DISCUSSION PAPER SERIES

Discussion paper No.261

Evolution of Information Projection Bias through Costly Communication in Overlapping Generations Organizations

Kohei Daido

(Kwansei Gakuin University)

Tomoya Tajika

(Nihon University)

October 2023

SCHOOL OF ECONOMICS

KWANSEI GAKUIN UNIVERSITY

1-155 Uegahara Ichiban-cho

Nishinomiya 662-8501, Japan

Evolution of Information Projection Bias through Costly Communication in Overlapping Generations Organizations *

Kohei Daido[†] Tomoya Tajika[‡]

October 18, 2023

Abstract

In organizations with overlapping generations, behavioral bias affects performance through promotion decisions. This study focuses on information projection bias and examines its effects on communication efforts and the overall performance of an organization adopting a performance-based promotion system to select the next-generation manager among current subordinates. We show that the bias generally disrupts communication between an incumbent manager and subordinates and that the expected overall performance is single-peaked with respect to the manager's bias. When considering the bias distribution among newly promoted managers, we find that a more biased group is likely to select a more biased manager. This trend becomes stronger over generations and the expected overall performance increases when the variety of the bias degree is restricted and communication efforts are complements. By contrast, in a competitive organization, the manager's bias diminishes over generations. Nonetheless, the overall performance decreases when the variety of the bias degree is sufficient. Our results contribute to the understanding of the effects of diversity in an organization on its performance.

JEL Classification: D82, D91, M51

Keywords: Information Projection Bias, Costly Communication, Overlapping Generations, Performance-based Promotion, Diversity

*We thank Takeshi Murooka, the participants at Hitotsubashi Summer Institute, DC conference in Japan, and a seminar at Waseda University for their valuable comments. Daido thanks JSPS KAKENHI Grant Number JP19K01568 and JP18H03640.

†Kwansei Gakuin University, email: daido@kwansei.ac.jp

‡Nihon University, email: tajika.tomoya@nihon-u.ac.jp

1 Introduction

In an organization, it is common for members to communicate with each other to exchange necessary work-related information. The worth of communication also depends on whether the other party has the information we require and their willingness to communicate it to us. The belief that the other party has information typically depends on whether we have the information or not. If a person has information, they tend to believe that others also have it; if they do not have it, they then tend to believe that others also do not have it. This phenomenon is called *information projection bias* (Madarász, 2012). Previous literature on information projection teaches us that information projection bias impedes communication because value-added communication occurs between people who have the information and those who do not. In organizations, a lack of communication due to this bias affects worker performance. Subsequently, the heterogeneity of the degree of bias among workers makes such an effect more significant. When a long-lasting organization adopts a system in which the worker who achieved the highest performance is promoted to the position of next-generation manager, the heterogeneity will influence the distribution of the promoted workers' bias. Moreover, the performance of the organization depends on the dynamics of the bias when interacting with new incoming workers. By observing these effects, we can evaluate such a performance-based promotion system from the perspective of information projection bias. This study aims to analyze the dynamics of the distribution of bias in an organization, its effects on the performance of the organization, and the efficiency of a performance-based promotion system.

In Section 2, we posit an overlapping generation (OLG) model in which one of the subordinates is promoted to the position of next-generation manager. All players, including the manager and subordinates, allocate their efforts between individual tasks and communication, and both contribute to performance. To improve performance via communication, managers, and subordinates must cooperate if the manager has information and the subordinates do not. Our main analysis supposes that communication efforts are complements.[1] When players decide to contribute to communication, they must determine whether their counterparts have information. This reasoning

[1] We provide supplementary material for the case where communication efforts are substitutes.

tends to be influenced by information projection bias: players project their information state onto others. The manager cares about the overall performance of the organization, which is the sum of the performance of all players in the organization, whereas the subordinate cares only about their performance. To confirm the effects of information projection bias on performance and the dynamics of the bias in the organization, we focus on a performance-based promotion system in which the subordinate who achieves the highest performance is promoted to the position of next-generation manager.

In Section 3, we first define the equilibrium when players have information projection bias. All players project their current information state (i.e., whether they have information or not) onto their counterpart's state. To focus on the case where communication is significant, we assume that managers have information but subordinates do not.[2] Under such a setup with additional assumptions, we show that information projection bias inhibits players from communicating with each other. Consequently, the expected overall performance is single-peaked with respect to the manager's bias, while the manager's bias decreases the expected performance of subordinates. The performance of each subordinate is single-peaked with respect to their bias; however, the peak is less than the manager's bias.[3] This implies that the subordinate yielding the highest performance tends to have a smaller bias than the manager.

Section 4 focuses on the performance-based promotion system to examine how the bias of a promoted subordinate depends on the current manager's bias and how the dynamics of bias influence the performance of the organization. In Section 4.1, we demonstrate that if the manager has a stronger bias, the subordinate with the stronger bias is more likely to be promoted. Although this property does not imply that bias becomes stronger over generations, we show that more biased groups are likely to select a more biased manager, and this trend becomes stronger over generations when the number of subordinates is finite and the variety of the bias degree is two. In this case, the upward trend of the manager's bias strength benefits the organization: the expected overall performance increases. Because of the complementarity of the value of communication, the

[2] In Section 5.1, we consider the case where subordinates possibly have information.

[3] Hereafter, we use female pronouns to refer to a manager and male pronouns to refer to a subordinate.

stronger the bias, the lower the effort exerted for communication by managers and subordinates. Consequently, when a manager is more strongly biased, subordinates with a stronger bias yield higher performance.[4]

The previous analysis limits the variety of the bias degree. Section 4.2 studies the case where the variety of the bias degree is sufficient in a large organization where the number of subordinates is infinite.[5] In contrast to Section 4.1, the manager's bias diminishes over time. Further, when bias distribution has continuum support, we demonstrate that the manager's bias converges to the least possible value. Although this increases communication by managers, the subordinates' contribution to communication remains unchanged because their bias distributions are the same. Subsequently, the manager's communication becomes excessive as the complementarity of communication efforts implies that subordinates require more communication efforts. This result leads to a decline in overall organizational performance in the long run.

Our results relate to findings in the literature on the effect of diversity in an organization on its performance.[6] In most studies, diversity causes a main trade-off for communication among workers in an organization: the negative effects of communication costs (difficulties) and the positive effects of information (or skill) complementarity. The type and source of diversity (e.g., ethnicity, nationality, or skill) depends on the research. This study captures the source of diversity as the variety of elements in the support of the subordinate's bias distribution. The diversity among subordinates affects the value of communication because of information projection bias between the manager and subordinates and competition for promotion among subordinates. Thus, this study examines how diversity among subordinates influences communication and organizational performance through the lens of information projection bias. Our results indicate that a less diversified

[4] We provide a numerical calculation to investigate the case in which the support of bias degree distribution has more elements in Appendix C. Consequently, a newly promoted manager's bias can be greater or smaller than that of the old incumbent manager. The expected overall performance increases in both cases but is more likely to increase when the bias increases.

[5] The reason for focusing on large organizations is for tractability. Further, examining large organizations enables us to consider a large variety of degree bias because such organizations potentially have a wide array of people with diversity in languages, races, or nationalities.

[6] For example, a field experiment by Lyons (2017) finds that team organization decreases outcomes when workers' skills are diverse and when they are nationally diverse. Lyons (2017) comprehensively introduces other related literature.

organization increases its performance under the performance-based promotion system.

In Section 5, we discuss some alternative models by extending the main model. We first examine the case of the manager's retrospective information projection: Although the subordinate projects his current information state to the manager's state, the manager projects her previous information state, which she had when she was a subordinate, to her subordinate's current state. Then, we discuss some possible scenarios of overconfidence as a psychological bias related to information projection bias.

Related Literature First, this study contributes to the literature on information projection bias. Madarász (2012) formalizes information projection bias in how agents project their private information and exaggerate how others know it. Madarász (2016) study interpersonal projection bias of belief by defining projection equilibrium and Danz, Madarász, and Wang (2018) experimentally test such bias in an agency setting.[7] They find that people project their information onto others and anticipate others' projections onto them, although people underestimate it.[8] This study extends the notion of information projection to an OLG model where all players project their information state onto others. This extension enables us to study whether bias is pervasive over time in an organization and the effects caused by the dynamics of bias on overall performance.

Next, communication in our model relies on the literature on costly communication (Bolton and Dewatripont , 1994; Garicano , 2000; Dewatripont and Tirole, 2005; Dessein and Santos , 2006; Dessein, Galeotti, and Santos , 2016). A recent study by Battiston, Blanes I Vidal, and Kirchmaier (2021) investigates the relationship between communication and performance in organizations. Through their field experiment, Battiston, Blanes I Vidal, and Kirchmaier (2021) show that sequential communication improves the performance of the receiver at the cost of its sender. Therefore, communication plays a key role in contexts in which the informed party helps the less-

[7] The definition of equilibrium in this study differs from that in Madarász (2016). In section 3, we explain the differences in detail.

[8] Related to information projection bias, Gagnon-Bartsch (2016), Gagnon-Bartsch, Pagnozzi, and Rosato (2021), and Gagnon-Bartsch and Rosato (2022) study taste projection. Bushong and Gagnon-Bartsch (forthcoming) show that the interpersonal taste projection bias leads to substantial and costly errors when people forecast others' behavior in a real-effort experiment.

informed to improve performance. In contrast to Battiston, Blanes I Vidal, and Kirchmaier (2021), we consider simultaneous communication to focus on players' information projection bias. Furthermore, we incorporate the notion of projection bias into communication and study its effects on players' contributions to communication. Thus, we examine organizational performance through communication influenced by information projection bias.

Finally, this study contributes to the literature on the evolution of psychological bias. Among such biases, substantial literature exists on the evolution of overconfidence. For example, Johnson and Fowler (2011) study the evolution of overconfidence, and their evolutionary model shows that overconfident populations are evolutionarily stable.[9] By contrast, focusing on information projection bias instead of overconfidence, we examine the effects of bias on the overall performance of an organization through communication and promotion, which are crucial for organizational practice. Moreover, we investigate the dynamics of bias distribution over generations from the perspective of stochastic dominance among the bias distributions and demonstrate the degeneration of bias in an organization.[10] Note that managers and subordinates have different bias distributions, which differ from standard models of evolution games. In particular, while the bias distribution of subordinates remains unchanged, that of managers evolves. Consequently, although the manager's bias degree diminishes, performance decreases in a large organization.

2 Model

We consider an OLG organization comprising a manager and n subordinates at each period t. We denote m as a manager and s as a typical subordinate. The manager at t is selected from the subordinates at $t-1$.[11] Each subject decides the effort allocation $(e_k, 1-e_k)$ where $e_k \in [0,1]$ and $k \in \{m, s\}$. e_k contributes to the individual task, and its outcome is represented by $u(e_k)$. $1-e_k$ pro-

[9] For other studies on overconfidence and its evolution, see, for example, Heller (2015) on preference for privately informed overconfident agents in a strategic situation; Bernardo and Welch (2001) on an overconfident individual with private information in a group; and Goel and Thakor (2008) on the selection of an overconfident CEO.

[10] Although the research topic differs from ours, Heller and Nehama (2023) study the evolution of risk attitudes in a population by focusing on the distributions of risk preferences.

[11] We omit t to economize the notation when it is obvious.

motes communication for sharing information when the subordinate does not have it, whereas the manager does. In each period, the manager and each subordinate obtain information with probabilities p_M and p_S, respectively. The required information differs by period. The state of information contributes to the performance of the subordinate.[12] If the subordinate obtains information by himself, the maximum performance from the state of information is archived, represented by V. Otherwise, the subordinate partially obtains it through costly communication with the manager, if she knows the information. For example, consider a situation where managers are more likely to possess information including work tips, visions of the ongoing project, and similar information through their experience relative to their subordinates.[13] Precisely, communication efforts by the subordinate and the manager, $(1 - e_s, 1 - e_m)$, can improve the performance from the state of information, represented by $v(e_s, e_m)$, which can be interpreted as the value of communication.

The subordinate s's performance depends on the performance shock and the performance of the individual task as well as the state of information, which is represented by $\pi_s = I_s V + (1 - I_s) I_m v(e_s, e_m) + u(e_s) + \theta_s$. $I_k \in \{0, 1\}$ is the indicator of whether player $k \in \{s, m\}$ knows the information. If player k knows it, $I_k = 1$, and otherwise, $I_k = 0$. The term θ_s is the performance shock of the subordinate s, distributed by the cumulative distribution function F on $\mathbb{R}$, and the average is denoted by $\bar{\theta}$. This term represents idiosyncratic factors contributing to performance (e.g., individual ability) other than information and individual tasks. We assume that this value is persistent over time. The organization's overall performance is supposed to be $\Pi = \theta_m + nu(e_m) + \sum_{j \in N} \pi_j$.[14] We suppose that each subordinate maximizes their performance π_i and the manager maximizes the overall performance Π, which implies that the manager and subordinates have no conflict of interest. Therefore, equilibrium is efficient when players do not have information projection bias. This setting enables us to focus on the pure effects of the bias on the players' decisions and the

[12] We can modify the model to ensure that the information also improves the manager's performance. However, when managers transmit information to their subordinates, this modification does not affect communication efforts, and the results do not change.

[13] As another example of information, consider a laboratory with one senior researcher (professor) and some junior researchers. The senior researcher has several ideas and delegates their implementation to the junior researchers. The ideas are the information in the model.

[14] We assume that the output of the manager's task $u(e_m)$ is n times more significant than that of the subordinate to exclude the effect of n from effort levels and simplify the analysis with respect to n.

outcome of the promotion. We refer to the promotion system as *performance-based* if the subordinate yielding the highest performance among the subordinates is promoted to the position of next-generation manager.

We assume the following properties on $v(\cdot)$ and $u(\cdot)$:

Assumption 1. $v\colon [0,1]^2 \to \mathbb{R}$ and $u\colon [0,1] \to \mathbb{R}$ are twice continuously differentiable and satisfy the following conditions.

(i) $V > v(e_s, e_m)$ for any e_s, e_m.

(ii) $v_1 < 0$, $v_2 < 0$, $u' \geqslant 0$.

(iii) $v_{11} < 0$, $v_{22} < 0$, $u'' < 0$.

(iv) $u''(e_s) \cdot u''(e_m) > v_{12}(e_s, e_m) \cdot v_{21}(e_s, e_m)$ for any e_s, e_m.

(v) $v_{12} > 0$.

(i) implies that the value of communication is less than that of initially having the information. (ii) assumes that the value of communication is increasing in both efforts on communication, $1 - e_s$ and $1 - e_m$, and the output of the individual task, u, is increasing in their effort on it. These effects are diminishing, as represented by (iii). (iv) implies the concavity of the overall performance concerning effort bundle e_s, e_m. By (v), we focus on the case where communication efforts have complementarity.

We also make the following assumption regarding the distribution of θ, f.

Assumption 2. $\lim_{x\to\infty} \frac{f(x-a)}{f(x-b)} = 0$ for each $b > a$.

For example, normal distributions satisfy this assumption.

Remark 1. In addition to the value of communication ($v(e_s, e_m)$), the formulation of the value of information, which depends on all players' information states, is also crucial. In particular, we suppose that the value of information is substitutable: if a manager or subordinate has the information, the subordinate obtains that value, as Table 1 (a) shows. Communication becomes

manager \ subordinate	Knows	Not
Knows	V	$v(e_s, e_m)$
Not	V	0

(a) Main model

manager \ subordinate	Knows	Not
Knows	$v(e_s, e_m)$	0
Not	0	0

(b) Complementarity of information

Table 1: The value of information

effective only when the manager has the information but the subordinate does not. Another possible setup is such that the knowledge is characterized by complementarity in the information value, as Table 1 (b) shows. In this case, the projection bias works in the opposite direction; as the degree of the bias strengthens, the exerted effort on communication increases.

Whether the value of information is a substitute or complement depends on the situations facing organizations. For example, when the necessary information for subordinates to improve performance is ubiquitous and difficult for them to find, the advice from their manager, who has the information, raises the value of the information. This situation falls under the substitutability of information, where at least one party who has the information is sufficient to improve performance. By contrast, when both the subordinate and manager have information, connecting such ubiquitous information has a synergy effect, which enhances the value of information. This situation can be applied to the complementarity of information, in which both parties' contributions are required to improve performance.

3 Equilibrium with Information Projection Bias

3.1 Definition of Equilibrium

We suppose that a subordinate misperceives the manager's state or information, which, in turn, influences his formation of belief; when the subordinate knows (or does not know) the information, he projects the fact to the manager's state, believing that the manager must (or must not) know the information. The manager also has an information projection bias: the manager projects her current information state to subordinates' states. The degree of information projection bias is represented by α for the typical subordinate and $\hat{\alpha}$ for the typical manager. We assume that when the subjects make a decision, they suppose that others have no bias.

Subordinate's game We assume that each subordinate s plays the following incomplete information game defined below. We first formulate two types of beliefs held by the subordinate, whose degree of bias is α. Let $P^{\alpha}_{I_s}(I_m = 1)$ be the probability with which, under the information state $I_s \in \{0, 1\}$, the subordinate with α believes that the manager knows the information ($I_m = 1$). As the subordinate has an information projection bias with a degree of α, this probability depends on the prior probability of the manager having the information (p_M) with probability $1 - \alpha$ and on the subordinate's information state (I_s) with probability α:

$$P^{\alpha}_{I_s}(I_m = 1) = (1 - \alpha)p_M + \alpha I_s.$$

We then consider the subordinate's decision. Each subordinate maximizes their performance, given that the aforementioned beliefs are common knowledge. When the subordinate obtains the information, the performance is $\theta_s + V + u(e_s)$, in which case $e_s = 1$ is optimal regardless of the subordinate's beliefs.[15] When the subordinate fails to gain the information, the expected performance is

$$\theta_s + P^{\alpha}_0(I_m = 1) \cdot v(e_s, e^*_m) + u(e_s).$$

Note that performance depends on the choice of the manager with I_m, represented by e^*_m. We assume that in the subordinate's belief, the manager knows the true probability. The subordinate expects that the choice of the manager is

$$e^*_m = \arg\max_{e_m} \left[nu(e_m) + \sum_{s \in N} \left\{ p_S \cdot (V + u(1)) + (1 - p_S) \cdot (v(e_s, e_m) + u(e_s)) \right\} \right].$$

[15]This assumption implies that there is no rational player in the current model. Even when $\alpha = 0$, each player believes that their counterpart knows the true probability, although this belief does not necessarily coincide with the true belief of the counterpart, owing to their bias.

We denote the subordinate's optimal choice by e_s^{α}.

Manager's game As with the case for the subordinate, the manager with bias $\hat{\alpha}$ plays the following incomplete information game. From the manager's perspective, the subordinate believes that the manager has the current information with probability p_M.

Given this setting, we consider the manager's choice. The manager maximizes the following expected overall performance:

$$nu(e_m) + \sum_{s \in N} \left\{ P_{I_m}^{\hat{\alpha}}(I_s = 1) \cdot (V + u(1)) + P_{I_m}^{\hat{\alpha}}(I_s = 1) \cdot (v(e_s^*, e_m) + u(e_s^*)) \right\}$$

where the subordinate's choice in the manager's game is given by

$$e_s^* = \arg\max_{e_s} \left[\theta_i + p_M \cdot v(e_s, e_{m,1}) + u(e_s) \right].$$

We consider the tuple $(e_s^{\alpha}, e_m^{\hat{\alpha}})$ as an equilibrium.

Remark 2. The definition of the equilibrium differs from the information projection equilibrium defined by Madarász (2012) in the following two aspects.

First, in Madarász (2012), information projection bias affects behavior via the belief of the other players' actions. He defines the information projection equilibrium as follows.[16]

Definition 1 (Madarász (2012)). A strategy profile σ^{α} is a α information projection equilibrium if there is a strategy profile σ^{+} such that for each i, σ_i^{α} is the best response against the strategy that σ_{-i}^{α} is taken with probability $1 - \alpha$, and σ_{-i}^{+} is taken with probability α. Here σ_{-i}^{+} is the best response against σ_i^{α} under the belief that the other players have the same information for player i.

However, in the current model, the knowledge of the manager and subordinates directly affects the payoffs. If the manager does not know the information, whatever the players do, their efforts in communication will never produce benefits for both players. For example, the difference occurs

[16] In Madarász (2021), similar to our study, the equilibrium is defined as a Bayesian Nash equilibrium with noncommon priors.

in the following situation. Suppose that $\alpha = 1$ and $v_1(e_s, 1) \neq 0$ and that the subordinate does not have information. In the setting of Madarász (2012), the subordinate believes that the manager behaves as if the manager does not have information, which implies that the manager never devotes effort to communication. In this case, the projected belief does not directly affect the subordinate's expected payoff, and the subordinate has an incentive to exert effort in communication. However, in our model, as the subordinate believes that the manager does not have the information with a probability of one, the subordinate never exerts effort for communication. Nevertheless, this difference is not crucial for our results.

Second, the degree of the projection bias is heterogeneous. On this point, Danz, Madarász, and Wang (2018) suppose that players know others' degrees of bias when the players have different degrees of bias. Conversely, this study supposes that players believe that others have no bias.

3.2 Equilibrium Effort Choices

Hereafter, we suppose that $p_S = 0$ and $p_M = 1$, focusing on the case where communication is crucial for performance.[17] First, considering the optimization problem of the subordinate who does not have information, we obtain the optimal effort level, e_s^α, as a solution to the following system of equations:

$$\begin{aligned} P_{I_s}^{\alpha}(I_m = 1) \cdot v_1(e_s^\alpha, e_m) + u'(e_s^\alpha) = 0, \\ u'(e_m) + (1 - p_S) \cdot v_2(e_s^\alpha, e_m) = 0. \end{aligned} \tag{1}$$

Next, the optimal effort level of the manager who has the information is a solution to the following a system of equations:

$$\begin{aligned} p_M \cdot v_1(e_s, e_m^{\hat{\alpha}}) + u'(e_s) = 0, \\ u'(e_m^{\hat{\alpha}}) + P_{I_m}^{\hat{\alpha}}(I_s = 0) \cdot v_2(e_s, e_m^{\hat{\alpha}}) = 0. \end{aligned} \tag{2}$$

[17] We examine the case of $p_S > 0$ in Section 5.1.

To guarantee the interior solution, we assume the following.

Assumption 3. (i) $u'(1) = 0$.

(ii) $-u'(0) > \max\{v_1(0,0), v_2(0,0)\}$.

Now, we obtain the following result by showing that the solutions to (1) and (2) uniquely exist.

Lemma 1. *Under Assumptions 1 and 3, for any* $\alpha \in [0,1]$, *there exists a unique equilibrium.*

The effects of bias on the optimal effort levels can be summarized as follows:

Lemma 2. $\frac{\partial e_m^\alpha}{\partial \alpha} > 0$ *and* $\frac{\partial e_s^\alpha}{\partial \alpha} > 0$.

Lemma 2 indicates that when the communication efforts have complementarity, communication efforts decrease in the projection bias (recall that communication efforts are captured by $1-e_k$). Communication is necessary when the information is asymmetric. As projection bias strengthens the belief that the players have the same information, they tend not to share it.

We provide an example that satisfies Assumption 1 and where the optimal effort choices behave as described in Lemma 2.

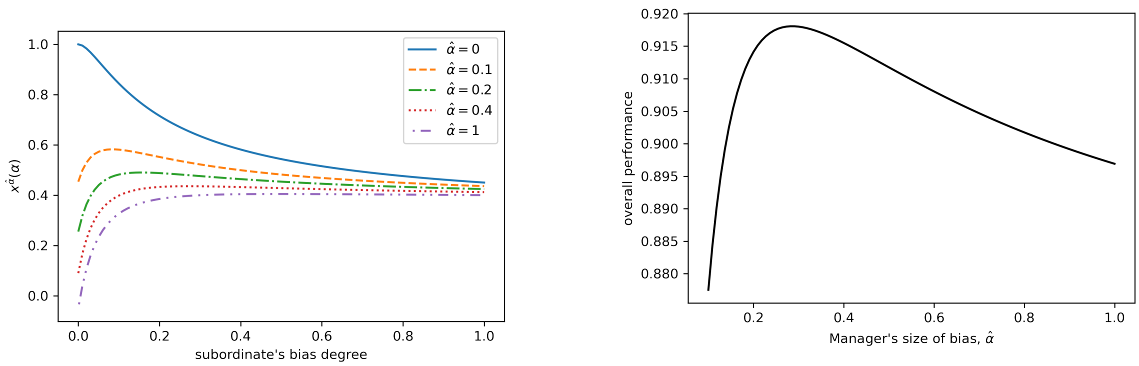
Example 1. Consider the following specification:

$$v(e_s, e_m) = -\frac{\gamma}{2}(e_s)^2 + \gamma_{ms}(1-e_m)(1-e_s) - \frac{\gamma}{2}(e_m)^2$$
$$u(e) = \frac{\zeta}{2}(1-(1-e)^2) = \zeta(e - e^2/2)$$

where γ and ζ are positive. Assumption 1 requires that $\max\{\gamma, \zeta\} > |\gamma_{ms}|$. When $\gamma_{ms} > 0$, communication efforts are complements. By the first order condition, we have

$$e_s^\alpha = e_m^\alpha = \frac{\zeta(\gamma+\zeta) - (1-\alpha)\gamma_{ms}(\gamma_{ms}+\gamma)}{((1-\alpha)\gamma+\zeta)(\gamma+\zeta) - (1-\alpha)\gamma_{ms}^2}$$

As Lemma 2 shows, $\frac{\partial e_s^\alpha}{\partial \alpha} > 0$ if $\zeta > \gamma_{ms}$, which is implied by Assumption 1.



(a) Subordinate's performance ($x^{\hat{\alpha}}(\alpha)$) and bias

(b) The overall performance and degree of the manager's bias (G is uniform distribution)

Figure 1: Single-peakedness of the subordinate's and overall performance

3.3 Overall Performance

When $p_S = 0$ and $p_M = 1$, the performance of the subordinate is $u(e_s^{\alpha}) + v(e_s^{\alpha}, e_m^{\hat{\alpha}}) + \theta_i$. Hereafter, we consider the performance attributed to efforts, $u(e_s^{\alpha}) + v(e_s^{\alpha}, e_m^{\hat{\alpha}})$, to focus on the roles of biases. The following lemma summarizes the properties of the relationship between the subordinate's performance and bias (Figure 1 (a) shows a numerical plot). Hereafter, we consider *symmetric* v, assuming that $v(e_s, e_m) = v(e_m, e_s)$ holds for any e_s, e_m.

Lemma 3. *Let $x^{\hat{\alpha}}(\alpha) = u(e_s^{\alpha}) + v(e_s^{\alpha}, e_m^{\hat{\alpha}})$, which is the subordinate's performance. Then,*

(a) $x^{\hat{\alpha}}(\alpha)$ is decreasing in $\hat{\alpha}$, and is single-peaked with respect to α.

(b) If v is symmetric, the maximizer of $x^{\hat{\alpha}}(\alpha)$ is no greater than $\hat{\alpha}$: $\arg\max_{\alpha} x^{\hat{\alpha}}(\alpha) \leqslant \hat{\alpha}$.

(c) If $v_{112} \leqslant 0$, $\arg\max_{\alpha} x^{\hat{\alpha}}(\alpha)$ is increasing in $\hat{\alpha}$.

(d) $\frac{\partial^2 x^{\hat{\alpha}}(\alpha)}{\partial\alpha\partial\hat{\alpha}} > 0$.

Let the bias of the manager, $\hat{\alpha}$, be fixed, and consider the overall performance when α follows the distribution function, G.

Subsequently, when the manager's bias is $\hat{\alpha}$, the expected overall performance (divided by $|N|$

and subtracting θ terms to focus on the effect of $\hat{\alpha}$) is

$$E[\Pi \mid \hat{\alpha}] = u(e_m^{\hat{\alpha}}) + \int dG(\alpha)[u(e_s^{\alpha}) + v(e_s^{\alpha}, e_m^{\hat{\alpha}})].$$

We have the following proposition:

Proposition 1. $\frac{\partial E[\Pi|\hat{\alpha}]}{\partial \hat{\alpha}}$ *is single-peaked with respect to* $\hat{\alpha}$. *If* v_{12} *is constant, the maximizer is increasing in* $\int dG(\alpha)e_s^{\alpha}$.

Note that in Example 1, v_2 is linear in the first argument. In this case, if $v_{21} > 0$, as $e_s^{\hat{\alpha}}$ is strictly increasing in $\hat{\alpha}$, $E[\Pi \mid \hat{\alpha}]$ is inverted U shape and its peak is $\varphi(\int dG(\alpha)e_s^{\alpha})$, where φ is the inverse of $e_s^{\hat{\alpha}}$ with respect to $\hat{\alpha}$.

By Lemma 2, as $\frac{\partial e_s^{\alpha}}{\partial \alpha} > 0$, the peak $\varphi(\int dG(\alpha)e_s^{\alpha})$ is increasing when G becomes larger in the sense of the first order stochastic dominance.

In summary, with some conditions, we obtain the following results:

1. The effort of each subordinate toward individual task, e_s^{α}, is increasing in α.

2. The effort of the manager toward individual task, $e_m^{\hat{\alpha}}$, is increasing in $\hat{\alpha}$.

3. The expected performance of each subordinate is single-peaked with respect to α and decreasing in $\hat{\alpha}$

4. The expected overall performance is single-peaked with respect to $\hat{\alpha}$.

First, information projection bias influences communication efforts as follows. The subordinate who has no information is more likely to believe that the manager does not have it. This belief makes the subordinate reluctant to communicate with the manager because communication has no value when the manager does not have information. This result is also true for the manager when communication efforts have complementarity.

Next, the degree of bias influences the performance of the subordinate as follows. As shown in statement 1, the low-bias subordinate exerts more communication efforts. Whether the subordinate's communication efforts raise performance depends on the level of the manager's communication efforts. When communication efforts have complementarity, performance decreases (increases) if the manager's communication effort level is low (high), which is more likely to happen when the manager's degree of bias is high (low), as shown in statement 2. Since the low-biased subordinate tends to believe that the manager's bias must be low, the negative effects of the subordinate's wasted communication efforts on performance are more serious when the manager is highly biased against the subordinate's belief.

As each player projects their bias onto others, the subordinate with the same degree of bias can coordinate their actions with their manager. Communication yields a benefit more than the player believes; a subordinate who has a smaller α than the manager's bias is likely to yield greater performance. This explains why the peak of $x^{\hat{\alpha}}(\alpha)$ is no greater than the manager's bias. From the perspective of the manager's bias, the expected total profit is greater if the manager's bias is closer to the average of the subordinates' biases.

4 Bias Distribution under the Performance-based

We explore how the bias of a promoted subordinate depends on the current manager's bias under the performance-based promotion system: the most productive subordinate is promoted. Let $\Gamma(\cdot \mid \hat{\alpha}, G)$ be the distribution function of the bias of the subordinate who yields the highest performance.[18] This distribution is conditioned on the case in which the manager's bias is $\hat{\alpha}$ and the current subordinates' bias distribution is G.

Suppose that the manager's bias $\hat{\alpha}$ is distributed by $\hat{G}$. Subsequently, we denote $\mathcal{G}_{\hat{G},G}$ as the unconditional distribution function of the bias of the subordinate with the highest performance. Precisely, $\mathcal{G}_{\hat{G},G}(\tilde{\alpha}) = \int d\hat{G}(\hat{\alpha})\Gamma(\tilde{\alpha} \mid \hat{\alpha}, G)$. Thus, $\mathcal{G}$ is the law of motion of the bias distribution under the performance-based promotion system. Specifically, if $\hat{G}_t$ is the current manager's bias distribution, the next-generation manager's bias distribution is $\hat{G}_{t+1} = \mathcal{G}_{\hat{G}_t,G}$. In the following

[18] The formal derivation is relegated to Appendix A.

subsections, we observe the properties of these distributions and their relationships with overall performance to determine whether the performance-based promotion system strengthens the bias and increases overall performance.

4.1 Law of Motion in Bias Distribution

This subsection investigates the condition under which the bias of the manager increases over time. To simplify the analysis, we also focus on $\mathrm{supp}(G) = \{0, 1\}$. By abusing the notation, we consider the value of the density function as the probability assigned to it. We have the following proposition.

Proposition 2. *Suppose that* $\mathrm{supp}(G) = \{0, 1\}$*, v is symmetric,* $v(1, e_m) = v(e_s, 1) = 0$ *for any e_s, e_m, and F is a normal distribution:*

(a) There is $\bar{g} \in (0, 1)$ such that $\mathcal{G}_{G,G} \succ_{\mathrm{sd}} G$ if $g(1) > \bar{g}$.

(b) $\bar{g}$ increases if the variance of F increases.

(c) If $g(1)$ is sufficiently large, the expected overall performance increases if the probability that the manager's bias is $\hat{\alpha} = 1$ increases.

Recall that $\mathcal{G}_{G,G}$ is the new manager's bias distribution when that of the current manager is G. In this special case, Proposition 2 states that the promotion is more likely to increase the manager's bias if the subordinates are more biased.

The intuition behind this proposition is as follows. Assuming that communication efforts are complements, when the manager exerts less effort on communication, less communication is more productive for subordinates, and vice versa. Subsequently, in biased groups, biased players can yield higher performance, which makes the selected manager's bias higher. This also improves overall performance. In such a biased group, no communication is a better option for the manager, as most subordinates exert no communication effort. As the manager's bias increases, communication diminishes, and overall performance improves.

We also demonstrate that the aforementioned tendency becomes stronger over generations. The following proposition shows that a highly biased subordinate is more likely to be promoted when the manager is highly biased.

Proposition 3. *For any distribution functions $\hat{G}, \bar{G}$, if $\hat{G} \succ_{\mathrm{sd}} \bar{G}$, $\mathcal{G}_{\hat{G},G} \succ_{\mathrm{sd}} \mathcal{G}_{\bar{G},G}$.*

Proposition 3 also implies the following corollary.

Corollary 1. *If $\mathcal{G}_{G,G} \succ_{\mathrm{sd}} G$, the stationary distribution also first order stochastically dominates G if $\hat{G}_0 = G$.*

In summary, the performance-based promotion system strengthens the manager's bias, which also improves overall performance. In this sense, a performance-based system works well, but it preserves the bias in managerial positions.

Although demonstrating the desirability of the performance-based system, we limit the variety of the bias degrees to two. If the support of G has more elements, how this system affects the manager's bias remains unclear. We numerically analyze such a case in Appendix C and show that overall performance improves when biases increase.

4.2 Competitive Organization

In this section, we examine the case for the support of bias degree distribution as a continuum. To guarantee the large variety of the bias degree, we suppose a sufficiently large number of subordinates: the limit $n \to \infty$, implying that promotion is highly competitive. We show that the bias distributions of the manager become smaller over generations in the sense of stochastic dominance.

Proposition 4. *Suppose that $v_{112} \leqslant 0$ and v is symmetric. In the limit $n \to \infty$, $\hat{G} \succ_{\mathrm{sd}} \mathcal{G}_{\hat{G},G}$.*

The intuition is as follows. Recall that the subordinate's performance, $x^{\hat{\alpha}}(\alpha)$, is single-peaked with respect to his degree of bias, and the degree of bias at which the function has the maximum value is less than that of the manager $\hat{\alpha}$ (Lemma 3, b). When the number of subordinates is infinite, there certainly exists a subordinate whose degree of bias maximizes performance. Then, the subordinate whose bias is higher than that of the manager is promoted to the position of next-generation manager with negligible probability.

We also note that the support of $\mathcal{G}_{G,G}$ can degenerate to the least possible value of bias degrees. If $\mathrm{supp}(G)$ is a closed interval, $\underset{\alpha' \in \mathrm{supp}(G)}{\arg\max}\, x^{\hat{\alpha}}(\alpha') < \hat{\alpha}$. Thus, in the limit $n \to \infty$, $\max \mathrm{supp}(\mathcal{G}_{\hat{G},G}) < \max \mathrm{supp}(\hat{G})$. This means that the largest value of the support decreases. Therefore, generation after generation, $\max \mathrm{supp}(\hat{G}_t)$ is a strictly decreasing sequence. On the contrary, as $\min \mathrm{supp}(G) = \underset{\alpha' \in \mathrm{supp}(G)}{\arg\max}\, x^{\hat{\alpha}}(\alpha')$ when $\hat{\alpha} = \min \mathrm{supp}(G)$, $\min \mathrm{supp}(\mathcal{G}_{\hat{G}_t,G}) = \min \mathrm{supp}(G)$. Subsequently, in the long run, the distribution of the manager's bias degenerates to the least possible value. The results are summarized as follows:

Corollary 2. *All assumptions of Proposition 4 hold. Suppose that the support of G is a closed interval. Consider the limit $n \to \infty$. Then, in the limit distribution $\hat{G}^*$,* $\mathrm{supp}(\mathcal{G}_{\hat{G}^*,G}) = \{\min \mathrm{supp}(G)\}$.

Suppose that all assumptions of Proposition 1 hold. Recall that by Proposition 1, the expected overall performance is single-peaked with respect to the manager's bias. Let α^* be the maximizer of the expected overall performance. Corollary 2 shows that the possible values of the manager's bias converge to the minimum value of $\mathrm{supp}(G)$. Therefore, if t is sufficiently large, $\max \mathrm{supp}(\mathcal{G}_{\hat{G}_t,G}) < \alpha^*$, and let t^* be the smallest t satisfying the inequality. After $t \geqslant t^*$, the manager with a larger bias among those with a positive measure yields greater overall performance. As $\hat{G}_t >_{\mathrm{sd}} \hat{G}_{t+1}$, generation after generation, the expected overall performance decreases.

The results have the following implications. Suppose that the size of an organization grows because of an increase in the number of subordinates. As the size of the bias diversifies, the bias degenerates in the organization; nonetheless, the performance decreases over time. This result implies that the performance-based promotion system is likely to be effective when the promotion is less competitive or the support of bias distribution is restricted, such as in either bias-free or bias-pervasive organizations.

Finally, we briefly refer to performance measurements for promotion. Although Proposition 2 suggests that performance-based promotion is effective, its drawbacks are suggested in Corollary 2. Under the performance-based promotion system in our model, the performance measurement is the overall performance of the subordinates, which includes the value of communication and the outcome of the individual task. One alternative measurement is communication effort. However, this does not work effectively because the cause of inefficiency is excessive communication, as in Corollary 2. Another possible measurement, evaluating the individual task, also does not work well; this implies that the manager's bias degree increases over generations. As the overall performance is single-peaked, it also diminishes in the long run. Under each performance measurement, the inefficiency comes from the dynamics of the managers' biases. One solution to cease such evolution is to select a manager randomly from current subordinates, independent of their performances. From this perspective, a random promotion could be more effective than performance-based systems in preventing the negative effects of eroding organizational performance.

5 Discussion

This section provides some alternative scenarios based on the main model presented to confirm the validity of our results.

5.1 Retrospective Projection and Informed Subordinates

In the main model, we assume that managers project their current information state to subordinates. However, there is another way of projecting the manager's knowledge: the manager may project their subordinate-era knowledge to their subordinate. We refer to this as a *retrospective projection*. This section considers a model in that the manager is partially biased by retrospective projection. Let I_m^S be the information state that describes whether manager m knew the information in the previous period.

Using this notion, in the modified model, the manager has belief $P^{\alpha}_{I_m, I_m^S}$, and in the belief,

$$P^{\hat{\alpha}}_{I_m, I_m^S}(I_s = 1) = (1 - \hat{\alpha})p_S + \hat{\alpha}[\beta I_m + (1 - \beta)I_m^S]$$

In this model, we assume a common β for all managers. Then, when $p_S = 0$, by replacing $e_m^{\hat{\alpha}}$ with $e_m^{\hat{\alpha}[\beta+(1-\beta)I_m^S]}$, we have a similar discussion as that on the main model.

Although the type of projection, current or retrospective, does not matter when $p_S = 0$, this difference changes the results if we suppose $p_S > 0$ and consider $n \to \infty$. The reason is that a subordinate receiving information certainly exists, and such a subordinate yields the highest performance with probability one. The following proposition is the formal statement:

Proposition 5. *Suppose that $p_S > 0$. Then, as $n \to \infty$, the distribution of α converges to G, and the probability that the subordinate with $I_s = 1$ promotes converges to* 1. *In the belief of the period $t + 1$ subordinates with $\alpha < 1$, the probability also converges to* 1.

Subsequently, under the performance-based promotion system, although the promotion system does not affect the bias distribution, the manager knew the information during her subordinate era with a probability of one. Although the latter does not have influence under current information projection, it affects the beliefs of the manager when we consider the retrospective projection. As the manager knew the information during her subordinate era, retrospective projection makes the manager believe that subordinates know the information. The manager then reduces her effort in communication. By contrast, under the random promotion system, when p_S is sufficiently small, the manager is less likely to know the subordinate-era information. The retrospective projection makes the manager believe that subordinates do not know the information currently. The manager then increases her effort in communication.

This implies that the performance-based promotion system reduces the manager's communication, which potentially deteriorates overall performance. To verify this formally, we compare the per-capita performance under the performance-based promotion system and that under the random promotion system (i.e., select a subordinate as the next-generation manager at random). Let the former be $\Pi_{PB}^{p_S,n}$, and the later be $\Pi_{rand}^{p_S,n}$. We also consider the limit $p_S \to 0$.

Proposition 6. *Suppose that v_2 and u' are concave, $p_S > 0$, and F is a normal distribution. Then, if* $\sup \operatorname{supp}(G)$ *is sufficiently small, and $\beta \approx 1$, $\lim_{p_S \to 0} \lim_{n \to \infty} \Pi_{rand}^{p_S,n} > \lim_{p_S \to 0} \lim_{n \to \infty} \Pi_{PB}^{p_S,n}$.*

This proposition shows the disadvantage of the performance-based promotion system. On the one hand, similar to the result from Corollary 2, the performance difference between the manager and the subordinate enlarges because the bias diversity makes a subordinate, who is less able on average, more likely to be promoted. This decreases the performance of the organization over time. On the other hand, unlike the result from Corollary 2, the lack of communication decreases the overall performance under the retrospective projection. When there is a subordinate who has information, this subordinate is more likely to be promoted, and this reduces communication. Note that the lack of communication does not affect the bias distribution.

5.2 Overconfidence

Although we focus on information projection bias about information states, we may be able to address our issues by applying different types of psychological bias.

One plausible bias is overconfidence, as referred to in Section 1. We have several possibilities for modeling overconfidence. One is the overconfidence in the players' knowledge. Thus, each player behaves as if he knows the information, even if he does not. Formally, each player who does not know the information maximizes the following.

$$u(e_s) + \beta V + (1-\beta)v(e_s, e_m) \qquad \text{(for uninformed subordinate)}$$

$$u(e_m) + \beta v(e_s, e_m) + (1-\beta)\cdot 0 + v(e_s) \qquad \text{(for uninformed manager)}$$

Here, β is the degree of the bias. In this model, the bias reduces the communication effort of uninformed subordinates and increases the communication effort of an uninformed manager; nevertheless, it does not affect the informed players. In the model of projection bias, the bias reduces the communication effort of the informed manager, and this is the difference.

Another possibility of modeling overconfidence is underestimating others' probabilities of obtaining the information. Similar to the above case, although the bias reduces the uninformed subordinate's communication effort, it increases the informed manager's communication effort because the manager's communication is effective when the subordinate does not know the information. When the manager underestimates the probability that the subordinate knows the information, the manager believes that more communication is required compared with the case without underestimation.

In these cases of overconfidence, the evolution of the bias distribution is different from our main model. When the manager's bias is strong, the subordinate whose bias is weak is more likely to be promoted since the manager exerts more effort in communication. This creates a cycle of communication effort for managers, which contrasts with the cases of Propositions 2 and 4.

References

Battiston, Diego, Jordi Blanes i Vidal, and Tom Kirchmaier (2021) "Face-to-Face Communication in Organizations," *Review of Economic Studies*, 88 (2), 574–609.

Bernardo, Antonio E. and Ivo Welch (2001) "On the Evolution of Overconfidence and Entrepreneurs," *Journal of Economics & Management Strategy*, 10 (3), 301–330.

Bolton, Patric and Mathias Dewatripont (1994) "The Firm as a Communication Network," *Quarterly Journal of Economics*, 109(4), 809–839.

Bushong, Benjamin and Tristan Gagnon-Bartsch, forthcoming, "Failures in Forecasting: An Experiment on Interpersonal Projection Bias," *Management Science*.
"Strategic Information Transmission,"*Econometrica*, 50 (6), 1431 – 1451.

Danz, David, Madarász, Kristóf, and Stephanie W. Wang (2018) "The Biases of Others: Anticipating Informational Projection in an Agency Setting" mimeo LSE and U of Pittsburgh.

David, Herbert A., and Haikady N. Nagaraja (2003) *Order Statistics*, Third Edition, John Wiley & Sons, inc, New Jersey.

Dessein, Wouter, Andrea Galeotti, and Tano Santos (2016) “Rational Inattention and Organizational Focus,” *American Economic Review*, 106(6), 1522–1536.

Dessein, Wouter, and Tano Santos (2006) “Adaptive Organizations,” *Journal of Political Economy*, 114 (5), 956 – 995.

Dewatripont, Mathias and Jean Tirole (2005) “Modes of Communication,” *Journal of Political Economy*, 113(6), 1217–1238.

Gagnon-Bartsch, Tristan (2016) “Taste Projection in Models of Social Learning,” *Working paper*

Gagnon-Bartsch, Tristan and Antonio Rosato (2022) “Quality is in the Eye of the Beholder: Taste Projection in Markets with Observational Learning,” *Working paper*.

Gagnon-Bartsch, Tristan, Marco Pagnozzi, and Antonio Rosato (2021) “Projection of Private Values in Auctions,” *American Economic Review*, 111(10), 3256–3298.

Garicano, Luis (2000) “Hierarchies and the Organization of Knowledge in Production,” *Journal of Political Economy*, 108(5), 874–904.

Goel, Anand M. and Anjan V. Thakor (2008) “Overconfidence, CEO Selection and Corporate Governance," *Journal of Finance* LXIII(6) 2737–2784.

Johnson, Dominic D. P. and James H. Fowler (2011) “The evolution of overconfidence," *Nature*, 477, 317–320.

Heller, Yuval (2015) “Overconfidence and Diversification," *American Economic Journal: Microeconomics*, 134–153.

Heller, Yuval, and Ilan Nehama (2023) “Evolutionary foundation for heterogeneity in risk aversion,” *Journal of Economic Theory*, 208.

Lyons, Elizabeth (2017) “Team Production in International Labor Markets: Experimental Evidence from the Field,” *American Economic Journal: Applied Economics*, 9(3), 70–104.

Madarász, Kristóf (2012) "Information Projection: Model and Applications," *Review of Economic Studies*, 79(3), 961–985.

Madarász, Kristóf (2016) "Projection Equilibrium: Definition and Applications to Social Investment, Communication and Trade," CEPR D.P.

Madarász, Kristóf (2021) "Bargaining under the Illusion of Transparency," *American Economic Review*, 111(11), 3500–3539.

Madarász, Kristóf, David Danz, and Stephanie W. Wang (2023) "Projective Thinking: Model, Evidence, and Applications," mimeo.

Shaked, Moshe, and J. George Shanthikumar (2007) *Stochastic Orders*, Springer.

Whitt, Ward (1985) "Uniform Conditional Variability Ordering of Probability Distributions," *Journal of Applied Probability*, 22(3), 619–633.

Appendix A Bias Distribution of Promoted Subordinates

In this section, we formally derive Γ, the distribution function of the subordinate's bias that yields the highest performance among his colleagues.

To see this, we calculate the subordinates' performance when the subordinate knows the information is $V + u(1)$. The performance when the subordinate does not have the information depends on the manager's decision and projection bias. Given that the manager has projection bias $\hat{\alpha}$, the performance of the subordinate with projection bias α is $\theta + x^{\hat{\alpha}}(\alpha)$.

Given π', let g^* be the distribution of α (pdf) that yields performance π_i, which is calculated as follows:

$$g^*(\alpha \mid \hat{\alpha}, \pi') = g(\alpha) \frac{f\left(\pi' - x^{\hat{\alpha}}(\alpha)\right)}{\int f\left(\pi' - x^{\hat{\alpha}}(\alpha')\right) dG(\alpha')}.$$

The term $f\left(\pi' - x^{\hat{\alpha}}(\alpha)\right)$ is the probability density function that the performance is π' conditioned on the information states of the manager and the subordinate.

Define $G(\alpha \mid \hat{\alpha}, \pi) = \int^{\alpha} g^*(\alpha' \mid \hat{\alpha}, \pi) d\alpha'$. Moreover, let $H(\pi \mid \hat{\alpha}, G)$ be the probability that the subordinate's performance is less than π given that the manager's bias is $\hat{\alpha}$, which is represented by

$$H(\pi \mid \hat{\alpha}, G) = \int^{\pi} d\pi' \left[\int f\left(\pi' - x^{\hat{\alpha}}(\alpha')\right) dG(\alpha') \right].$$

Then, let $\Gamma(\alpha \mid \hat{\alpha}, G)$ denote the distribution of α of the promoted subordinate who yields the greatest performance among n subordinates.

$\Gamma(\alpha \mid \hat{\alpha}, G)$ is calculated as follows:

$$\Gamma(\alpha \mid \hat{\alpha}, G) := \int G(\alpha \mid \hat{\alpha}, \pi) \underbrace{nH'(\pi \mid \hat{\alpha}, G)[H(\pi \mid \hat{\alpha}, G)]^{n-1}}_{\text{density that the highest performance is } \pi} d\pi.$$

Appendix B Proofs

Appendix B.1 Proofs in Section 3

Proof of Lemma 1. Consider (1) which determines the subordinate's optimal effort level. By Assumption 3, for each e_m, there uniquely exists $\varphi(e_m) \in [0, 1]$ such that

$$P^{\alpha}_{I_s}(I_m = 1) v_1(\varphi(e_m), e_m) + u'(\varphi(e_m)) = 0.$$

Similarly, for each e_s, there also uniquely exists $\psi(e_s) \in [0, 1]$ such that

$$u'(\psi(e_s)) + (1 - p_S) v_2(e_s, \psi(e_s)) = 0.$$

The solution to (1) satisfies $\varphi(e_m) = e_s$ and $\psi(e_s) = e_m$. By $v_{11} < 0, v_{22} < 0$ and $u'' < 0$, the implicit function theorem implies that ψ and φ are continuous, which implies the existence of the solution.

We now prove the uniqueness. Assuming the contrary, we have two solutions, (e_s, e_m) and (e'_s, e'_m), which satisfy

$$P^{\alpha}_{I_s}(I^M_m = 1)v_1(e_s, e_m) + u'(e_s) = P^{\alpha}_{I_s}(I^M_m = 1)v_1(e'_s, e'_m) + u'(e'_s),$$
$$u'(e_m) + (1 - p_S)v_2(e_s, e_m) = u'(e'_m) + (1 - p_S)v_2(e'_s, e'_m).$$

Via Taylor's expansion,

$$\begin{pmatrix} P^{\alpha}_{I_s}(I^M_m = 1)v_{11} + u''(e_s) & v_{12} \\ (1 - p_S)v_{21} & u''(e_m) + (1 - p_S)v_{22} \end{pmatrix} \begin{pmatrix} e_s - e'_s \\ e_m - e'_m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

By Assumption 1, the left matrix has a positive determinant, which implies $e_s - e'_s = 0$ and $e_m - e'_m = 0$. This shows the uniqueness of the equilibrium.

Note that the solution is not a corner one if $\alpha \in (0, 1)$. By $v_1 < 0$, $v_2 < 0$ and $u'(1) = 0$, $(e_s, e_m) = (1, 1)$ does not satisfy (1). Similarly, by assumption 3, $(e_s, e_m) = (0, 0)$ also does not satisfy (1). A similar proof is applicable to (2), which determines the informed manager's optimal effort level. □

Proof of Lemma 2. **Subordinate:** Taking the total differentiation for the system of equations (1), we have

$$\frac{\partial e^{\alpha}_s}{\partial \alpha} = \frac{p_M v_1(u''(e_m) + (1 - p_S)v_{22})}{\det(D_S)}$$

where

$$D_S = \begin{pmatrix} (1 - \alpha)p_M v_{11}(e_s, e_m) + u''(e_s) & (1 - \alpha)p_M v_{12}(e_s, e_m) \\ (1 - p_S)v_{21}(e_s, e_m) & u''(e_m) + (1 - p_S)v_{22}(e_s, e_m). \end{pmatrix}$$

By Assumption 1 (iv), $\det(D_S) > 0$; moreover, the numerator of $\frac{\partial e^{\alpha}_s}{\partial \alpha}$ is also positive.

Manager: Taking the total differentiation for the system of equations (2),

$$\frac{\partial e_m^\alpha}{\partial \alpha} = \frac{(1-p_S)v_2[p_M v_{11} + u''(e_s)]}{\det(D_M)},$$

where

$$D_M = \begin{pmatrix} p_M v_{11}(e_s, e_m) + u''(e_s) & p_M v_{12}(e_s, e_m) \\ (1-\alpha)(1-p_S)v_{21}(e_s, e_m) & u''(e_m) + (1-\alpha)(1-p_S)v_{22}(e_s, e_m) \end{pmatrix}$$

By Assumption 1 (iv), we can also show that $\det(D_M) > 0$. Then, $\frac{\partial e_m^\alpha}{\partial \alpha}$ is positive.

□

Proof of Lemma 3. (a) Note that $u(e) + v(e, e_m)$ is concave in e and decreasing in e_m. By Lemma 2, $e_m^{\hat{\alpha}}$ and e_s^α are increasing in α, and $x^{\hat{\alpha}}(\alpha)$ is single-peaked with respect to α and decreasing in $\hat{\alpha}$.[19]

(b) Now, we consider the maximizer, $\arg\max_\alpha x^{\hat{\alpha}}(\alpha)$. Let $e_s^*(\hat{\alpha})$ be the effort level of the subordinate that the manager with degree $\hat{\alpha}$ believes.

Claim 1. *If v is symmetric, $e_s^{\hat{\alpha}} \geqslant e_s^*(\hat{\alpha})$.*

Proof of Claim 1. Note that $(e_s^*(\hat{\alpha}), e_m^{\hat{\alpha}})$ is the solution to

$$v_1(e_s, e_m) + u'(e_s) = 0,$$
$$u'(e_m) + (1-\hat{\alpha})v_2(e_s, e_m) = 0.$$

By the symmetry of v, $v_2(e_s, e_m) = v_1(e_m, e_s)$. Then, if $e_s^*(\hat{\alpha}) > e_m^{\hat{\alpha}}$, by the assumptions that $v_1 \leqslant 0$, $v_{12} > 0$, and $v_{11} < 0$,

$$v_1(e_s^*(\hat{\alpha}), e_m^{\hat{\alpha}}) = v_2(e_m^{\hat{\alpha}}, e_s^*(\hat{\alpha})) < v_2(e_s^*(\hat{\alpha}), e_m^{\hat{\alpha}}) \leqslant (1-\hat{\alpha})v_2(e_s^*(\hat{\alpha}), e_m^{\hat{\alpha}}).$$

[19] Note that the concave function is single-peaked, and any monotonic transformation of a single-peaked function is single-peaked.

Note also that by $u'' < 0$, $u'(e_m^*(\hat{\alpha})) > u(e_s^{\hat{\alpha}})$. Then, $0 = u'(e_m^*(\hat{\alpha})) + v_1(e_s^*(\hat{\alpha}), e_m^{\hat{\alpha}}) < (1 - \hat{\alpha})v_2(e_s^*(\hat{\alpha}), e_m^{\hat{\alpha}}) + u'(e_m^{\hat{\alpha}}) = 0$, which is a contradiction. Therefore, $e_m^{\hat{\alpha}} \geqslant e_s^*(\hat{\alpha})$.

On the other hand, $e_s^{\hat{\alpha}}$ is the solution to

$$(1 - \hat{\alpha})v_1(e_s, e_m) + u'(e_s) = 0,$$
$$u'(e_m) + v_2(e_s, e_m) = 0,$$

By the symmetry of v, $v_2(e_s, e_m) = v_1(e_m, e_s)$, and therefore, $e_s^{\hat{\alpha}} = e_m^{\hat{\alpha}}$. Then, we conclude that $e_s^{\hat{\alpha}} \geqslant e_s^*(\hat{\alpha})$. □

By Claim 1, we prove that the maximizer of $x^{\hat{\alpha}}(\alpha)$ is less than $\hat{\alpha}$. The derivative of $x^{\hat{\alpha}}(\alpha)$ with respect to α is calculated as

$$\frac{\partial x^{\hat{\alpha}}(\alpha)}{\partial \alpha} = \frac{\partial e_s^{\alpha}}{\partial \alpha}\left[u'(e_s^{\alpha}) + v_1(e_s^{\alpha}, e_m^{\hat{\alpha}})\right].$$

Suppose that $\alpha > \hat{\alpha}$. Then, by $u'' < 0$, $v_{11} < 0$, $v_1 < 0$, $e_s^*(\hat{\alpha}) < e_s^{\alpha}$, and monotonicity of $e_s^*(\alpha)$,

$$\begin{aligned} u'(e_s^{\alpha}) + v_1(e_s^{\alpha}, e_m^{\hat{\alpha}}) &\leqslant u'(e_s^*(\alpha)) + v_1(e_s^*(\alpha), e_m^{\hat{\alpha}}) \\ &< u'(e_s^*(\hat{\alpha})) + v_1(e_s^*(\hat{\alpha}), e_m^{\hat{\alpha}}) \\ &< u'(e_s^*(\hat{\alpha})) + (1 - \hat{\alpha})v_1(e_s^*(\hat{\alpha}), e_m^{\hat{\alpha}}) = 0, \end{aligned}$$

which implies that $\frac{\partial x^{\hat{\alpha}}(\alpha)}{\partial \alpha} < 0$. Then, $x^{\hat{\alpha}}(\alpha)$ has the decreasing region in $(\hat{\alpha}, 1)$, and thus the maximizer is less than $\hat{\alpha}$.

(c) As for the comparative statics of the maximizer $\alpha^* = \arg\max_{\alpha} x^{\hat{\alpha}}(\alpha)$, since the maximizer α^* satisfies $\frac{\partial x^{\hat{\alpha}}(\alpha^*)}{\partial \alpha} = 0$, and thus, $v_1(e_s^{\alpha^*}, e_m^{\hat{\alpha}}) - (1 - \alpha^*)v_1(e_s^{\alpha^*}, e_m^*(\alpha^*)) = 0$. Differentiating $v_1(e_s^{\alpha}, e_m^{\hat{\alpha}}) - (1 - \alpha)v_1(e_s^{\alpha}, e_m^*(\alpha))$ by α yields that

$$\left[v_{11}(e_s^{\alpha}, e_m^{\hat{\alpha}}) - (1 - \alpha)v_{11}(e_s^{\alpha}, e_m^*(\alpha))\right]\frac{\partial e_s^{\alpha}}{\partial \alpha} + v_1(e_s^{\alpha}, e_m^*(\alpha)) - (1 - \alpha)v_{12}(e_s^{\alpha}, e_m^*(\alpha))\frac{\partial e_m^{\alpha}}{\partial \alpha}.$$

If $v_{112} \leqslant 0$, as the maximizer α^* satisfies $\alpha^* < \hat{\alpha}$, the first term is negative. Further, by $v_{11} < 0$ and $v_{12} \geqslant 0$, the remaining terms are all negative. The total derivative implies that the maximizer is increasing in $\hat{\alpha}$ as $v_{12} > 0$.

(d) Finally, as for the cross derivative,

$$\frac{\partial^2 x^{\hat{\alpha}}(\alpha)}{\partial\alpha\partial\hat{\alpha}} = \frac{\partial e_s^\alpha}{\partial\alpha}\frac{\partial e_m^{\hat{\alpha}}}{\partial\hat{\alpha}} v_{12}(e_s^\alpha, e_m^{\hat{\alpha}}) > 0.$$

□

Proof of Proposition 1. Note that $E[\Pi \mid \hat{\alpha}]$ is concave with respect to $e_m^{\hat{\alpha}}$. Since $e_m^{\hat{\alpha}}$ is increasing in $\hat{\alpha}$, $E[\Pi \mid \hat{\alpha}]$ is single-peaked with respect to $\hat{\alpha}$.

To find the peak of $E[\Pi \mid \hat{\alpha}]$, we differentiate the expected overall performance with respect to e_m, which yields

$$\begin{aligned}\frac{\partial E[\Pi \mid \hat{\alpha}]}{\partial e_m^{\hat{\alpha}}} &= u'(e_m^{\hat{\alpha}}) + \int dG(\alpha) v_2(e_s^\alpha, e_m^{\hat{\alpha}}) \\ &= \int dG(\alpha) v_2(e_s^\alpha, e_m^{\hat{\alpha}}) - v_2(e_s^*(\hat{\alpha}), e_m^{\hat{\alpha}})\end{aligned}$$

where $e_s^*(\hat{\alpha})$ is the effort level of a subordinate, which the manager with degree $\hat{\alpha}$ believes. The assumption that v_{12} is constant implies that v_2 is linear with respect to the first argument,

$$\frac{\partial E[\Pi \mid \hat{\alpha}]}{\partial e_m^{\hat{\alpha}}} = v_2(\textstyle\int dG(\alpha) e_s^\alpha, e_m^{\hat{\alpha}}) - v_2(e_s^*(\hat{\alpha}), e_m^{\hat{\alpha}}).$$

Then, since $\frac{\partial e_m^{\hat{\alpha}}}{\partial\hat{\alpha}} > 0$ by Lemma 2, $e_s^{\hat{\alpha}} > \int dG(\alpha) e_s^\alpha$ if and only if $\frac{\partial E[\Pi|\hat{\alpha}]}{\partial\hat{\alpha}} < 0$. As $e_s^*(\alpha)$ is increasing in α, this completes the proof. □

Appendix B.2 Proofs in Section 4

Proof of Proposition 3. The proof follows from the following lemma.

Lemma 4. *For any $\hat{\alpha}, \hat{\alpha}'$ with $\hat{\alpha} > \hat{\alpha}'$, $\Gamma(\cdot \mid \hat{\alpha}, G)$ first order stochastically dominates $\Gamma(\cdot \mid \hat{\alpha}', G)$.*

Proof of Lemma 4. As discussed in Appendix A, we have

$$\int G(\tilde{\alpha} \mid \hat{\alpha}, \pi) H'(\pi \mid \hat{\alpha}, G) d\pi = \int^{\tilde{\alpha}} g(\alpha) \int f(\pi - x^{\hat{\alpha}}(\alpha)) d\pi d\alpha.$$

By differentiating $\Gamma(\alpha \mid \hat{\alpha}, G)$ with $\hat{\alpha}$,

$$\begin{aligned} \frac{\partial \Gamma(\tilde{\alpha} \mid \hat{\alpha}, G)}{\partial \hat{\alpha}} = \int^{\tilde{\alpha}} g(\alpha) \int_{-\infty}^{\infty} \Big[&- \frac{\partial x^{\hat{\alpha}}(\alpha)}{\partial \hat{\alpha}} f'(\pi - x^{\hat{\alpha}}(\alpha)) H(\pi \mid \hat{\alpha}, G) \\ &+ (n-1) f(\pi - x^{\hat{\alpha}}(\alpha)) H_{\hat{\alpha}}(\pi \mid \hat{\alpha}, G) \Big] (H(\pi \mid \hat{\alpha}, G))^{n-2} n d\pi d\alpha \end{aligned}$$

where $H_{\hat{\alpha}}$ is a derivative of H with respect to $\hat{\alpha}$;

$$\begin{aligned} H_{\hat{\alpha}}(\pi \mid \hat{\alpha}, G) &= - \int \frac{\partial x^{\hat{\alpha}}(\alpha)}{\partial \hat{\alpha}} \int^{\pi} f'(\pi' - x^{\hat{\alpha}}(\alpha)) d\alpha d\pi' \\ &= - \int \frac{\partial x^{\hat{\alpha}}(\alpha)}{\partial \hat{\alpha}} f(\pi - x^{\hat{\alpha}}(\alpha)) d\alpha. \end{aligned}$$

Note that integration by parts yields

$$\begin{aligned} &\int_0^{\tilde{\alpha}} d\alpha g(\alpha) \int_{-\infty}^{\infty} d\pi \frac{\partial x^{\hat{\alpha}}(\alpha)}{\partial \hat{\alpha}} [f'(\pi - x^{\hat{\alpha}}(\alpha)) H(\pi \mid \hat{\alpha}, G) \\ &\quad +(n-1) f(\pi - x^{\hat{\alpha}}(\alpha)) H'(\pi \mid \hat{\alpha}, G)] (H(\pi \mid \hat{\alpha}, G))^{n-2} n \\ &= \int_0^{\tilde{\alpha}} g(\alpha) \left[\frac{\partial x^{\hat{\alpha}}(\alpha)}{\partial \hat{\alpha}} f(\pi - x^{\hat{\alpha}}) (H(\pi \mid \hat{\alpha}, G))^{n-1} n \right]_{-\infty}^{\infty} d\alpha = 0. \end{aligned} \tag{3}$$

Adding (3) to $\frac{\partial \Gamma(\alpha \mid \hat{\alpha}, G)}{\partial \hat{\alpha}}$ yields

$$\begin{aligned} \frac{\partial \Gamma(\tilde{\alpha} \mid \hat{\alpha}, G)}{\partial \hat{\alpha}} = \int^{\tilde{\alpha}} d\alpha g(\alpha) \int_{-\infty}^{\infty} d\pi [&b_{\pi, \hat{\alpha}}(\alpha) H_{\hat{\alpha}}(\pi \mid \hat{\alpha}, G) \\ &+ \frac{\partial x^{\hat{\alpha}}(\alpha)}{\partial \hat{\alpha}} f(\pi - x^{\hat{\alpha}}(\alpha)) H'(\pi \mid \hat{\alpha}, G)] \\ &\times (H(\pi \mid \hat{\alpha}, G))^{n-2} n(n-1). \end{aligned}$$

Note that because Γ is a cumulative density function on $[0,1]$, $\Gamma(1 \mid \hat{\alpha}, G) = 1$ and $\Gamma(0 \mid \hat{\alpha}, G) = 0$, and then $\frac{\partial \Gamma(1 \mid \hat{\alpha}, G)}{\partial \hat{\alpha}} = \frac{\partial \Gamma(0 \mid \hat{\alpha}, G)}{\partial \hat{\alpha}} = 0$. Moreover,

$$
\begin{aligned}
\frac{\partial \Gamma(\tilde{\alpha} \mid \hat{\alpha}, G)}{\partial \hat{\alpha}} &= \int^{\tilde{\alpha}} (d\alpha) g(\alpha) \int (d\alpha') g(\alpha') \int_{-\infty}^{\infty} (d\pi) f(\pi - x^{\hat{\alpha}}(\alpha)) f(\pi - x^{\hat{\alpha}}(\alpha')) \\
&\quad \times \left[\frac{\partial x^{\hat{\alpha}}(\alpha)}{\partial \hat{\alpha}} - \frac{\partial x^{\hat{\alpha}}(\alpha')}{\partial \hat{\alpha}} \right] (H(\pi \mid \hat{\alpha}, G))^{n-2} n(n-1), \\
&= \begin{cases} \displaystyle\int_{-\infty}^{\infty} (d\pi) W(1 \mid \pi) W(\tilde{\alpha} \mid \pi) \\ \displaystyle \quad \times \left[\int (d\alpha) w(\alpha \mid \pi, \tilde{\alpha}) \frac{\partial x^{\hat{\alpha}}(\alpha)}{\partial \hat{\alpha}} - \int (d\alpha') w(\alpha' \mid \pi) \frac{\partial x^{\hat{\alpha}}(\alpha')}{\partial \hat{\alpha}} \right] \\ \quad \times (H(\pi \mid \hat{\alpha}, G))^{n-2} n(n-1) \end{cases}
\end{aligned} \tag{4}
$$

where

$$
w(\hat{\alpha} \mid \pi) = \frac{g(\alpha') f(\pi - x^{\hat{\alpha}}(\alpha'))}{W(1 \mid \pi)}, W(\tilde{\alpha} \mid \pi) = \int^{\tilde{\alpha}} d(\alpha') g(\alpha') f(\pi - x^{\hat{\alpha}}(\alpha')),
$$

$$
\text{and } w(\alpha' \mid \pi, \tilde{\alpha}) = \begin{cases} \frac{g(\alpha') f(\pi - x^{\hat{\alpha}}(\alpha'))}{W(\tilde{\alpha} | \pi)} & \text{if} \quad \alpha' < \tilde{\alpha} \\ 0 & \text{if} \quad \alpha' \geqslant \tilde{\alpha}. \end{cases}
$$

Note that both $w(\hat{\alpha} \mid \pi)$ and $w(\hat{\alpha} \mid \pi, \tilde{\alpha})$ are probability density functions with respect to α' and $w(\hat{\alpha} \mid \pi)$ first order stochastically dominates $w(\hat{\alpha} \mid \pi, \tilde{\alpha})$ as they are single-crossing (see, e.g., Shaked and Shanthikumar, 2007, Theorem 1.A.12). Note also that

$$
\frac{\partial x^{\hat{\alpha}}(\alpha)}{\partial \hat{\alpha}} = v_2 \frac{\partial e_m^{\hat{\alpha}}}{\partial \hat{\alpha}}. \tag{5}
$$

By (5), $\frac{\partial x^{\hat{\alpha}}(\alpha)}{\partial \hat{\alpha}}$ is increasing in α as e_s^{α} is increasing in α, and $\frac{\partial e_m^{\hat{\alpha}}}{\partial \hat{\alpha}} > 0$ by Lemma 2. Then, since $w(\hat{\alpha} \mid \pi)$ first order stochastically dominates $w(\hat{\alpha} \mid \pi, \tilde{\alpha})$, the second term of (4) is negative for each π. Therefore, $\Gamma(\tilde{\alpha} \mid \hat{\alpha}', G) > \Gamma(\tilde{\alpha} \mid \hat{\alpha}, G)$ for each $\tilde{\alpha}$, $\hat{\alpha}$, and $\hat{\alpha}'$ with $\hat{\alpha} > \hat{\alpha}'$. This implies the first-order stochastic dominance with higher $\hat{\alpha}$. □

By Lemma 4, $\Gamma(\tilde{\alpha} \mid \hat{\alpha}, G)$ is decreasing in $\hat{\alpha}$. Therefore, if $\hat{G} \gtrsim_{\text{sd}} \tilde{G}$, $\mathcal{G}_{\hat{G},G}(\tilde{\alpha}) = \int d\hat{G}(\hat{\alpha}) \Gamma(\tilde{\alpha} \mid \hat{\alpha}, G) \leqslant \mathcal{G}_{\tilde{G},G}(\tilde{\alpha})$, which completes the proof. □

Proof of Proposition 2. (a) First, we show that $x^0(0) > x^0(1)$ and $x^1(0) < x^1(1)$ under the conditions stated in Proposition 2.

Claim 2. *If* $v(1, e_m) = v(e_s, 1) = 0$ *for any* e_s *and* e_m, $x^0(0) > x^0(1)$ *and* $x^1(0) < x^1(1)$.

Proof of Claim 2. **The case for** $\hat{\alpha} = 0$**:** By Lemma 3 (b), if $\hat{\alpha} = 0$, $x^0(0) > x^0(1)$ since the maximizer of $x^{\hat{\alpha}}(\alpha)$ is less than $\hat{\alpha}$

The case for $\hat{\alpha} = 1$**:** If $\alpha = 1$, the first order condition of the subordinate is $u'(e_s) = 0$, which implies that $e_s^1 = 1$. On the contrary, if $\alpha = 0$, $e_s^0 < 1$.

As for the manager with $\hat{\alpha} = 1$, the optimal effort is $e_m^1 = 1$. Then, $x^1(\alpha) = u(e_s^\alpha)$, which implies that $x^1(0) < x^1(1)$. □

Let $\gamma_{G,G}(\alpha)$ be the probability that the selected subordinate's bias is α. Then,

$$\gamma_{G,G}(\alpha) = g(\alpha) \sum_{\hat{\alpha}} g(\hat{\alpha}) \int d\pi f(\pi - x^{\hat{\alpha}}(\alpha)) n \left[\sum_{\alpha'} g(\alpha') F(\pi - x^{\hat{\alpha}}(\alpha')) \right]^{n-1}.$$

Let define $\tilde{\gamma}(\alpha \mid \alpha', \hat{\alpha})$ as

$$\tilde{\gamma}(\alpha \mid \alpha', \hat{\alpha}) = \int d\pi f(\pi - x^{\hat{\alpha}}(\alpha)) n [F(\pi - x^{\hat{\alpha}}(\alpha'))]^{n-1}.$$

Note that by the convexity of exponential $n-1$, $\gamma_{G,G}(\alpha) \leqslant g(\alpha) \sum_{\alpha',\hat{\alpha}} g(\hat{\alpha}) g(\alpha') \tilde{\gamma}(\alpha \mid \alpha', \hat{\alpha})$.

Suppose that f is the normal distribution with mean μ and variance σ^2. Then,

$$\tilde{\gamma}(\alpha \mid \alpha', \hat{\alpha}) = \int e^{-\frac{x\Delta}{\sigma} - \frac{\Delta^2}{2\sigma^2}} n\varphi(x)[\Phi(x)]^{n-1} dx$$

where $\Delta = x^{\hat{\alpha}}(\alpha') - x^{\hat{\alpha}}(\alpha)$, $x = \frac{\pi - x^{\hat{\alpha}}(\alpha) - \mu}{\sigma}$, and φ and Φ are pdf and cdf of the standard normal distribution respectively.

By considering $\bar{u}(x) = e^{-\frac{x\Delta}{\sigma} - \frac{\Delta^2}{2\sigma^2}}$, $\tilde{\gamma}(\alpha \mid \alpha', \hat{\alpha})$ is the expectation of $\bar{u}(x)$ when x is followed by distribution Φ^n. Note that $\bar{u}'(x) > 0$ if and only if $\Delta < 0$. Note also that $\int e^{-\frac{x\Delta}{\sigma} - \frac{\Delta^2}{2\sigma^2}} \varphi(x) dx = 1$ since $\int e^{-\frac{x\Delta}{\sigma}} \varphi(x) dx$ is the moment-generating function of the standard normal distribution, whose value is $\int e^{-\frac{x\Delta}{\sigma}} \varphi(x) dx = e^{\frac{\Delta^2}{2\sigma^2}}$. Then, since $[\Phi]^n \succ_{\text{sd}} \Phi$, $\tilde{\gamma}(\alpha \mid \alpha', \hat{\alpha}) > 1$ if and only if $\Delta < 0$.

Now we consider the value of $\gamma_{G,G}(0)$. Note that $\tilde{\gamma}(0 \mid 0, 1) = \tilde{\gamma}(1 \mid 0, 0) = 1$. Then,

$$\gamma_{G,G}(0) \leqslant [g(1)[g(0) + g(1)\tilde{\gamma}(0 \mid 1, 1)] + g(0)[g(0) + g(1)\tilde{\gamma}(0 \mid 1, 0)]]g(0),$$

which implies that if $g(1)\tilde{\gamma}(0 \mid 1, 1) + g(0)\tilde{\gamma}(0 \mid 1, 0) < 1$, $\gamma_{G,G}(0) < g(0)$. As $x^1(0) - x^1(1) < 0$ and $x^0(0) - x^0(1) > 0$, $\tilde{\gamma}(0 \mid 1, 1) < 1 < \tilde{\gamma}(0 \mid 1, 0)$. Then, if $g(1) > \frac{\tilde{\gamma}(0|1,0)-1}{\tilde{\gamma}(0|1,0)-\tilde{\gamma}(0|1,1)}$, $\gamma_{G,G}(0) < g(0)$ holds. This implies that $\mathcal{G}_{G,G} \succ_{\text{sd}} G$.

(b) Note that if $\frac{\partial\tilde{\gamma}(0|1,1)}{\partial\sigma} > 0 > \frac{\partial\tilde{\gamma}(0|1,0)}{\partial\sigma}$, $\frac{\tilde{\gamma}(0|1,0)-1}{\tilde{\gamma}(0|1,0)-\tilde{\gamma}(0|1,1)}$ is increasing in σ. To show this, we prepare the following claim.

Claim 3. *$\tilde{\gamma}(\alpha \mid \alpha', \hat{\alpha})$ decreases if Δ/σ increases.*

By this claim, $\tilde{\gamma}(\alpha \mid \alpha', \hat{\alpha})$ is decreasing in σ if and only if $\Delta < 0$, and this completes the proof.

Proof of Claim 3. Recall that

$$\tilde{\gamma}(\alpha \mid \alpha', \hat{\alpha}) = \int e^{-\frac{x\Delta}{\sigma} - \frac{\Delta^2}{2\sigma^2}} n\varphi(x)[\Phi(x)]^{n-1}dx.$$

Let $r = \Delta/\sigma$ and define

$$Q(r) = e^{-r^2/2} \int e^{-rx} n\varphi(x)[\Phi(x)]^{n-1}dx.$$

Then,

$$Q'(r) = -e^{-r^2/2} \int (r + x)e^{-rx} n\varphi(x)[\Phi(x)]^{n-1}dx.$$

Since φ is the standard normal distribution, we can show that $\int(r + x)e^{-rx} n\varphi(x)dx = 0$. Note that the integrand is positive if and only if $x > -r$. Then,

$$\begin{aligned}\int (r+x)e^{-rx} n\varphi(x)[\Phi(x)]^{n-1}dx &= \int^{-r} (r+x)e^{-rx} n\varphi(x)[\Phi(x)]^{n-1}dx + \int_{-r} (r+x)e^{-rx} n\varphi(x)[\Phi(x)]^{n-1}dx \\ &\geqslant \int^{-r} (r+x)e^{-rx} n\varphi(x)[\Phi(r)]^{n-1}dx + \int_{-r} (r+x)e^{-rx} n\varphi(x)[\Phi(r)]^{n-1}dx \\ &= [\Phi(r)]^{n-1} \int (r+x)e^{-rx} n\varphi(x)dx = 0.\end{aligned}$$

Therefore, $Q'(r) < 0$. This implies that $\tilde{\gamma}(\alpha \mid \alpha', \hat{\alpha})$ decreases if Δ/σ increases. □

(c) When the manager's bias distribution is $\hat{g}$, the expected per-capita overall performance is

$$\hat{g}(1)[u(1) + g(1)x^1(1) + g(0)x^1(0)] + \hat{g}(0)[u(e_m^0) + g(1)x^0(1) + g(0)x^0(0)].$$

This is increasing in $\hat{g}(1)$ if

$$[u(1) - u(e_m^0)] + g(1)[x^1(1) - x^0(1)] + g(0)[x^0(1) - x^0(0)] > 0.$$

Note that $u(1) > u(e_m^0)$, $x^1(1) > x^0(1)$, and $x^0(1) < x^0(0)$. Then, if $g(1)$ is sufficiently large, the overall performance is increasing in $\hat{g}(1)$. □

Appendix B.3 Proofs in Section 5

Proof of Proposition 4. Let $z\colon [0,1]^2 \to \{1, 0\}$ so that $z(\alpha, \hat{\alpha}) = 1$ if $\alpha = \underset{\alpha' \in \mathrm{supp}(G)}{\arg\max}\, x^{\hat{\alpha}}(\alpha')$, and otherwise, 0. Then, we show that $\mathcal{G}_{\hat{G},G}$ is independent of G other than its support in the limit.

Lemma 5. *Consider the limit $n \to \infty$. Then, $\mathcal{G}'_{\hat{G},G}(\alpha) = \int d\hat{G}(\hat{\alpha}) z(\alpha, \hat{\alpha})$.*

Proof of Lemma 5. Note that

$$\mathcal{G}'_{\hat{G},G}(\alpha) = \int d\hat{G}(\hat{\alpha}) \int_{-\infty}^{\infty} \frac{f(\pi - x^{\hat{\alpha}}(\alpha))g(\alpha)}{\int f(\pi - x^{\hat{\alpha}}(\alpha'))dG(\alpha')} [(H(\pi \mid \hat{\alpha}, G))^n]' d\pi.$$

Note that $[(H(\pi \mid \hat{\alpha}, G))^n]' = H'(\pi \mid \hat{\alpha}, G)n(H(\pi \mid \hat{\alpha}, G))^{n-1} \to 0$ as $n \to \infty$ if $\pi < \infty$. This implies that at the limit $n \to \infty$,

$$\mathcal{G}'_{\hat{G},G}(\alpha) = \lim_{\pi\to\infty} \int d\hat{G}(\hat{\alpha}) \lim_{\pi\to\infty} \frac{f(\pi - x^{\hat{\alpha}}(\alpha))g(\alpha)}{\int f(\pi - x^{\hat{\alpha}}(\alpha'))dG(\alpha')}.$$

By Assumption 2,

$$\lim_{\pi\to\infty} \frac{f(\pi - x^{\hat{\alpha}}(\alpha))g(\alpha)}{\int f(\pi - x^{\hat{\alpha}}(\alpha'))dG(\alpha')} = \begin{cases} 1 & \text{if } \alpha = \arg\max_{\alpha'} x^{\hat{\alpha}}(\alpha') \\ 0 & \text{otherwise.} \end{cases}$$

This completes the proof. $\square$

By Lemma 3, let $\chi(\hat{\alpha}) := \arg\max_{\alpha} x^{\hat{\alpha}}(\alpha)$, which is increasing in $\hat{\alpha}$ and less than $\hat{\alpha}$. Then, $\mathcal{G}_{\hat{G},G}(\alpha) = \hat{G}(\chi^{-1}(\alpha)) \geqslant \hat{G}(\alpha)$. $\square$

Proof of Corollary 2. First note that $\mathcal{G}$ is a Markov process, and it has a unique limit distribution. We denote the limit distribution by $\hat{G}^*$.

Next, we show that $\operatorname{supp}(\mathcal{G}_{\hat{G},G})$ is an interval if $\operatorname{supp} G$ is an interval. By the continuity of u and v, each e_s and e_m is continuous with respect to α and $\hat{\alpha}$. Then, $\arg\max_{\alpha'\in\operatorname{supp}(G)} x^{\hat{\alpha}}(\alpha')$ is also continuous in $\hat{\alpha}$.

Note also that when $\hat{\alpha} = \min\operatorname{supp}(G)$,

$$\min\operatorname{supp}(G) = \arg\max_{\alpha'\in\operatorname{supp}(G)} x^{\hat{\alpha}}(\alpha').$$

Then, the intermediate theorem implies that for each $\hat{\alpha} \in \operatorname{supp}(\mathcal{G}_{\hat{G},G})$ and each $\hat{\alpha}' < \hat{\alpha}$, $\hat{\alpha}' \in \operatorname{supp}(\mathcal{G}_{\hat{G},G})$. This shows that $\operatorname{supp}(\mathcal{G}_{\hat{G},G})$ is an interval.

Since $\max\operatorname{supp}(\hat{G}_t)$ is a strictly decreasing sequence on the real number, it has a convergent. Let the convergent be $\bar{\alpha}$. Suppose by contradiction that $\bar{\alpha} > \min\operatorname{supp}(G)$. Then, $\bar{\alpha} > \arg\max_{\alpha'\in\operatorname{supp}(G)} x^{\bar{\alpha}}(\alpha')$, which implies that $\bar{\alpha} \notin \operatorname{supp}(\mathcal{G}_{\hat{G}^*,G})$. This contradicts the supposition that $\hat{G}^*$ is a limit distribution. $\square$

Proof of Proposition 5. Note that for each $\pi < \infty$, $nH'(\pi \mid \hat{\alpha}, G)[H(\pi \mid \hat{\alpha}, G)]^{n-1} \to 0$ as $n \to \infty$, because, for each $\pi < \infty$, $H(\pi \mid \hat{\alpha}, G) < 1$. Therefore,

$$\lim_{n\to\infty} \Gamma(\alpha \mid \hat{\alpha}, G) = \lim_{\pi\to\infty} G(\alpha \mid \hat{\alpha}, \pi).$$

Recall that by Assumption 1, $u(1)+V > u(e)+v(e,e')$: the subordinate who knows the information initially yields the highest performance. Note that given π, the probability that the subordinate with $I_i = 1$ is promoted is

$$\int d\alpha g(\alpha)\left[\frac{p_S f(\pi - (u(1)+V))}{p_S f(\pi - (u(1)+V)) + (1-p_S) f(\pi - x^{\hat{\alpha}}(\alpha))}\right].$$

As long as $p_S > 0$, under Assumptions 1 and 2, this probability converges to 1 as $n \to \infty$. This is the main difference when considering the limit $p_S \to 0$ with fixing n as in Section 4.1.

Note that the performance of a subordinate who knows the information initially is independent of the bias parameter α. Therefore, under Assumptions 1 and 2, we can show that[20]

$$\lim_{n\to\infty} \Gamma(\alpha \mid \hat{\alpha}, G) = G(\alpha)$$

for each α. Replacing p_S with $P_{\alpha}^{I_s}(I_m^S = 1)$, we can show the same result as long as $\alpha < 1$. □

Proof of Proposition 6. By Proposition 5, we can focus on the case where the manager's information state is $I_m^S = 1$. As we consider the limit $p_S \to 0$, the equilibrium effort level of the subordinate, $\tilde{e}_s^{\alpha}$, is a solution to

$$P_0^{\alpha}(I_m^S = 1) v_1(\tilde{e}_s^{\alpha}, e_m) + u'(\tilde{e}_s^{\alpha}) = 0,$$

$$u'(e_m) + v_2(\tilde{e}_s^{\alpha}, e_m) = 0.$$

Therefore, as for the subordinate, the optimization problem is the same as the case for the main

[20] Indeed, as $G(\alpha \mid \hat{\alpha}, \pi) = \int dG(\alpha) \frac{p_S f(\pi-(u(1)+V)) + (1-p_S) b_\pi(\alpha)}{p_S A_\pi + (1-p_S) B_\pi}$ and $u(1) + V > \max^{\alpha} x^{\hat{\alpha}}(\alpha)$, by taking the limit $\pi \to \infty$, $\frac{p_S f(\pi-(u(1)+V)) + (1-p_S) b_\pi(\alpha)}{p_S f(\pi-(u(1)+V)) + (1-p_S) B_\pi} \to 1$ by Assumption 2.

model, and then $e_s^\alpha = \tilde{e}_s^\alpha$ for each α.

Similarly, the optimal effort level of the manager, $\tilde{e}_m^{\hat{\alpha}}$ is a solution to

$$v_1(e_s, \tilde{e}_m^{\hat{\alpha}}) + u'(e_s) = 0,$$

$$u'(\tilde{e}_m^{\hat{\alpha}}) + P_{1,I_m^S}^{\hat{\alpha}}(I_s = 0) v_2(e_s, e_{m,1}^{\hat{\alpha}}) = 0.$$

Then, the difference to (2) is that $P_{1,I_m^S}^{\hat{\alpha}}(I_s = 0) = 1 - \hat{\alpha}(1-(1-\beta)(1-I_m^S)) = 1 - \hat{\alpha}[\beta + (1-\beta)I_m^S] \geqslant 1 - \hat{\alpha}$, and it is $P_1^{\hat{\alpha}}(I_s = 0) = 1 - \hat{\alpha}$ at (2). Therefore, $\tilde{e}_m^{\hat{\alpha}} = e_m^{\hat{\alpha}[\beta+(1-\beta)I_m^S]}$.

Using this notation, we calculate the per-capita performance in the performance-based promotion system as follows:

$$\lim_{p_S \to 0} \lim_{n\to\infty} \Pi_{PB}^{p_S,n} = \lim_{n\to\infty} \frac{1}{n} (\mathbb{E}[\theta^{\max,n}]) + \iint dG(\alpha) dG(\hat{\alpha}) \left[u(\tilde{e}_m^{\hat{\alpha}}) + u(\tilde{e}_s^\alpha) + v(\tilde{e}_s^\alpha, \tilde{e}_m^{\hat{\alpha}}) \right],$$

where $\theta^{\max,n}$ is a random variable being followed by cdf, $[F(\cdot)]^n$.

On the contrary, under the random promotion system, since almost all managers have no information during their subordinate periods, $P_{1,I_m^S}^{\hat{\alpha}}(I_s = 0) = \hat{\alpha}$, which is the same as that of

$$\lim_{p_S \to 0} \lim_{n\to\infty} \Pi_{rand}^{p_S,n} = \lim_{n\to\infty} \frac{1}{n} \left(\bar{\theta}\right) + \iint dG(\alpha) dG(\hat{\alpha}) \left[u(e_m^{\hat{\alpha}}) + u(e_s^\alpha) + v(e_s^\alpha, e_m^{\hat{\alpha}}) \right].$$

Suppose that F is a normal distribution. The value $\theta^{\max,n}$ follows a distribution $\frac{x}{\sqrt{2\ln n}} + b_n$, where x follows a Gumbel distribution[21] and $b_n = (2\ln(n))^{1/2} - \frac{(1/2)\ln(4\pi\ln(n))}{(2\ln(n))^{1/2}}$.[22] Then, $\mathbb{E}[\theta^{\max,n}/n] \to 0$ as $n \to \infty$. This implies that the per-capita gain from selecting the most productive subordinate converges to 0.

We now consider the remaining parts. To consider the difference in the per-capita overall

[21] The Gumbel distribution has cdf $\exp(-e^{-x})$ on $x \in \mathbb{R}$.

[22] See David and Nagaraja (2003), Example 10.5.3, p.302.

performance other than noise terms θ, we denote

$$\Pi(I_m^S) = \iint dG(\alpha)dG(\hat{\alpha})[u(e_s^\alpha) + u(\tilde{e}_m^{\hat{\alpha}}) + v(e_s^\alpha, \tilde{e}_m^{\hat{\alpha}})]$$
$$= \iint dG(\alpha)dG(\hat{\alpha})[u(e_s^\alpha) + u(e_m^{\rho(\hat{\alpha})}) + v(e_s^\alpha, e_m^{\rho(\hat{\alpha})}))],$$

where $\rho(\hat{\alpha}) = \hat{\alpha}(\beta+(1-\beta)(1-I_m^S) \leqslant \hat{\alpha}$. Then, $\lim_{p_S\to 0}\lim_{n\to\infty}\Pi_{rand}^{p_S,n} = \Pi(0)$ and $\lim_{p_S\to 0}\lim_{n\to\infty}\Pi_{PB}^{p_S,n} = \Pi(1)$. Differentiating Π with respect to I_m^S yields

$$\frac{\partial \Pi}{\partial I_m^S} = \iint dG(\alpha)dG(\hat{\alpha})[u'(e_m^{\rho(\hat{\alpha})}) + [v_2(e_s^\alpha, e_m^{\rho(\hat{\alpha})})]]\frac{\partial e_m^{\rho(\hat{\alpha})}}{\partial \rho}\hat{\alpha}(1-\beta).$$

Claim 4. *1. If α is sufficiently small, $\frac{\partial e_m^\alpha}{\partial \alpha}\alpha$ is increasing in α.*

2. If $\hat{\alpha} \geqslant \alpha$, $u'(e_m^{\hat{\alpha}}) + v_2(e_s^\alpha, e_m^{\hat{\alpha}}) \leqslant u'' \cdot [e_m^{\hat{\alpha}} - e_m^\alpha]$.

3. If $\hat{\alpha} < \alpha$, $u'(e_m^{\hat{\alpha}}) + v_2(e_s^\alpha, e_m^{\hat{\alpha}}) \leqslant v_{21} \cdot [e_s^\alpha - e_s^{\hat{\alpha}}]$.

Proof of Claim 4. 1. Note that differentiating $\frac{\partial e_m^\alpha}{\partial \alpha}\alpha$ gives

$$\frac{\partial e_m^\alpha}{\partial \alpha} + \frac{\partial^2 e_m^\alpha}{\partial \alpha^2}\alpha. \tag{6}$$

Then, as $\frac{\partial e_m^\alpha}{\partial \alpha} > 0$, if α is sufficiently small, (6) is positive.

2. If $\hat{\alpha} > \alpha$,

$$\begin{aligned} u'(e_m^{\hat{\alpha}}) + v_2(e_s^\alpha, e_m^{\hat{\alpha}}) &\leqslant u'(e_m^{\hat{\alpha}}) + v_2(e_s^\alpha, e_m^*(\alpha)) \\ &\leqslant u'(e_m^{\hat{\alpha}}) - u'(e_m^*(\alpha)) && \text{(by FOC in problem (1))} \\ &\leqslant u'' \cdot [e_m^{\hat{\alpha}} - e_m^*(\alpha)] && \text{(by concavity of } u'\text{)} \\ &\leqslant u'' \cdot [e_m^{\hat{\alpha}} - e_m^\alpha] < 0. && \text{(by } e_m^\alpha \geqslant e_m^*(\alpha) \text{ and } u'' < 0\text{)} \end{aligned}$$

3. If $\hat{\alpha} < \alpha$,

$$\begin{aligned}
u'(e_m^{\hat{\alpha}}) + v_2(e_s^{\alpha}, e_m^{\hat{\alpha}}) &\leqslant u'(e_m^{\hat{\alpha}}) + v_2(e_s^{\alpha}, e_m^{*}(\hat{\alpha})) && \text{(by } v_{22} < 0\text{)} \\
&\leqslant u'(e_m^{*}(\hat{\alpha})) + v_2(e_s^{\alpha}, e_m^{*}(\hat{\alpha})) && \text{(by } u'' < 0\text{)} \\
&= -v_2(e_s^{\hat{\alpha}}, e_m^{*}(\hat{\alpha})) + v_2(e_s^{\alpha}, e_m^{*}(\hat{\alpha})) && \text{(by FOC in problem (1))} \\
&\leqslant v_{21} \cdot [e_s^{\alpha} - e_s^{\hat{\alpha}}]. && \text{(by concavity of } v_2\text{)}
\end{aligned}$$

□

If $\beta \approx 1$, as $\rho(\alpha) \approx \alpha$, by Claim 4,

$$\begin{aligned}
\frac{1}{1-\beta}\frac{\partial \Pi}{\partial I_m^S} &\leqslant \int_{\alpha \leqslant \hat{\alpha}} dG(\alpha)dG(\hat{\alpha})u'' \cdot [e_m^{\hat{\alpha}} - e_m^{\alpha}]\frac{\partial e^{\hat{\alpha}}}{\partial \hat{\alpha}}\hat{\alpha} + \int_{\alpha > \hat{\alpha}} dG(\alpha)dG(\hat{\alpha})v_{21} \cdot [e_s^{\alpha} - e_s^{\hat{\alpha}}]\frac{\partial e^{\hat{\alpha}}}{\partial \hat{\alpha}}\alpha \\
&= \int_{\alpha \leqslant \hat{\alpha}} dG(\alpha)dG(\hat{\alpha})u'' \cdot [e_m^{\hat{\alpha}} - e_m^{\alpha}]\frac{\partial e^{\hat{\alpha}}}{\partial \hat{\alpha}}\hat{\alpha} + \int_{\alpha < \hat{\alpha}} dG(\alpha)dG(\hat{\alpha})v_{21} \cdot [e_s^{\hat{\alpha}} - e_s^{\alpha}]\frac{\partial e^{\alpha}}{\partial \alpha}\alpha \\
&\leqslant \int_{\alpha \leqslant \hat{\alpha}} dG(\alpha)dG(\hat{\alpha})[u'' \cdot [e_m^{\hat{\alpha}} - e_m^{\alpha}] + v_{21} \cdot [e_s^{\hat{\alpha}} - e_s^{\alpha})]\frac{\partial e^{\hat{\alpha}}}{\partial \hat{\alpha}}\hat{\alpha}.
\end{aligned}$$

By symmetry of v, $e_s^{\hat{\alpha}} = e_m^{\hat{\alpha}}$. Also note that $0 > v_{21} + u''$ is implied by Assumption 1(iii). Then, we conclude that $\frac{\partial \Pi}{\partial I_m^S} < 0$. □

Appendix C Numerical Analyses

This section presents a numerical analysis under the setting given in Example 1:

$$\begin{aligned}
v(e_s, e_m) &= -\frac{\gamma}{2}(e_s)^2 + \gamma_{ms}(1-e_m)(1-e_s) - \frac{\gamma}{2}(e_m)^2 \\
u(e) &= \frac{\zeta}{2}(1-(1-e)^2) = \zeta(e - e^2/2)
\end{aligned}$$

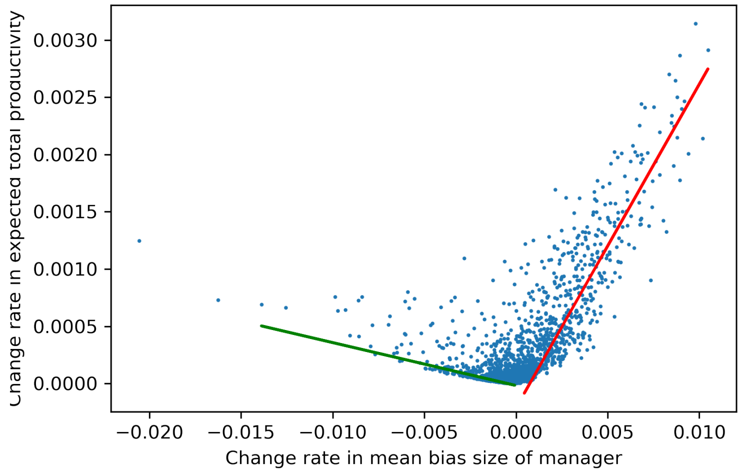


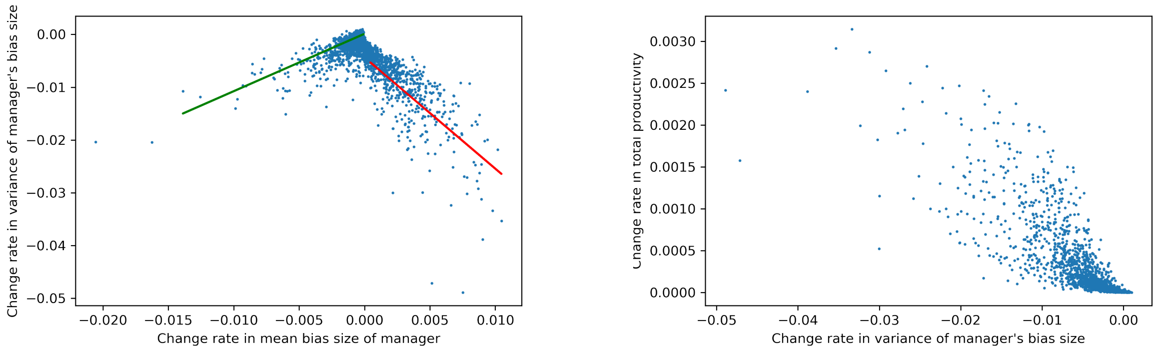
Figure 2: The change rates in the average bias and overall performance

where γ and ζ are positive and $\gamma_{ms} > 0$.[23] We randomly pick 2000 distributions with their support having 10 elements of $[0, 1]$ as G, and compute the distribution of bias α of the subordinate with the highest performance when subordinates and the manager share the same bias distribution G. This means that we calculate $\mathcal{G}_{G,G}$ for each chosen distribution G.

We first examine the relationship between the rate of change in the average bias of promoted subordinates and overall performance. Here, we consider the bias distribution of the subordinate with the highest performance as the bias distribution of the next-generation manager. For each $H \in \{G, \mathcal{G}_{G,G}\}$, let $\alpha_H = \int \alpha dH(\alpha)$ and $\bar{\Pi}_H = \int dH(\hat{\alpha}) \left[u(e_m^{\hat{\alpha}}) + \int dG(\alpha) \left(u(e_s^{\alpha}) + v(e_s^{\alpha}, e_m^{\hat{\alpha}}) \right) \right]$ be the mean of distribution H and the expected overall performance when the manager's bias follows distribution H, respectively. Then, we can compute the expected rate of change in the overall performance $(\frac{\bar{\Pi}_{\mathcal{G}_{G,G}} - \bar{\Pi}_G}{\bar{\Pi}_G})$ and the degree of managers' biases $(\frac{\alpha_{\mathcal{G}_{G,G}} - \alpha_G}{\alpha_G})$. Figure 2 depicts the result of numerical calculation.

Next, we compute the relationships between the rate of change in the average and the variance of bias and between the rate of change in the average of bias and overall performance. Figures 3 (a) and 3 (b) illustrate the results of these numerical calculations.

[23] In the current analysis, we set $n = 5$, $\gamma = 0.05$, $\zeta = 1$, $\gamma_{ms} = 1$, and F is the normal distribution with mean 5 and variance 4.



(a) The change rates in the average and variance of the bias

(b) The change rates in the variance of the bias and overall performance

Figure 3: The effects on variance and the relationship with performance and average bias

We can summarize the above results as follows:

1. The rate of change in overall performance is generally positive. When the rate of change in the average α is positive, the rate of change in overall performance is high. (Figure 2)

2. The higher the rate of change in bias, the greater the decrease in the rate of change in variance. (Figure 3 (a))

3. The greater the decrease in the rate of change in variance, the greater the increase in the rate of change in performance. (Figure 3 (b))

The results imply that the variance of the subordinate's bias distribution has an important role in explaining the effects of bias on performance.

We examine these observations by comparing $\Gamma(\cdot \mid \hat{\alpha}, G)$ with the original bias distribution of subordinates, G. Note that Γ is the distribution of selected subordinates. Through such a comparison, we can determine which type of subordinate is more likely to be promoted. We show some stochastic dominance properties between the above distributions; either stochastic dominance or uniform variability order (Whitt, 1985) holds. For any two distributions F and $\tilde{F}$ with densities f and $\tilde{f}$, F is *uniformly less variable* than $\tilde{F}$ if $f/\tilde{f}$ is single-peaked, and neither F nor $\tilde{F}$ first-order stochastically dominates the others. In general, if F is uniformly less varying than $\tilde{F}$, and they have the same mean, then, for any convex function ψ, $\int \psi(x)d\tilde{F}(x) \geqslant \int \psi(x)dF(x)$. Thus, the variance is greater for $\tilde{F}$. $\tilde{F} >_{\text{uv}} F$ represents that if F is uniformly less variable than $\tilde{F}$. Furthermore, for any distributions, F and $\tilde{F}$, if F first-order stochastically dominates $\tilde{F}$, we denote $F >_{\text{sd}} \tilde{F}$.

Proposition 7. *Suppose that the set of $\hat{\alpha}$ equals the support of G. There are $\hat{\alpha}_1$ and $\hat{\alpha}_2$ with that* $\min(\text{supp}(G)) \leqslant \hat{\alpha}_1 \leqslant \hat{\alpha}_2 \leqslant \max(\text{supp}(G))$ *such that*

(a) $G >_{\text{sd}} \Gamma(\cdot \mid \hat{\alpha}, G)$ *if* $\hat{\alpha} < \hat{\alpha}_1$

(b) $G >_{\text{uv}} \Gamma(\cdot \mid \hat{\alpha}, G)$ *if* $\hat{\alpha} \in (\hat{\alpha}_1, \hat{\alpha}_2)$.

(c) $\Gamma(\cdot \mid \hat{\alpha}, G) >_{\text{sd}} G$ *if* $\hat{\alpha} > \hat{\alpha}_2$.

Furthermore, we have the following properties. If v is symmetric, $\min(\text{supp}(G)) < \alpha_1$. *If* $\hat{\alpha}_2 < \max(\text{supp}(G))$, $\hat{\alpha}_1 < \hat{\alpha}_2$.

Statement (a) in proposition 7 suggests that if the manager's bias is sufficiently small, the promoted subordinate is likely to have a smaller bias relative to the other colleagues. In contrast, statement (c) means that if the manager's bias is sufficiently large, the promoted subordinate is more likely to have greater bias than others. Statement (b) considers the intermediate case. This shows that, although the degree of bias cannot be compared in terms of stochastic dominance, the promoted subordinate's bias is less varied than that of the original group. The intuition is straightforward. As the subordinate's performance is single-peaked with respect to his bias, the promoted subordinate's bias is likely to be around the peak.

Proposition 7 also shows that if case (c) holds for some $\hat{\alpha}$, it implies that there is $\hat{\alpha}$ that case (b) holds. Similarly, if case (b) holds for some $\hat{\alpha}$, it implies that there is $\hat{\alpha}$ that case (a) holds. However, the reverse relation may not hold. There is G such that for any $\hat{\alpha}$, case (a) holds or for any $\hat{\alpha}$, case (a) or (b) holds. When the original subordinate's bias distribution G stochastically dominates new managers' bias distribution, $\mathcal{G}_{G,G}$, while it shows the existence of $\hat{\alpha}$ that case (a) holds, there may be no α such that case (b) and (c) hold. In contrast, when $\mathcal{G}_{G,G}$ stochastically dominates G, case (c) holds for some α. Therefore, there is $\hat{\alpha}$ such that case (b) holds, and then the variance is likely to decrease. This result is consistent with the results obtained via numerical calculation, which shows that the size of improvement of performance is larger when the mean degree of bias among new managers increases. This is because, in the case of the increment of bias, either case (b) or (c) needs to hold. In such a case, the variance of the bias tends to decrease further, which further improves performance.

Appendix C.1 Proof in Appendix C

Proof of Proposition 7. We assume that $\min(\mathrm{supp}(G)) = 0$ and $\max(\mathrm{supp}(G)) = 1$ for notational simplicity, but this does not lose the generality of the proof.

Note that the density of $\Gamma(\cdot \mid \hat{\alpha}, G)$ is rewritten as

$$\begin{aligned}\Gamma'(\alpha \mid \hat{\alpha}, G) &= g(\alpha)\left[\int d\pi f(\pi - x^{\hat{\alpha}}(\alpha))n\left[\int dG(\alpha')F(\pi - x^{\hat{\alpha}}(\alpha'))\right]^{n-1}\right] \\ &= g(\alpha)\left[\int d\pi f(\pi)n\left[\int dG(\alpha')F(\pi + x^{\hat{\alpha}}(\alpha) - x^{\hat{\alpha}}(\alpha'))\right]^{n-1}.\right]\end{aligned}$$

Let us define σ as follows.

$$\sigma_{\hat{\alpha}}(\alpha) := \frac{\Gamma'(\alpha \mid \hat{\alpha}, G)}{g(\alpha)} = \int d\pi f(\pi)n\left[\int dG(\alpha')F(\pi + x^{\hat{\alpha}}(\alpha) - x^{\hat{\alpha}}(\alpha'))\right]^{n-1}.$$

Since $x^{\hat{\alpha}}(\alpha)$ is single-peaked with respect to α and F is increasing, $\sigma_{\hat{\alpha}}(\alpha)$ is single-peaked with respect to α.

$\sigma_{\hat{\alpha}}(\alpha) > 1 > \sigma_{\hat{\alpha}}(\alpha')$ since $\int dG(\alpha) = \int dG(\alpha)\sigma_{\hat{\alpha}}(\alpha) = 1$, for some α and α'. Therefore, because $\sigma_{\hat{\alpha}}(\cdot)$ is single-peaked, exactly one of the following holds:

Case (a) $\sigma_{\hat{\alpha}}(0) > 1 > \sigma_{\hat{\alpha}}(1)$.

Case (b) $1 > \max\{\sigma_{\hat{\alpha}}(0), \sigma_{\hat{\alpha}}(1)\}$.

Case (c) $\sigma_{\hat{\alpha}}(1) > 1 > \sigma_{\hat{\alpha}}(0)$.

In case (a) and (c), since $g(\alpha) > \Gamma'(\alpha)$ if and only if $\sigma_{\hat{\alpha}}(\alpha) < 1$, G and Γ are single-crossing. Then, by Shaked and Shanthikumar (2007, Theorem 1.A.12), $G >_{\mathrm{sd}} \Gamma$ in case (a), and $\Gamma >_{\mathrm{sd}} G$ in case (c). In case (b), we can show that $G >_{\mathrm{uv}} \Gamma$ by the definition of the uniform variance order.

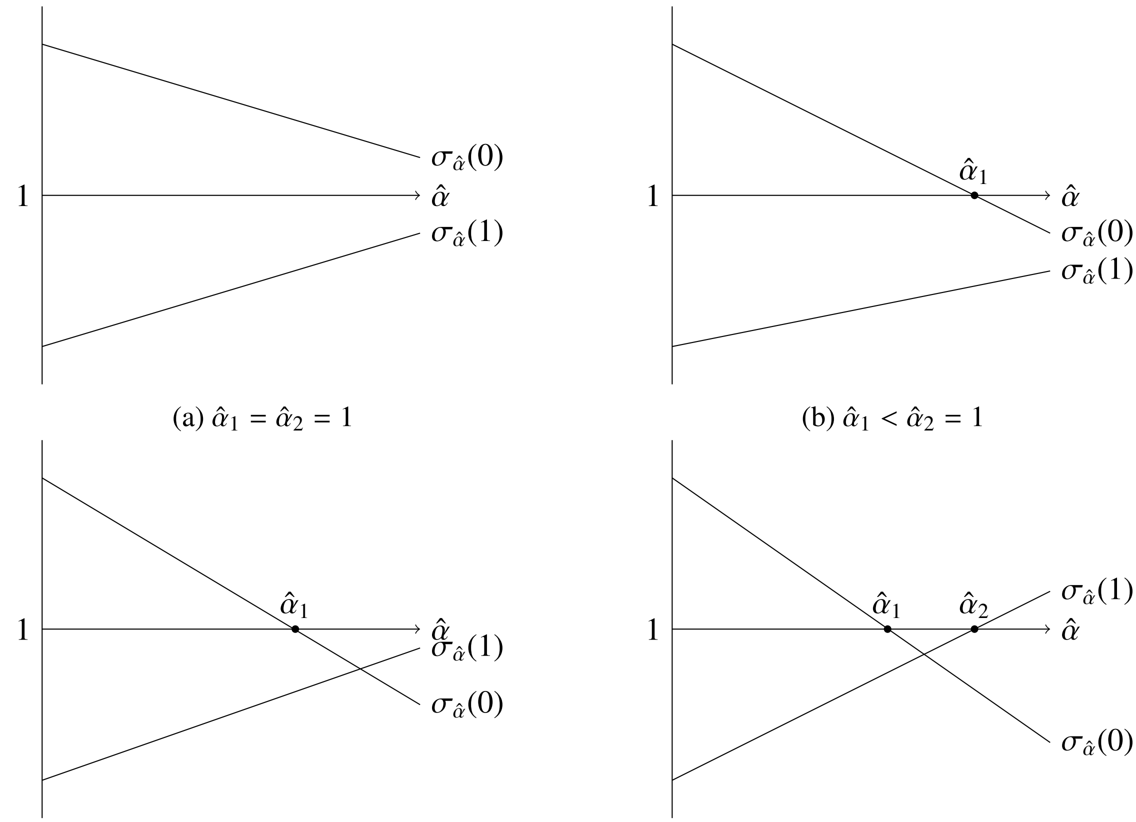
Now, we investigate the cases of (a)–(c). Note that if v is symmetric, by Lemma 2 (b), the maximizer of $x^{\hat{\alpha}}(\cdot)$ is no greater than $\hat{\alpha}$. Then, $x^0(\cdot)$ is decreasing, which implies that $\sigma_0(0) > 1 > \sigma_0(1)$. Consequently, $\hat{\alpha} = 0$ is case (a).

To verify the case of other $\hat{\alpha}$, we investigate the relation between $\hat{\alpha}$ and the stochastic order. Note that

$$\begin{aligned}\frac{\partial \sigma_{\hat{\alpha}}}{\partial \hat{\alpha}}(\alpha) = &\int d\pi f(\pi) n(n-1)\left[\int dG(\alpha')F(\pi + x^{\hat{\alpha}}(\alpha) - x^{\hat{\alpha}}(\alpha'))\right]^{n-2} \\ &\times \int dG(\alpha') f(\pi + x^{\hat{\alpha}}(\alpha) - x^{\hat{\alpha}}(\alpha'))\left(\frac{\partial x^{\hat{\alpha}}}{\partial \hat{\alpha}}(\alpha) - \frac{\partial x^{\hat{\alpha}}}{\partial \hat{\alpha}}(\alpha')\right).\end{aligned}$$

Since $\frac{\partial^2 x^{\hat{\alpha}}(\alpha)}{\partial \alpha \partial \hat{\alpha}} > 0$ by $v_{12} > 0$, $\sigma_{\hat{\alpha}}(0)$ is decreasing in $\hat{\alpha}$, and $\sigma_{\hat{\alpha}}(1)$ is increasing in $\hat{\alpha}$. Then, as Figure 4 illustrates, $\sigma_{\hat{\alpha}}(\cdot)$ is in case (a) for each $\hat{\alpha} < \hat{\alpha}_1$, in case (b) for each $\hat{\alpha} \in (\hat{\alpha}_1, \hat{\alpha}_2)$, and in case (c) for each $\hat{\alpha} > \hat{\alpha}_2$.

Finally, we note that $\hat{\alpha}_1 = \hat{\alpha}_2 < 1$ is impossible. If this is the case, as Figure 5 shows, both $\sigma_\alpha(1)$ and $\sigma_\alpha(0)$ are larger than 1. As $\sigma_{\hat{\alpha}}$ is single-peaked, this implies that $\int dG(\alpha)\sigma_{\hat{\alpha}}(\alpha) > 1$, which is a contradiction. □



(a) $\hat{\alpha}_1 = \hat{\alpha}_2 = 1$

(b) $\hat{\alpha}_1 < \hat{\alpha}_2 = 1$

(c) $\hat{\alpha}_1 < \hat{\alpha}_2 = 1$

(d) $\hat{\alpha}_1 < \hat{\alpha}_2 < 1$

Figure 4: Patterns of the relations between $\sigma_{\hat{\alpha}}(0)$ and $\sigma_{\hat{\alpha}}(1)$

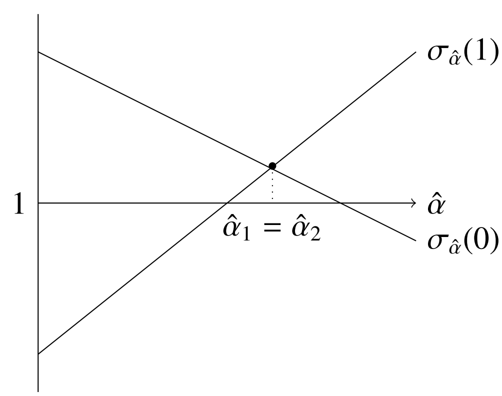


Figure 5: The case for $\hat{\alpha}_1 = \hat{\alpha}_2 < 1$

Supplementary material for

Evolution of Information Projection Bias through Costly Communication in Overlapping Generations Organizations

Kohei Daido and Tomoya Tajika

This material analyzes the case where $v_{12}(e_s, e_m) < 0$ for each e_s, e_m. Almost the same results as in the case of $v_{12} > 0$ hold for the one-shot game except for overall performance.

Lemma 6. *Suppose that $v_{12} < 0$. The following statements hold. (a) Both $\frac{\partial e_m^\alpha}{\partial \alpha}$ and $\frac{\partial e_s^\alpha}{\partial \alpha}$ are positive, (b) $x^{\hat{\alpha}}(\alpha)$ is decreasing in $\hat{\alpha}$ and is single-peaked with respect to α, (c) $\arg\max_\alpha x^{\hat{\alpha}}(\alpha) \leqslant \hat{\alpha}$, (d) $\frac{\partial^2 x^{\hat{\alpha}}(\alpha)}{\partial \alpha \partial \hat{\alpha}} < 0$, and (e) $E[\Pi \mid \hat{\alpha}]$ is decreasing with respect to $\hat{\alpha}$.*

One prominent difference from the case of complementarity in communication efforts is that the overall performance decreases in the manager's bias. This is single-peaked in the case of complementarity.

The evolution dynamics of bias are also different from the complementarity case, where as Proposition 2 shows, bias can increase. However, in the case of substitutability, this does not occur, as the following proposition shows:

Proposition 8. *If $v_{12} < 0$, $G \succ_{\text{sd}} \Gamma(\cdot \mid \hat{\alpha}, G)$ for any $\hat{\alpha}$.*

Proofs

Proof of Lemma 6. (a) As in the proof of Lemma 2,

$$\frac{\partial e_s^\alpha}{\partial \alpha} = \frac{p_M v_1 (u''(e_m) + (1 - p_S) v_{22})}{\det(D_S)}$$
$$\frac{\partial e_m^\alpha}{\partial \alpha} = \frac{(1 - p_S) v_2 [p_M v_{11} + u''(e_s)]}{\det(D_M)},$$

which implies that $\frac{\partial e_m^\alpha}{\partial \alpha}$ and $\frac{\partial e_s^\alpha}{\partial \alpha}$ are positive.

(b) Note that $u(e) + v(e, e_m)$ is concave in e, and decreasing in e_m. By Lemma 2, e_m^α and e_s^α are increasing in α, and $x^{\hat{\alpha}}(\alpha)$ is single-peaked with respect to α and decreasing in $\hat{\alpha}$.

(c) Now we consider the maximizer. Let $e_s^*(\hat{\alpha})$ be the effort level of the subordinate that the manager with degree $\hat{\alpha}$ believes.

Claim 5. *If $v_{12} < 0$, $e_s^{\hat{\alpha}} \geqslant e_s^*(\hat{\alpha})$.*

Proof of Claim 5. We consider the following auxiliary problem:

$$\begin{cases} \varphi(\beta) v_1(e_s, e_m) + u'(e_s) = 0, \\ u'(e_m) + \psi(\beta) v_2(e_s, e_m) = 0, \end{cases} \tag{7}$$

where

$$\varphi(\beta) = \beta + (1 - \beta)(1 - \alpha), \qquad \text{and} \qquad \psi(\beta) = (1 - \beta) + \beta(1 - \alpha).$$

Let $(\hat{e}_s^\beta, \hat{e}_m^\beta)$ be the solution to problem (7). We can show that $\hat{e}_s^1 = e_s^*(\alpha)$ and $\hat{e}_s^0 = e_s^\alpha$. The derivative with respect to β is given by

$$\frac{\partial e_s^\beta}{\partial \beta} = -\frac{\alpha}{\det(D_\beta)} \left[v_1 [\psi(\beta) v_{11} + u''] + v_2 [\varphi(\beta) v_{12}] \right] < 0,$$

where

$$D_\beta = \begin{pmatrix} \varphi(\beta) v_{11} + u'' & \varphi(\beta) v_{12} \\ \psi(\beta) v_{21} & \psi(\beta) v_{22} + u'' \end{pmatrix}.$$

Then, we conclude that $e_s^\alpha \geqslant e_s^*(\alpha)$. $\square$

Using this claim, we prove that the maximizer of $x^{\hat{\alpha}}(\alpha)$ is less than $\hat{\alpha}$. The derivative of $x^{\hat{\alpha}}(\alpha)$

with respect to α is calculated as follows.

$$\frac{\partial x^{\hat{\alpha}}(\alpha)}{\partial \alpha} = \frac{\partial e_s^{\alpha}}{\partial \alpha}\left[u'(e_s^{\alpha}) + v_1(e_s^{\alpha}, e_m^{\hat{\alpha}})\right].$$

Suppose that $\alpha > \hat{\alpha}$. Then, by $u'' < 0$, v_{11} and $e_s^*(\hat{\alpha}) < e_s^{\alpha}$, and monotonicity of e_s^{α},

$$\begin{aligned} u'(e_s^{\alpha}) + v_1(e_s^{\alpha}, e_m^{\hat{\alpha}}) &\leqslant u'(e_s^*(\alpha)) + v_1(e_s^*(\alpha), e_m^{\hat{\alpha}}) \\ &< u'(e_s^*(\hat{\alpha})) + v_1(e_s^*(\hat{\alpha}), e_m^{\hat{\alpha}}) = 0, \end{aligned}$$

which implies that $\frac{\partial x^{\hat{\alpha}}(\alpha)}{\partial \alpha} < 0$. Then, $x^{\hat{\alpha}}(\alpha)$ has decreasing region in $(\hat{\alpha}, 1)$, and thus the maximizer is less than $\hat{\alpha}$.

(d) As for the cross derivative,

$$\frac{\partial^2 x^{\hat{\alpha}}(\alpha)}{\partial \alpha \partial \hat{\alpha}} = \frac{\partial e_s^{\alpha}}{\partial \alpha}\frac{\partial e_m^{\hat{\alpha}}}{\partial \hat{\alpha}} v_{12}(e_s^{\alpha}, e_m^{\hat{\alpha}}) < 0.$$

(e) First note that $E[\Pi \mid \hat{\alpha}]$ is concave with respect to $e_m^{\hat{\alpha}}$. As $e_m^{\hat{\alpha}}$ is increasing in $\hat{\alpha}$, $E[\Pi \mid \hat{\alpha}]$ is single-peaked with respect to $\hat{\alpha}$. Moreover, $\frac{\partial E[\Pi|\hat{\alpha}]}{\partial \hat{\alpha}} < 0$ at $\hat{\alpha} = 0$. To show this, note that

$$E[\Pi \mid \hat{\alpha}] = u(e_m^{\hat{\alpha}}) + \int [u(e_s^{\alpha}) + v(e_s^{\alpha}, e_m^{\hat{\alpha}})] d\alpha,$$

and

$$\frac{\partial E[\Pi \mid \hat{\alpha}]}{\partial \hat{\alpha}} = \frac{\partial e_m^{\hat{\alpha}}}{\partial \hat{\alpha}}\left[u'(e_m^{\hat{\alpha}}) + \int v_2(e_s^{\alpha}, e_m^{\hat{\alpha}}) d\alpha\right].$$

Note also that $u(e_m^0) + v_2(e_s^*(0), e_m^0) = 0$ by the first order condition for the manager's problem.

Then, by $v_{12} < 0$, $e_s^*(\alpha) < e_s^\alpha$, and e_s^α is increasing in α,

$$
\begin{aligned}
0 &= u(e_m^0) + v_2(e_s^*(0), e_m^0) \\
&> u(e_m^0) + v_2(e_s^0, e_m^0) \\
&> u(e_m^0) + v_2(e_s^\alpha, e_m^0)
\end{aligned}
$$

for each α. This shows that $\frac{\partial E[\Pi|\hat{\alpha}]}{\partial \hat{\alpha}} > 0$. Then, by the single-peakedness of $E[\Pi \mid \hat{\alpha}]$, it is decreasing in $\hat{\alpha}$. □

Proof of Proposition 8. As in the proof of Proposition 7, let

$$
\sigma_{\hat{\alpha}}(\alpha) := \frac{\Gamma'(\alpha \mid \hat{\alpha}, G)}{g(\alpha)} = \int d\pi f(\pi) n \left[\int dG(\alpha') F(\pi + x^{\hat{\alpha}}(\alpha) - x^{\hat{\alpha}}(\alpha')) \right]^{n-1}.
$$

As $\bar{x}^{\hat{\alpha}}(\alpha)$ is single-peaked with respect to α, either one of the following holds.

Case (a) $\sigma_{\hat{\alpha}}(0) > 1 > \sigma_{\hat{\alpha}}(1)$.

Case (b) $1 > \max\{\bar{\sigma}_{\hat{\alpha}}(0), \sigma_{\hat{\alpha}}(1)\}$.

Case (c) $\sigma_{\hat{\alpha}}(1) > 1 > \sigma_{\hat{\alpha}}(0)$.

Note that as the peak is less than $\hat{\alpha}$, $x^0(\alpha)$ is decreasing in α. Then, $\bar{\sigma}_{\hat{\alpha}}(\alpha)$ is also decreasing in α. Then, $\sigma_{\hat{\alpha}}(0) > 1 > \sigma_{\hat{\alpha}}(1)$ holds for $\hat{\alpha} = 0$. Note also that

$$
\begin{aligned}
\frac{\partial \bar{\sigma}_{\hat{\alpha}}}{\partial \hat{\alpha}}(\alpha) = & \int d\pi f(\pi) n(n-1) \left[\int dG(\alpha') F(\pi + x^{\hat{\alpha}}(\alpha) - x^{\hat{\alpha}}(\alpha')) \right]^{n-2} \\
& \times \int dG(\alpha') f(\pi + \bar{x}^{\hat{\alpha}}(\alpha) - \bar{x}^{\hat{\alpha}}(\alpha')) \left(\frac{\partial \bar{x}^{\hat{\alpha}}}{\partial \hat{\alpha}}(\alpha) - \frac{\partial \bar{x}^{\hat{\alpha}}}{\partial \hat{\alpha}}(\alpha') \right).
\end{aligned}
$$

As $\frac{\partial^2 x^{\hat{\alpha}}(\alpha)}{\partial \alpha \partial \hat{\alpha}} < 0$ by Lemma 6 (d), $\sigma_{\hat{\alpha}}(0)$ is increasing in $\hat{\alpha}$. This implies that for any $\hat{\alpha}$, $\sigma_{\hat{\alpha}}(0) > 1$. As in the proof of Proposition 7, this implies that $G >_{\text{sd}} \Gamma(\cdot \mid \hat{\alpha}, G)$. □