increasing the total variation in his rating (the denominator term). However, it also encourages him by increasing how much of this variation can be attributed to their similar talents (the numerator term). Part (a) reveals that the latter effect dominates for a low-variance agent. By the same logic, a sufficiently high-variance agent is given RPE, i.e., $\lambda_i^* < 0$ if $\sigma_i > \sigma_\varepsilon$. In this case, the principal worries that team-based incentives would introduce too much uncertainty in the agent's rating and de-motivate him. Although RPE would reduce the covariance between the agent's talent and rating, the principal errs on the side of reducing the aggregate variance of his rating for incentives if the agent is sufficiently unknown to the market.

One implication of part (a) is that agents may receive sharply different performance evaluations: if $\sigma_1 < \sigma_\varepsilon < \sigma_2$, agent 1 faces CPE, whereas agent 2 faces RPE. For example, using (15) and rounding λ_i^*, the optimal ratings are: $r_1^* \approx y_1 + .16y_2$ and $r_2^* \approx y_2 - .08y_1$ for $\sigma_1 = .5$, $\sigma_\varepsilon = 1$, $\sigma_2 = 1.5$, and $\rho = .5$.[13] Another implication is that the principal would never induce a head-to-head competition between (ex-post) identical agents, i.e., $r_i^* \neq y_i - y_{-i}$ for $\mu_i = \mu$, $\sigma_i = \sigma$ and $\rho = 1$. Otherwise, by perfectly filtering out their talents, such ratings would become pure noises for the agents, inducing no costly effort.

[13] *Wolfram Mathematica* was used for all numerical examples.

Part (a) further indicates that the agents' ratings would be more dependent on each other's performance the more correlated their talents are. This is because correlation increases the incentive role of the rating choice, as captured by the term $\lambda_i \rho \sigma_i \sigma_{-i}$ in (14). Therefore, an agent facing CPE or RPE will face a steeper one with a higher correlation. Such fine-tuning of the rating system is consistent with part (b): the equilibrium effort strictly increases with talent correlation. This means that if she could, the principal would hire equally-talented agents, perhaps those with identical education and training. Notice that this prediction starkly contrasts with the full output disclosure in Lemma 1. By tying their reputations ex-ante in her ratings, the principal induces the agents to (partially) internalize the reputational spillover between them. It also contrasts with the traditional view of contract theory. As reviewed in the Introduction, this theory prescribes RPE when agents' performances are subject to common shocks, here their unknown talents, which RPE helps filter out. However, with career concerns, such filtering also weakens the relationship between an agent's talent and performance rating, dampening his incentive to improve the latter.

Part (c) reinforces part (a) in that the better known agent is more likely to be given team incentives. Interestingly, $\lambda_i^* < 1$, so that simply disclosing the team output, $y_i + y_{-i}$, is never optimal, even for ex-post identical agents. Also, part (c) adds to part (b) in that the less known agent works harder in equilibrium, as he has more to gain in reputation. This finding aligns with the benchmarks. Finally, part (d) confirms that the principal indeed elicits more effort with an optimal rating system than individual output disclosure or IPE (with indifference at $\sigma_i = \sigma_\varepsilon$).

5 Public performance ratings and CPE

Unlike in the previous section, suppose agent i's market also learns his peer's rating, r_{-i}. The principal cannot benefit from such full rating disclosure since she would then have to worry about the market pinpointing individual outputs and estimating talents as in the full output disclosure, which would significantly discourage effort. Nevertheless, rating disclosure may be beyond the principal's control. For instance, the manager may not prevent the employees from considering the same jobs and having their recommendation letters compared.[14] Therefore, if the principal expects all her ratings to become public, she will likely choose them differently from their "confidential" counterparts in Section 4. In fact, as shown below, public

[14]Likewise, professors may not compel students to apply to different PhD programs.

ratings that improve upon full output disclosure can only feature team incentives or CPE.

Note that when the market learns both r_1 and r_2, it can uniquely infer each output y_i, provided $\lambda_1 \lambda_2 \neq 1$. Since this is strategically equivalent to full output disclosure where $\lambda_1 = \lambda_2 = 0$, eq.(8) implies the total equilibrium effort:

$$X^{FO} \equiv \psi(\alpha_1) + \psi(\alpha_2) \text{ if } \lambda_1 \lambda_2 \neq 1. \tag{20}$$

Alternatively, the principal can set $\lambda_1 \lambda_2 = 1$ to obscure individual outputs. This no-output-disclosure constraint already rules out heterogeneous evaluations where one agent receives CPE and the other RPE. In fact, by (2), it means perfectly correlated public ratings:

$$r_1(\mathbf{y}) = \lambda_1 r_2(\mathbf{y}). \tag{21}$$

Since scaling of a rating has no incentive consequence, eq.(21) implies that the principal effectively promises to report the same rating, say (r_1, r_1), or ones with opposite signs $(r_1, -r_1)$. To understand when the principal prefers public ratings with no output disclosure, I first write from (14) the total equilibrium effort without this constraint:

$$X^{NO}(\lambda_1, \lambda_2) \equiv \psi\left(\Omega(\lambda_1; \sigma_1, \sigma_2, \rho)\right) + \psi\left(\Omega(\lambda_2; \sigma_2, \sigma_1, \rho)\right), \tag{22}$$

where $\Omega(.)$ represents the LHS of (14).

Let $\Delta(\lambda_1, \lambda_2) \equiv X^{NO}(\lambda_1, \lambda_2) - X^{FO}$. Clearly, if $\Delta(\lambda_1, \lambda_2) > 0$ for some $\lambda_1 \lambda_2 = 1$, the principal strictly prefers public ratings with no output disclosure to full output disclosure, and vice versa. To determine, she checks the sign of the following value:

$$\max_{(\lambda_1, \lambda_2) \in \mathbb{R}^2} \Delta(\lambda_1, \lambda_2) \text{ s.t. } \lambda_1 \lambda_2 = 1. \tag{PR}$$

Without its constraint, the optimal confidential ratings in Proposition 1 would uniquely solve (PR). In particular, $\Delta(\lambda_1^*, \lambda_2^*) > 0$ since $x_i^* > x_i^{FO}$ by Lemma 2 and Proposition 1(d). However, it is verified that $\lambda_1^* \lambda_2^* \neq 1$.[15] That is, if they were made public, the optimal confidential ratings would reveal individual outputs, reducing incentives to those under full output disclosure. The following result shows the principal can sometimes overcome this issue.

Proposition 2 *Besides (C1), suppose $c'''(x) > 0$ for $x > 0$. Then, a solution to (PR), denoted by $(\lambda_1^{**}, \lambda_2^{**})$, exists. In particular, $\Delta(\lambda_1^{**}, \lambda_2^{**}) > 0$ if and only if $\sigma_1 \neq \sigma_2$ and ρ is sufficiently close to* 1*. Furthermore, if $\Delta(\lambda_1^{**}, \lambda_2^{**}) > 0$, then*

[15]Since λ_i^* in (15) is strictly increasing in σ_ε^2, it readily follows that $\lambda_1^* \lambda_2^* < \rho^2$.

(a) $\lambda_1^{**} = \frac{1}{\lambda_2^{**}} > 0$,

(b) $\lambda_1^{**} > \lambda_1^*$ *and* $\lambda_2^{**} > \lambda_2^*$, *and*

(c) $X^{NO}(\lambda_1^{**}, \lambda_2^{**})$ *is strictly increasing in* ρ.

Proof. Besides (C1), suppose $c'''(x) > 0$ for $x > 0$. Then, we also have $\psi'' < 0$ in (3). Notice that $\lambda_i = \pm\infty$ for some i cannot solve (PR); otherwise, agent i would exert no effort, which would be strictly suboptimal given that $\psi(0) = 0$ and $\psi'' < 0$. Hence, the feasible set $\{(\lambda_1, \lambda_2) \in \mathbb{R}^2 | \lambda_1\lambda_2 = 1\}$ is compact. Moreover, since $\Delta(\lambda_1, \lambda_2)$ is continuous, there exists a solution $(\lambda_1^{**}, \lambda_2^{**})$ to (PR).

Suppose $\Delta\left(\lambda_1^{**}, \lambda_2^{**}\right) > 0$, but, to the contrary, $\sigma_1 = \sigma_2 = \sigma$. I first argue $\lambda_1^{**} = \lambda_2^{**} = 1$ for any $\rho > 0$.

Using (22) and $\lambda_1\lambda_2 = 1$, define

$$\widehat{\Delta}(\lambda_1) \equiv \Delta(\lambda_1, 1/\lambda_1).$$

Then, for $\sigma_1 = \sigma_2 = \sigma$,

$$\widehat{\Delta}(\lambda_1) = \psi\left(\tfrac{\sigma^2 + \lambda_1\rho\sigma^2}{(1+\lambda_1^2)(\sigma^2+\sigma_\varepsilon^2)+2\lambda_1\rho\sigma^2}\right) + \psi\left(\tfrac{\lambda_1^2\sigma^2 + \lambda_1\rho\sigma^2}{(1+\lambda_1^2)(\sigma^2+\sigma_\varepsilon^2)+2\lambda_1\rho\sigma^2}\right) - X^{FO}.$$

Straightforward algebra reveals

$$\widehat{\Delta}'(\lambda_1) = -\frac{A}{D}\psi'\left(\tfrac{\sigma^2 + \lambda_1\rho\sigma^2}{(1+\lambda_1^2)(\sigma^2+\sigma_\varepsilon^2)+2\lambda_1\rho\sigma^2}\right) + \frac{B}{D}\psi'\left(\tfrac{\lambda_1^2\sigma^2 + \lambda_1\rho\sigma^2}{(1+\lambda_1^2)(\sigma^2+\sigma_\varepsilon^2)+2\lambda_1\rho\sigma^2}\right),$$

where $A = \sigma_\varepsilon^2(\rho\lambda_1^2 + 2\lambda_1 - \rho) + \sigma^2(\rho\lambda_1^2 + 2\lambda_1 + \rho)$, $B = \sigma_\varepsilon^2\left(-\rho\lambda_1^2 + 2\lambda_1 + \rho\right) + \sigma^2(\rho\lambda_1^2 + 2\lambda_1 + \rho)$ and $D = \left((1+\lambda_1^2)\left(\sigma^2+\sigma_\varepsilon^2\right) + 2\lambda_1\rho\sigma^2\right)^2/\sigma^2$. Clearly, $A - B = 2\sigma_\varepsilon^2\rho(\lambda_1^2 - 1)$ and $D > 0$.

If $\lambda_1^2 > 1$, then $A > B$, and, because $\psi'' < 0$, $\widehat{\Delta}'(\lambda_1) < 0$. Symmetrically, $\lambda_1^2 < 1$ implies $\widehat{\Delta}'(\lambda_1) > 0$. Hence, $\widehat{\Delta}'(\lambda_1) = 0$ requires $\lambda_1^2 = 1$. Moreover, $\widehat{\Delta}(1) = 2\psi\left(\frac{\sigma^2(1+\rho)}{2(\sigma^2+\sigma_\varepsilon^2)+2\rho\sigma^2}\right) - X^{FO}$ and $\widehat{\Delta}(-1) = 2\psi\left(\frac{\sigma^2(1-\rho)}{2(\sigma^2+\sigma_\varepsilon^2)+2\rho\sigma^2}\right) - X^{FO}$. Clearly, $\widehat{\Delta}(1) > \widehat{\Delta}(-1)$ for $\rho > 0$, given $\psi' > 0$. Thus, $\lambda_1^{**} = 1 = \lambda_2^{**}$ for $\sigma_1 = \sigma_2$. Furthermore, since X^{FO} is strictly decreasing in ρ by Lemma 1, $\widehat{\Delta}(1)$ is strictly increasing in ρ, with $\widehat{\Delta}(1) = 0$ for $\rho = 1$. This implies $\Delta\left(\lambda_1^{**}, \lambda_2^{**}\right) \leq 0$, a contradiction. Hence, $\sigma_1 \neq \sigma_2$.

Conversely, suppose $\sigma_1 \neq \sigma_2$ and $\rho = 1$. I first show $\widehat{\Delta}(\lambda_1) > 0$ for some $\lambda_1 \in \mathbb{R}$. It is verified that $\widehat{\Delta}(\frac{\sigma_2}{\sigma_1}) = 0$ and

$$sgn\left[\widehat{\Delta}'(\frac{\sigma_2}{\sigma_1})\right] = sgn\left[\psi'\left(\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2+\sigma_\varepsilon^2}\right) - \psi'\left(\frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2+\sigma_\varepsilon^2}\right)\right].$$

If $\sigma_1 > \sigma_2$, then $\widehat{\Delta}'(\frac{\sigma_2}{\sigma_1}) > 0$ since $\psi'' < 0$. Thus, $\widehat{\Delta}(\frac{\sigma_2}{\sigma_1} + \epsilon) > 0$ for some $\epsilon > 0$. Similarly, if $\sigma_1 < \sigma_2$, then $\widehat{\Delta}(\frac{\sigma_2}{\sigma_1} - \epsilon) > 0$ for some $\epsilon > 0$.

To prove the converse for ρ close to 1, I next prove part (a): $\widehat{\Delta}(\lambda_1) > 0$ implies $\lambda_1 > 0$. Since, as argued above, the principal will optimally induce both agents to exert effort, she will pick (λ_1, λ_2) such that

$$\sigma_1^2 + \lambda_1\rho\sigma_1\sigma_2 > 0 \quad \text{and} \quad \sigma_2^2 + \lambda_2\rho\sigma_1\sigma_2 > 0. \tag{23}$$

Furthermore, $\widehat{\Delta}(\lambda_1) > 0$ implies some agent exerts more effort than full output disclosure. If this is agent 1, (20) and (22) require

$$\alpha_1 < \frac{\sigma_1^2 + \lambda_1\rho\sigma_1\sigma_2}{\left(\sigma_1^2 + \sigma_\varepsilon^2\right) + 2\lambda_1\rho\sigma_1\sigma_2 + \lambda_1^2\left(\sigma_2^2 + \sigma_\varepsilon^2\right)}$$

$\Longleftrightarrow$

$$-\frac{\rho\sigma_1\sigma_2}{\sigma_2^2 + \sigma_\varepsilon^2} < \lambda_1 < \rho\frac{\sigma_2}{\sigma_1}\frac{\sigma_\varepsilon^2}{\left(\sigma_\varepsilon^2 + (1-\rho^2)\sigma_2^2\right)}. \tag{24}$$

If agent 2, then symmetrically,

$$-\frac{\rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_\varepsilon^2} < \lambda_2 < \rho\frac{\sigma_1}{\sigma_2}\frac{\sigma_\varepsilon^2}{\left(\sigma_\varepsilon^2 + (1-\rho^2)\sigma_1^2\right)}. \tag{25}$$

Now suppose, to the contrary, $\lambda_1 < 0$. Then, given $\lambda_2 = 1/\lambda_1$, (23) and (24) imply

$$-\frac{\rho\sigma_1\sigma_2}{\sigma_2^2 + \sigma_\varepsilon^2} < \lambda_1 < -\frac{\rho\sigma_1}{\sigma_2} \Longrightarrow \sigma_\varepsilon^2 < 0, \text{ a contradiction.}$$

From (23) and (25),

$$-\frac{\sigma_1}{\rho\sigma_2} < \lambda_1 < -\frac{\sigma_1^2 + \sigma_\varepsilon^2}{\rho\sigma_1\sigma_2} \Longrightarrow \sigma_\varepsilon^2 < 0, \text{ a contradiction.}$$

Hence, $\widehat{\Delta}(\lambda_1) > 0$ implies $\lambda_1 > 0$, as claimed.

Now, given $\lambda_1^{**} > 0$, we have $\partial\Omega(\lambda_1^{**}; ., \rho)/\partial\rho > 0$ and $\partial\Omega(1/\lambda_1^{**}; ., \rho)/\partial\rho > 0$, where $\Omega(.)$ represents the LHS of (14). Since X^{FO} is decreasing in ρ, $\widehat{\Delta}(\lambda_1^{**})$ is strictly increasing in ρ by the envelope argument. Hence, if $\widehat{\Delta}(\lambda_1^{**}) > 0$ for $\rho = 1$, it also holds for ρ sufficiently close to 1, as desired.

For part (b), let us write the Lagrangian for (PR):

$$\mathcal{L} = \Delta(\lambda_1, \lambda_2) + \nu(1 - \lambda_1\lambda_2).$$

Then, using (22), the first-order conditions for (PR) can be re-stated:

$$\psi'\left(\Omega(\lambda_1^{**};.)\right)\Omega'(\lambda_1^{**};.) = v^{**}\lambda_2^{**} \text{ and } \psi'\left(\Omega(\lambda_2^{**};.)\right)\Omega'(\lambda_2^{**};.) = v^{**}\lambda_1^{**}. \tag{26}$$

Clearly, $v^{**} \neq 0$; otherwise, $(\lambda_1^*, \lambda_2^*)$ would be the solution to (PR). But $\lambda_1^*\lambda_2^* \neq 1$, as noted above. Next, since $\lambda_1^{**}\lambda_2^{**} = 1$ and $\psi' > 0$, multiplying (26) side by side reveals

$$\Omega'(\lambda_1^{**};.)\Omega'(\lambda_2^{**};.) > 0. \tag{27}$$

Recall from the proof of Proposition 1 that $\Omega(\lambda_i;.)$ is single-peaked in λ_i, with a unique maximizer λ_i^*. Therefore, $\lambda_i^{**} \neq \lambda_i^*$. Suppose, to the contrary, that $\lambda_1^{**} < \lambda_1^*$. Then, $\Omega'(\lambda_1^{**};.) > 0$ and thus, $\Omega'(\lambda_2^{**};.) > 0$ by (27). This implies $\lambda_2^{**} < \lambda_2^*$, and because $\lambda_i^{**} > 0$, we have $\lambda_1^{**}\lambda_2^{**} = 1 < \lambda_1^*\lambda_2^*$. But this means $\lambda_i^* \geq 1$ for some i, contradicting Proposition 1(c). Hence, $\lambda_1^{**} > \lambda_1^*$ and, similarly, $\lambda_2^{**} > \lambda_2^*$.

Finally, since X^{FO} is independent of (λ_1, λ_2), any pair $(\lambda_1^{**}, \lambda_2^{**}) \in \mathbb{R}_{++}$ that solves (PR) would also solve: $\max_{(\lambda_1,\lambda_2)\in\mathbb{R}^2} X^{NO}(\lambda_1, \lambda_2)$ s.t. $\lambda_1\lambda_2 = 1$. Hence, by the envelope argument, part (c) follows. ■

Proposition 2(a) shows that optimal public ratings that elicit more effort than full output disclosure must provide team incentives or CPEs. To understand, note that the market would easily infer individual outputs if one agent faced CPE and the other RPE, i.e., $\lambda_i > 0$ and $\lambda_{-i} < 0$. The principal cannot provide both agents with RPEs, either, because one would need to be much steeper than the other to obscure their outputs. However, the agent facing a steep RPE would exert little or no effort, given the substantially low, or even negative, covariance between his talent and rating. Such unequal effort allocation cannot be optimal for the principal because of the convex marginal cost of effort, i.e., $c''' > 0$, leaving her with CPE for each agent.

The principal can design CPEs that induce more effort than full output disclosure for agents of ex-ante heterogeneous but highly correlated talents, which is when they care most about their collective reputation. To illustrate, let $\sigma_1 = 3$, $\sigma_2 = 2$, $\sigma_\varepsilon = 1$, and $\rho = 1$. Also let $c(x) = x^3/3$ so that $\psi(x) = \sqrt{x}$. Then, $\lambda_1^{**} \approx 1.24$ and $\Delta\left(\lambda_1^{**}, 1/\lambda_1^{**}\right) \approx .02$, implying the public rating: $r_1^{**} \approx y_1 + 1.24y_2$ for both agents. Hence, agents know they will receive the same rating but have different marginal impacts on it. Notice that r_1^{**} starkly contrasts with the confidential ones from Proposition 1, which would have $\lambda_1^* = -.69$ and $\lambda_2^* = -.29$, referring to RPEs. In fact, Proposition 2(b) reveals that even when optimal confidential ratings are all CPEs, i.e., $\sigma_1, \sigma_2 < \sigma_\varepsilon$, public ratings call for steeper ones, making agents more concerned about collective reputation.[16]

Proposition 2 further says that the principal cannot strictly improve upon full output disclosure when agents are equally known ($\sigma_1 = \sigma_2$). The reason is that with a strictly convex marginal cost, the optimal public rating that obscures individual performances is the total output: $r_1^{**} = y_1 + y_2$. However, given the normal priors, such a rating would carry the same information about talents as full output disclosure if agents' talents were perfectly correlated, and strictly less information otherwise. Thus, the principal would be strictly worse off if she insisted on concealing the individual performances of agents with imperfectly correlated talents. Interestingly, a similar result holds even for ex-ante heterogenous agents when the cost of effort is quadratic. For completeness, I formalize these observations in:

Proposition 3 *If (**1**) $c''' > 0$ and $\sigma_1 = \sigma_2$, or (**2**) $c(x) = kx^2$, then $\Delta\left(\lambda_1^{**}, \lambda_2^{**}\right) \leq 0$, with equality whenever $\rho = 1$.*

Proof. Suppose $c''' > 0$ and $\sigma_1 = \sigma_2$. Then, as established in the previous proof, $\lambda_1^{**} = \lambda_2^{**} = 1$, and $\Delta\left(1,1\right) = \widehat{\Delta}(1)$ is strictly increasing in ρ, and $\widehat{\Delta}(1) = 0$ for $\rho = 1$.

Next, suppose $c(x) = kx^2$ and let $k = 1/2$ without loss. Then, $\psi(z) = z$, and $\widehat{\Delta}'(\lambda_1) = 0$ has these roots:

$$\lambda_1 = \frac{\sigma_2^2 - \sigma_1^2 \pm \sqrt{(\sigma_2^2 - \sigma_1^2)^2 + 4\left(\rho\sigma_1\sigma_2\right)^2}}{2\rho\sigma_1\sigma_2}.$$

It can be verified that $\widehat{\Delta}(\lambda_1^+) > \widehat{\Delta}(\lambda_1^-)$. Hence, $\lambda_1^{**} = \lambda_1^+$, and $x_i^{**} > 0$. Moreover, $\widehat{\Delta}(\lambda_1^{**}) \leq 0$, with equality for $\rho = 1$, as claimed. ■

Unlike confidential ratings, public ratings are affected by the shape of *marginal* cost. But this is expected: public ratings cannot be treated independently to hide the individual performances, forcing the principal to also consider marginal cost of effort across the agents.[17]

Remark 3 *As a robustness check, I have considered an extension in the appendix where the principal can also add random errors to public ratings to conceal individual outputs. Formally, I let $R_i(\mathbf{y}) = r_i(\mathbf{y}) + e_i$, where $e_i \overset{iid}{\sim} N(0, \sigma_e^2)$. The principal chooses $(\lambda_1, \lambda_2, \sigma_e)$. I show that she would*

[16] For example, $\lambda_1^{**} \approx .43$ and $\lambda_2^{**} \approx 2.34$ whereas $\lambda_1^* \approx .02$ and $\lambda_2^* \approx .65$ for $\sigma_1 = .8$, $\sigma_2 = .1$, $\sigma_\varepsilon = 1$, $\rho = .9$, and $c(x) = x^3/3$.

[17] If the marginal cost of effort were strictly concave, i.e., $c'''(x) < 0$ for $x > 0$ so that $\psi'' > 0$, (PR) would produce a corner solution where only one agent exerts effort under public ratings, making the multi-agent problem uninteresting.

not benefit from adding small errors ($\sigma_e \approx 0$) under any cost function and would be strictly worse off by any $\sigma_e > 0$ under the quadratic cost.

6 An extension: correlated noise and RPE

I have assumed thus far that agents face independent production noises, which is reasonable if the noises are specific to projects or individuals. However, they may also be positively correlated due to agents' similar working conditions. In this context, I demonstrate such a correlation unequivocally favors competitive incentives. This finding is consistent with contract theory: RPE helps isolate career concerns by filtering out exogenous noise. Therefore, when agents primarily work for their careers, determining which common shock (talent or noise) RPE mostly filters becomes important for incentive design.

Formally, consider confidential ratings in Section 4 and let $Corr(\varepsilon_1, \varepsilon_2) = \rho_\varepsilon \in (0, 1]$. Then, eq.(14) is modified as follows:

$$\frac{\sigma_i^2 + \lambda_i \rho \sigma_i \sigma_{-i}}{(\sigma_i^2 + \sigma_\varepsilon^2) + 2\lambda_i(\rho\sigma_i\sigma_{-i} + \rho_\varepsilon \sigma_\varepsilon^2) + \lambda_i^2\left(\sigma_{-i}^2 + \sigma_\varepsilon^2\right)} = c'(x_i^*), \tag{28}$$

where, as expected, noise correlation only affects the rating's total variance. Maximizing its LHS yields

$$\lambda_{i,\rho_\varepsilon}^* = \frac{\sigma_i}{\rho\sigma_{-i}}\left[\sqrt{1 + \rho^2 \frac{\sigma_{-i}^2}{\sigma_{-i}^2 + \sigma_\varepsilon^2}\left(\frac{\sigma_\varepsilon^2}{\sigma_i^2} - \left(1 + 2\frac{\rho_\varepsilon \sigma_\varepsilon^2}{\rho\sigma_i\sigma_{-i}}\right)\right)} - 1\right]. \tag{29}$$

It is verified that $\lambda_{i,\rho_\varepsilon}^* \in \mathbb{R}$. Moreover, compared with (15), $\lambda_{i,\rho_\varepsilon}^* < \lambda_i^*$. Hence, RPE is more likely with correlated noise.

7 Discussion and conclusion

This paper makes two related points. First, organizations with career-concerned agents should be cautious about RPE as an incentive mechanism. Although RPE can help isolate common exogenous shocks agents may face, it can also lower the correlation between one's performance measure and ability, dampening career incentives. Second, such organizations will likely adopt CPE and thus prefer agents with similar abilities, not diverse.

As mentioned in the Introduction, government agencies are prominent examples of these organizations. Interestingly, the U.S. Office of Personnel Management (OPM) has warned federal agencies against RPE: “By law, forced distribution of employees among levels of performance, or grading on the curve, is prohibited, because employees are required to be assessed against documented standards of performance versus an individual’s performance relative to others.”[18] Although the OPM has listed several reasons outside my model, including concerns for employee morale and biased ratings, its warning against RPE aligns with this paper’s message. However, my analysis further suggests CPE over IPE because (1) federal agencies often hire employees with specific and, thus, highly correlated talents for their missions, and (2) employee ratings will likely be compared for internal promotions. Indeed, Van der Hoek, Groeneveld, and Kuipers (2018) and McNabb and Swenson (2021) noted the increased emphasis on collaboration within and across public organizations.

While forced distribution ratings, or stack ranking, are used in the private sector, they appear less common in industries that value creativity, innovation, and collaboration; see Wijayanti et al. (2024) for a literature survey. In an influential news article, Eichenwald (2012) argued that Microsoft lost top talent and market value due to its stack ranking in the previous decade. The company abandoned the practice in 2013. Cohan (2013) documented that Adobe Systems, too, removed its stack ranking around the same time and gained 68% in stock value. This company stated, “We now use Check In – a system that fits with Adobe’s corporate culture of collaboration and creativity.” By mentioning similar reasons, other major companies that switched away from stack ranking include Accenture, Deloitte, and General Electric.

I close the paper with another prediction of my analysis: career-concerned agents would always prefer full output disclosure and let the market make its own inference about their talents. In equilibrium, they cannot “fool” the market and receive a higher offer, on average, than their prior mean talents, which formally refers to the law of iterated expectations. Thus, agent i’s equilibrium expected payoff is: $\mu_i - c(x_i^d)$, where d denotes the principal’s disclosure policy. And, x_i^{FO} is the lowest effort. Moreover, conditional on full output disclosure, agents would want the principal to hire peers with the most similar skills, i.e., $\rho = 1$. While this aligns with the principal’s preference, the reason is very different: the principal wants $\rho = 1$ only if she optimally obscures individual outputs.

[18] https://www.fedweek.com/issue-briefs/opm-sets-new-standards-on-performance-ratings/

Appendix: Public Ratings with Strategic Noise

Here, I prove the robustness claims made in Remark 3. Suppose the principal can also introduce random errors to agents' ratings:

$$R_i(\mathbf{y}) = r_i(\mathbf{y}) + e_i,$$

where $e_i \overset{iid}{\sim} N(0, \sigma_e^2)$, and e_i is independent of all the other random variables.

The principal ex-ante chooses $(\lambda_1, \lambda_2, \sigma_e)$, where $\sigma_e = 0$ refers to the analysis in Section 5. Conditional on observing $(R_i(\mathbf{y}), R_{-i}(\mathbf{y}))$ and conjecturing $\mathbf{x}^{**}$, the (Bayesian) market estimates agent i's talent to be:[19]

$$E[\eta_i | R_i(\mathbf{y}), R_{-i}(\mathbf{y}), \mathbf{x}^{**}] = \mu_i + \widehat{\alpha}_i \left(R_i(\mathbf{y}) - E[R_i(\mathbf{y}^{**})]\right) + \widehat{\beta}_{-i}(R_{-i}(\mathbf{y}) - E[R_{-i}(\mathbf{y}^{**})]),$$

where $\mathbf{y}^{**} \equiv \mathbf{y}|\mathbf{x}^{**}$ and

$$\widehat{\alpha}_i = \frac{Cov(\eta_i, R_i(\mathbf{y}^{**}))Var(R_{-i}(\mathbf{y}^{**})) - Cov(\eta_i, R_{-i}(\mathbf{y}^{**}))Cov(R_i(\mathbf{y}^{**}), R_{-i}(\mathbf{y}^{**}))}{Var(R_i(\mathbf{y}^{**}))Var(R_{-i}(\mathbf{y}^{**})) - Cov^2(R_i(\mathbf{y}^{**}), R_{-i}(\mathbf{y}^{**}))} \tag{A-1}$$

$$\widehat{\beta}_{-i} = \frac{Cov(\eta_i, R_{-i}(\mathbf{y}^{**}))Var(R_i(\mathbf{y}^{**})) - Cov(\eta_i, R_i(\mathbf{y}^{**}))Cov(R_i(\mathbf{y}^{**}), R_{-i}(\mathbf{y}^{**}))}{Var(R_i(\mathbf{y}^{**}))Var(R_{-i}(\mathbf{y}^{**})) - Cov^2(R_i(\mathbf{y}^{**}), R_{-i}(\mathbf{y}^{**}))}. \tag{A-2}$$

Agent i solves

$$\max_{x_i} E\left[E[\eta_i | R_i(\mathbf{y}), R_{-i}(\mathbf{y}), \mathbf{x}^{**}] | x_i, x_{-i}^{**}\right] - c(x_i).$$

Notice that

$$E\left[E[\eta_i | R_i(\mathbf{y}), R_{-i}(\mathbf{y}), \mathbf{x}^{**}] | x_i, x_{-i}^{**}\right] = \mu_i + \left(\widehat{\alpha}_i + \lambda_{-i}\widehat{\beta}_{-i}\right)(x_i - x_i^{**}).$$

Hence, x_i^{**} must satisfy the first-order condition:

$$c'(x_i^{**}) = \widehat{\alpha}_i + \lambda_{-i}\widehat{\beta}_{-i},$$

or equivalently, letting $\psi \equiv c'^{-1}$,

$$x_i^{**} = \psi\left(\widehat{\alpha}_i + \lambda_{-i}\widehat{\beta}_{-i}\right).$$

Under public ratings, the principal then solves

$$\max_{(\lambda_1, \lambda_2, \sigma_e) \in \mathbb{R}^2 \times \mathbb{R}_+} \psi\left(\widehat{\alpha}_1 + \lambda_2\widehat{\beta}_2\right) + \psi\left(\widehat{\alpha}_2 + \lambda_1\widehat{\beta}_1\right). \tag{PR1}$$

[19]Again, notice that the random vector $(\eta_i, R_i(\mathbf{y}), R_{-i}(\mathbf{y}))$ is jointly normal.

Remark A1. *Unlike (PR), (PR1) is not subject to the no-disclosure constraint $\lambda_1\lambda_2 = 1$. This is because, with the additional noises, e_i's, in the ratings, the market cannot pin down individual outputs so long as $\sigma_e > 0$.*

Since the focus here is on the choice of σ_e, I fix (λ_1, λ_2) and consider the sub-program:

$$\max_{\sigma_e \in \mathbb{R}_+} \psi\left(\widehat{\alpha}_1 + \lambda_2\widehat{\beta}_2\right) + \psi\left(\widehat{\alpha}_2 + \lambda_1\widehat{\beta}_1\right) \equiv J(\sigma_e).$$

Proposition A1. *Fix (λ_1, λ_2), and consider any cost function satisfying (C1). Then,*

$$J'(0) = 0.$$

That is, the principal cannot strictly benefit from introducing a sufficiently small but positive noise, e_i, to the ratings.

Proof. Fix (λ_1, λ_2), and define $g_i(\sigma_e) = \widehat{\alpha}_i(\sigma_e) + \lambda_{-i}\widehat{\beta}_{-i}(\sigma_e)$, where I make the dependence on σ_e explicit. Then, $J(\sigma_e) = \psi(g_1(\sigma_e)) + \psi(g_2(\sigma_e))$. From (A-1),

$$\widehat{\alpha}_i(\sigma_e) = \frac{Cov(\eta_i, r_i)[Var(r_{-i}) + \sigma_e^2] - Cov(\eta_i, r_{-i})Cov(r_i, r_{-i})}{[Var(r_i) + \sigma_e^2][Var(r_{-i}) + \sigma_e^2] - Cov^2(r_i, r_{-i})}.$$

Suppose $\lambda_i\lambda_{-i} \neq 1$ so that $r_i \neq \lambda_i r_{-i}$ and thus, $Var(r_i)Var(r_{-i}) - Cov^2(r_i, r_{-i}) \neq 0$. Then, it follows that $\widehat{\alpha}_i'(0) = 0$. Similarly, $\widehat{\beta}_{-i}'(\sigma_e) = 0$, implying that $g_i'(0) = 0$ and, in turn, $J'(0) = 0$.

Next suppose $\lambda_i\lambda_{-i} = 1$. Then,

$$\begin{aligned} g_i(\sigma_e) &= \frac{(1+\lambda_i^2)(\sigma_i^2 + \lambda_i\rho\sigma_i\sigma_{-i})}{(1+\lambda_i^2)\left[\left(\sigma_i^2 + \sigma_\varepsilon^2\right) + 2\lambda_i\rho\sigma_i\sigma_{-i} + \lambda_i^2\left(\sigma_{-i}^2 + \sigma_\varepsilon^2\right)\right] + \lambda_i^2\sigma_e^2} \\ g_{-i}(\sigma_e) &= \frac{\lambda_i(1+\lambda_i^2)(\sigma_{-i}^2 + \lambda_i\rho\sigma_i\sigma_{-i})}{(1+\lambda_i^2)\left[\left(\sigma_i^2 + \sigma_\varepsilon^2\right) + 2\lambda_i\rho\sigma_i\sigma_{-i} + \lambda_i^2\left(\sigma_{-i}^2 + \sigma_\varepsilon^2\right)\right] + \lambda_i^2\sigma_e^2}, \end{aligned} \tag{A-3}$$

which can be obtained either by noticing $r_i = \lambda_i r_{-i}$ as in the text or by taking limits of $\widehat{\alpha}_i(\sigma_e)$ and $\widehat{\beta}_{-i}(\sigma_e)$ as $\lambda_{-i} \to 1/\lambda_i$. Evidently, $g_i(0) = \Omega(\lambda_i; \sigma_i, \sigma_{-i}, \rho)$ and $g_{-i}(0) = \Omega(1/\lambda_i; \sigma_{-i}, \sigma_i, \rho)$, as it should be, where $\Omega(\lambda_i; \sigma_i, \sigma_{-i}, \rho)$ is defined in the text. Furthermore, $g_i'(0) = g_{-i}'(0) = 0$, implying that $J'(0) = 0$, as claimed. ■

Remark A2. *Note from (A-3) that $J'(\sigma_e) < 0$ for all $\sigma_e > 0$ whenever $\lambda_i\lambda_{-i} = 1$. Intuitively, if the principal chooses to obscure individual performances by perfectly correlating agents' ratings, she will be strictly worse off by introducing significant noise for the same purpose.*

Corollary A1. *Fix (λ_1, λ_2), and consider $c(x) = kx^2$. Then, $J'(\sigma_e) < 0$ for all $\sigma_e > 0$. That is, with the quadratic cost of effort, the principal is strictly worse off introducing additional noise to agents' ratings.*

Proof. Without loss of generality, let $c(x) = x^2/2$. Then,

$$\begin{aligned} J(\sigma_e) &= g_1(\sigma_e) + g_2(\sigma_e) \\ &= \frac{A\sigma_e^2 + B}{\sigma_e^4 + (A+C)\sigma_e^2 + D}, \end{aligned}$$

where

$$\begin{aligned} A &= \sigma_1^2(1+\lambda_2^2) + \sigma_2^2(1+\lambda_1^2) + 2\rho\sigma_1\sigma_2(\lambda_1+\lambda_2) > 0, \\ B &= (\lambda_1\lambda_2 - 1)^2 \left[(\sigma_1^2+\sigma_2^2)\sigma_\varepsilon^2 + 2\sigma_1^2\sigma_2^2(1-\rho^2)\right] \geq 0, \\ C &= (2+\lambda_1^2+\lambda_2^2)\sigma_\varepsilon^2 > 0, \end{aligned}$$

and

$$D = (\lambda_1\lambda_2 - 1)^2 \left[(\sigma_1^2+\sigma_\varepsilon^2)(\sigma_2^2+\sigma_\varepsilon^2) - \rho^2\sigma_1^2\sigma_2^2\right] \geq 0.$$

It is verified that $(B-D)A + BC > 0$. Hence,

$$J'(\sigma_e) = -\frac{A\sigma_e^4 + 2B\sigma_e^2 + (B-D)A + BC}{\left(\sigma_e^4 + (A+C)\sigma_e^2 + D\right)^2} 2\sigma_e < 0 \text{ for all } \sigma_e > 0.$$

■

References

[1] Arya, Anil, and Brian Mittendorf. “The benefits of aggregate performance metrics in the presence of career concerns.” *Management Science* 57, no. 8 (2011): 1424-1437.

[2] Auriol, Emmanuelle, Guido Friebel, and Lambros Pechlivanos. “Career concerns in teams.” *Journal of Labor Economics* 20, no. 2 (2002): 289-307.

[3] Bar-Isaac, Heski. “Something to prove: reputation in teams.” *RAND Journal of Economics* 38, no. 2 (2007): 495-511.

[4] Bloomfield, Matthew J., Catarina Marvao, and Giancarlo Spagnolo. “Relative performance evaluation, sabotage and collusion.” *Journal of Accounting and Economics* (2023): 101608.

[5] Chalioti, Evangelia. “Team production, endogenous learning about abilities and career concerns.” *European Economic Review* 85 (2016): 229-244.

[6] Che, Yeon-Koo, and Seung-Weon Yoo. “Optimal incentives for teams.” *American Economic Review* 91, no. 3 (2001): 525-541.

[7] Cohan, Peter. “Adobe’s Stock Up 68% Since it Dumped Stack Ranking, Will Microsoft Follow?” *Forbes* (2013)

[8] Dewatripont, Mathias, Ian Jewitt, and Jean Tirole. “The economics of career concerns, part I: Comparing information structures.” *Review of Economic Studies* 66, no. 1 (1999): 183-198.

[9] Eichenwald, Kurt. “Microsoft’s lost decade.” *Vanity Fair* 54, no. 8 (2012).

[10] Ekmekci, Mehmet. “Sustainable reputations with rating systems.” *Journal of Economic Theory* 146, no. 2 (2011): 479-503.

[11] Fleckinger, Pierre. “Correlation and relative performance evaluation.” *Journal of Economic Theory* 147, no. 1 (2012): 93-117.

[12] Fleckinger, Pierre, David Martimort, and Nicolas Roux. “Should They Compete or Should They Cooperate? The View of Agency Theory.” *Journal of Economic Literature*, Forthcoming.

[13] Gibbons, Robert, and Kevin J. Murphy. "Relative performance evaluation for chief executive officers." *ILR Review* 43, no. 3 (1990): 30-S.

[14] Green, Jerry R., and Nancy L. Stokey. "A comparison of tournaments and contracts." *Journal of Political Economy* 91, no. 3 (1983): 349-364.

[15] Gromb, Denis, and David Martimort. "Collusion and the organization of delegated expertise." *Journal of Economic Theory* 137, no. 1 (2007): 271-299.

[16] Holmström, Bengt. "Moral hazard and observability." *Bell Journal of Economics* (1979): 74-91.

[17] Holmström, Bengt. "Moral hazard in teams." *RAND Journal of Economics* (1982): 324-340.

[18] Holmström, Bengt. "Managerial incentive problems: A dynamic perspective." *Review of Economic Studies* 66, no. 1 (1999): 169-182.

[19] Hörner, Johannes, and Nicolas S. Lambert. "Motivational ratings." *Review of Economic Studies* 88, no. 4 (2021): 1892-1935.

[20] Itoh, Hideshi. "Incentives to help in multi-agent situations." *Econometrica* (1991): 611-636.

[21] Legros, Patrick, and Steven A. Matthews. "Efficient and nearly-efficient partnerships." *Review of Economic Studies* 60, no. 3 (1993): 599-611.

[22] Lazear, Edward P., and Sherwin Rosen. "Rank-order tournaments as optimum labor contracts." *Journal of Political Economy* 89, no. 5 (1981): 841-864.

[23] Liu, Qingmin, and Andrzej Skrzypacz. "Limited records and reputation bubbles." *Journal of Economic Theory* 151 (2014): 2-29.

[24] Ma, Ching-To. "Unique implementation of incentive contracts with many agents." *Review of Economic Studies* 55, no. 4 (1988): 555-572.

[25] McNabb, David E., and Carl R. Swenson. *Collaboration in Government: Forms and Practices*. Routledge, 2021.

[26] Meagher, Kieron, and Suraj Prasad. "Career concerns and team talent." *Journal of Economic Behavior & Organization* 129 (2016): 1-17.

[27] Meyer, Margaret A. "The dynamics of learning with team production: Implications for task assignment." *Quarterly Journal of Economics* 109, no. 4 (1994): 1157-1184.

[28] Meyer, Margaret A., and John Vickers. "Performance comparisons and dynamic incentives." *Journal of Political Economy* 105, no. 3 (1997): 547-581.

[29] Miller, Nolan H. "Efficiency in partnerships with joint monitoring." *Journal of Economic Theory* 77, no. 2 (1997): 285-299.

[30] Mookherjee, Dilip. "Optimal incentive schemes with many agents." *Review of Economic Studies* 51, no. 3 (1984): 433-446.

[31] Nalebuff, Barry J., and Joseph E. Stiglitz. "Prizes and incentives: towards a general theory of compensation and competition." *Bell Journal of Economics* (1983): 21-43.

[32] Neeman, Zvika, Aniko Öry, and Jungju Yu. "The benefit of collective reputation." *RAND Journal of Economics* 50, no. 4 (2019): 787-821.

[33] Ortega, Jaime. "Power in the firm and managerial career concerns." *Journal of Economics & Management Strategy* 12, no. 1 (2003): 1-29.

[34] Pei, Harry. "Reputation Effects under Short Memories." Working Paper, 2023.

[35] Rodina, David. "Information design and career concerns." Working paper, 2017.

[36] Tong, Y.L. *The Multivariate Normal Distribution*. Switzerland, Springer New York, 2012.

[37] Wijayanti, Aniek, Mahfud Sholihin, and Ertambang Nahartyo. "What do we know about the forced distribution system: a systematic literature review and opportunities for future research." *Management Review Quarterly* (2024): 1-42.

[38] Van der Hoek, Marieke, Sandra Groeneveld, and Ben Kuipers. "Goal setting in teams: Goal clarity and team performance in the public sector." *Review of Public Personnel Administration* 38, no. 4 (2018): 472-493.

[39] Yildirim, Huseyin. "The economics of career concerns in teamwork," Working Paper, 2024.