

Article

Hierarchies and Promotions in Political Institutions: Accountability and Selection

B. Pablo Montagnes ^{1,*}, Stephane Wolton ^{2,3}  and Junyan Jiang ⁴ ¹ Department of Political Science, Emory University, Atlanta, GA 30322, USA² Department of Government, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK; s.wolton@lse.ac.uk³ Center for Economic Policy Research, 75007 Paris, France⁴ Department of Political Science, Columbia University, 420 W 118th St, New York, NY 10027, USA; jj3160@columbia.edu

* Correspondence: pablo.montagnes@emory.edu

Abstract

Hierarchies are common in political settings. From judges to elected politicians, as well as from activists to bureaucrats, political agents compete for promotion to higher positions. This paper studies political tournaments and their impact on two aspects of political performance: accountability and selection. While larger tournaments discourage effort, they improve selection. We also discuss the optimal design of tournaments as a function of the principal's objectives and the features of the environment. We find that tournaments of size two (such as two-candidate elections) are generally suboptimal. Our analysis also highlights that increased desirability of promotion always increases effort but can reduce the optimal tournament size under certain conditions. We also present a range of other comparative statics.

Keywords: political hierarchies; tournaments; accountability and selection; political institutions



Academic Editor: Ulrich Berger

Received: 7 March 2025

Revised: 19 June 2025

Accepted: 20 June 2025

Published: 4 July 2025

Citation: Montagnes, B. P., Wolton, S., & Jiang, J. (2025). Hierarchies and Promotions in Political Institutions: Accountability and Selection. *Games*, 16(4), 34. <https://doi.org/10.3390/g16040034>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Hierarchies are common in political settings. Electoral careers often begin at the local level and progress to the provincial or national level (Myerson, 2006). Judges move through the various strata of the judicial system (Epstein et al., 2003). Authoritarian leaders fill vacancies by drawing on administrators at lower levels (Lu & Landy, 2014). Bureaucrats are promoted through various grades of pay and responsibility (Cameron et al., 2013). Political parties have their own set of rules and structures, with members competing for positions on party lists or in leadership roles (Michels, 1915). Legislators seek ministerial positions (Eggers & Spirling, 2016) or assignments to desirable committees (Fong & McCrain, 2024; Groseclose & Stewart, 1998). Even electoral campaigns can be understood in this light, where candidates seek to win office by proving themselves in the eyes of a voter.

Hierarchical structures allow a principal (e.g., a party leader, the president, or an authoritarian ruler) to select the most capable agent (legislator, bureaucrat, or local administrator) to fill a vacant position (cabinet post, agency head, or provincial position). Hierarchies thus presuppose the possibility of promotions. From the perspective of agents, promotions are prizes to be won by besting fellow contestants. From the perspective of the principal, promotions are a way to select the most competent agents while also incentivizing effort. Promotion-based hierarchies are thus best understood as tournaments with a twist.

Rather than just being about incentives—the focus of a large portion of literature since [Lazear and Rosen \(1981\)](#)—the tournament we study also serves as a tool for selection. In our setting, effort is a by-product of selection. Effort occurs as an attempt to mislead the principal, which is an attempt that fails in equilibrium. We emphasize that, within the class of symmetric equilibria, the effort level is uniquely pinned down—something that is not guaranteed in all tournament models that rely on effort incentives. Furthermore, the unique equilibrium level of effort we obtain is equal to the highest level of effort agents exert in traditional tournament models.

Our first key result shows that selection in promotion-based tournaments comes at almost no cost. The principal can use a tournament to select the ablest agent and still obtain a very high level of effort from everyone involved. Yet, this does not imply that selection and accountability—defined as the level of effort agents exert—always go hand in hand.¹

Increasing the number of agents competing with each other can yield an accountability–selection tradeoff for the principal. Having more agents competing always improves the expected ability of the selected agent but also reduces the amount of effort each agent exerts. This has implications for the design of the optimal tournament (i.e., the number of participants allowed to compete), which is a question we briefly study at the end of our paper.

We show that in some settings (e.g., if the principal cares about average performance), a principal may choose to restrict the number of participants in a tournament to provide greater incentives for effort. Yet, we nonetheless find that whenever the principal cares about effort in addition to selection, tournaments with two competing agents are generally dominated by larger tournaments. This result suggests that two-candidate elections are suboptimal from the electorate’s perspective unless the cost of running larger elections is very high.

Our paper also provides a host of additional comparative statics—some are intuitive, while others are more subtle. Effort is increasing with the value of the prize (as agents find promotion more valuable) and decreasing with the heterogeneity of ability and the informativeness of agents’ outputs as signals of their skills (since in both cases, effort is less likely to determine promotion). In turn, selection always improves with the variance in agents’ types, as the principal only cares about the highest draw of ability. More surprisingly, perhaps, even though effort increases with the value of the prize, the optimal tournament size decreases with the desirability of promotion whenever the principal cares about the promoted agent’s effort. The key force behind this result is that the principal weighs the gain and loss of adding one competitor when designing a tournament. A higher prize exacerbates the negative effect on the effort of increasing the pool of competitors. In turn, greater desirability of promotion leads to a larger tournament when the principal cares about total effort, since a larger prize enhances the gain from increased size.

Literature Review. The theoretical analysis of tournaments is not new. Beginning with the seminal work of [Lazear and Rosen \(1981\)](#), tournament incentives have been studied extensively (see [Waldman, 2013](#), for a review). Early results emphasized that tournaments are dominated by (sufficiently rich) contracts when it comes to incentive provision, since they introduce noise into compensation schemes ([Green & Stokey, 1983](#)).

Later studies justified tournaments as an optimal response to specific features of the environment, such as common shocks ([Maskin et al., 2000](#)) or commitment problems ([Malcomson, 1986](#)), or assumed a tournament and characterized the optimal prize structure ([Krishna & Morgan, 1998](#), builds on [Nalebuff & Stiglitz, 1983](#)). [Connelly et al. \(2014\)](#) provided a useful review of the large literature on tournaments as a tool for accountability, from both a theoretical and an empirical angle.

For a long time, tournament models failed to consider selection. This was one of the critiques raised by [Prendergast \(1999\)](#). Since then, tremendous progress has been made in addressing this gap. While some works have focused exclusively on selection (e.g., [Clark & Riis, 2001](#); [Ryvkin, 2010](#)), others, like us, have incorporated both effort and selection. [Gürtler and Gürtler \(2015\)](#) showed how two agents may be motivated to exert greater effort when they are known to have different average abilities. [Höchtel et al. \(n.d.\)](#) looked at a two-stage tournament and found that selection benefits from heterogeneous competitors competing early with each other. [Levy and Zhang \(2024\)](#) argue that, if agents are homogeneous in their ability, pooled contests (where all participants compete with each other) may be better than two-stage tournaments.

While these contributions are important, we advance the literature along two fronts. First, we return to the traditional model of the tournament ([Höchtel et al., n.d.](#); [Levy & Zhang, 2024](#), use a contest function) and solve a tournament game with promotion involving a large number of agents ([Gürtler & Gürtler, 2015](#), whereas prior work only considered a two-agent tournament). Second, we discuss the gains and losses from increasing the size of the tournament, taking into consideration both accountability and selection.²

By considering accountability and selection, our work is related to agency models of elections (see [Ashworth, 2012](#), for a review). Canonical political career concerns models (e.g., [Ashworth, 2005](#)) are, however, a special case of political tournaments in which only the performance of one agent is observed and compared to a benchmark. Our richer setup allows us to compare how electoral incentives differ from other types of tournaments and generate novel comparative statics relating the optimal size of a tournament to model parameters.

Further, while electoral accountability models focus on bottom-up accountability, promotion-based hierarchies are best understood as top-down accountability problems. [Ghosh and Waldman \(2010\)](#) studied how promotions can be used to screen agents. They focused on a one-to-one relationship between firm owners and workers and so had nothing to say about the size effect of promotion tournaments. Our approach, instead, is rooted in the idea that top-down accountability is a many-to-one problem, and agents are rarely judged in isolation. [Montagnes and Wolton \(2019\)](#) and [Wang et al. \(2022\)](#) studied a related question. However, they assumed that the principal has coarse information about her agents' performances and emphasized the role of punishments (purges in authoritarian countries) rather than rewards.

2. Formal Model of Tournaments with Selection

We consider a one-period game with a principal and a set of available agents with cardinality $\bar{n} > 2$. The principal first decides the number of agents $n \in \{1, \dots, \bar{n}\}$ to be considered for promotion to a unique higher-up position valued by all agents. The n randomly selected agents then compete by exerting costly effort to manipulate the principal's promotion decision. We thus describe the subgame played by the selected agents as an n -agent tournament.

Each agent i is characterized by their ability θ_i . Following canonical models of tournaments (e.g., [Lazear & Rosen, 1981](#)) and career concerns models (e.g., [Ashworth, 2005](#); [Holmström, 1999](#)), we assume that θ_i is unknown to all players at the beginning of the game. It is, however, common knowledge that for all $i \in \{1, \dots, n\}$, θ_i is independently and identically distributed (i.i.d.) over the interval $[-\bar{\theta}, \bar{\theta}]$ (possibly unbounded) according to the cumulative distribution function (CDF) F_θ , with variance σ_θ^2 and symmetric and single-peaked probability density function (pdf) f_θ .³ We further assume that F_θ and f_θ are continuously differentiable (i.e., $F_\theta \in \mathcal{C}^2$). The principal cares about the promoted agent's ability. She, however, only observes a noisy signal of agents' skills before making her

promotion decision for those considered for promotion; otherwise, she observes nothing. More specifically, an agent considered for promotion produces an output—denoted q_i for all $i \in \{1, \dots, n\}$ —which consists of the sum of the agent's ability θ_i , their effort a_i , and a noise (luck) term η_i :

$$q_i = a_i + \theta_i + \eta_i \quad (1)$$

Effort a_i is unobserved by the principal. As in career concerns models, agents try to manipulate the signal available to the principal. However, unlike career concerns models, n agents simultaneously compete for promotion by exerting effort. The noise term η_i is also unknown to all players, but it is common knowledge that η_i is i.i.d. over the interval $[-\bar{\epsilon}, \bar{\epsilon}]$ (possibly unbounded) according to the CDF $F_\eta \in \mathcal{C}^2$ with variance σ_η^2 and symmetric and single-peaked pdf f_η .⁴ Notice that the output function only incorporates independent shocks. This assumption is without loss of generality since, as extensively noted in the literature (e.g., [Green & Stokey, 1983](#); [Maskin et al., 2000](#)), any shock that would similarly impact all agents would be differenced out by the principal, who evaluates agents on their relative rather than absolute performance.

An agent receives a payoff V if promoted and 0 if otherwise. The agent also pays a cost $c(a_i)$ to exert effort and manipulate the signal q_i available to the principal. We assume that $c(\cdot)$ is \mathcal{C}^∞ , increasing, and convex. Agent i 's utility function thus takes the following form (with $\mathbb{I}_{\{\text{promoted}\}}$ being the indicator function equal to 1 if i is promoted):

$$U_i(a_i) = V\mathbb{I}_{\{\text{promoted}\}} - c(a_i) \quad (2)$$

The principal's utility depends on both the ability of the promoted agent and the agents' efforts. We assume that increasing the size of the tournament is costly, with the convex cost function $C_P(n)$. The principal's utility can be represented as

$$U_P(n) = \omega g(a_i, \dots, a_{\bar{n}}) + (1 - \omega)\mathbb{I}_{\{i=\text{promoted}\}}\theta_i - C_P(n), \quad (3)$$

where $\omega \in [0, 1]$ is the weight on agents' effort, $g(a_i, \dots, a_{\bar{n}})$ is a continuous function weakly increasing in all its arguments,⁵ and $\mathbb{I}_{\{i=\text{promoted}\}}$ is an indicator function equal to 1 if i is promoted and 0 if otherwise.

To summarize, the timing of the game is as follows:

1. Nature draws relative abilities (θ_i) and noise (η_i) for each agent $i \in \{1, \dots, \bar{n}\}$, and these values are unobserved by all players.
2. The principal selects the number of agents to be considered for promotion, where $n \in \{1, \dots, \bar{n}\}$. For notational convenience and without loss of generality, we assume that the first n agents are selected.
3. Each selected agent i simultaneously chooses their effort a_i .
4. For each selected agent i , the output q_i is observed by all players.
5. The principal promotes one agent among the n selected agents. The game ends, and payoffs are realized.

The equilibrium concept is Perfect Bayesian Equilibrium in pure strategies. We focus on symmetric equilibria. Two well-known aspects of Perfect Bayesian Equilibrium are worth stressing: (i) All players consistently anticipate other players' strategies so that (ii) agent i and the principal have the same expectation about agent j 's effort for all $j \neq i$. Throughout, we posit variance parameters for ability and noise that place the game in the well-behaved region identified in the tournament literature—conditions under which a single symmetric equilibrium is generally obtained. We adopt that equilibrium for all subsequent analysis (This assumption is common in the literature; see [Green & Stokey, 1983](#); [Krishna & Morgan, 1998](#); [Nalebuff & Stiglitz, 1983](#)).⁶

As a preliminary, note that given the principal's objective, she always promotes the agent with the highest expected ability. To form a correct assessment of agents' skills, she adjusts for agents' anticipated effort from the observed outputs. Since her anticipation is correct in equilibrium, we can study promotion and effort separately. We thus depart from the usual (backwards) order of analysis and focus first on agents' effort and how it varies with the model parameters and the size of the tournament.

3. Agents' Effort

We first consider effort in a tournament of size n . Agents not competing for promotion (i.e., agents $m \in \{n+1, \bar{n}\}$) do not exert any effort. Likewise, in a tournament of size one, agent 1 exerts no effort. Therefore, in what follows, we focus on tournaments of size at least 2. We introduce the following notation: let a_j^e be the principal's anticipation of j 's effort.

Agent $i \in \{1, \dots, n\}$ is promoted if and only if the principal expects their ability to be the highest. This occurs whenever they produce the highest output net of anticipated effort, i.e., $q_i - a_i^e \geq \max_{j \neq i} q_j - a_j^e$. Agent i , however, does not observe q_j and must form beliefs about other agents' performances based on (i) their anticipation of the realization of their ability (θ_j) and luck (η_j) and (ii) their effort (a_j). Denote \hat{a}_j^e as i 's anticipation of j 's effort for all $j \neq i$. The probability that i is promoted given their output q_i is $Pr(Promoted|q_i) := Pr\left(q_i - a_i^e \geq \max_{j \neq i} (\theta_j + \eta_j + \hat{a}_j^e - a_j^e)\right)$. Imposing the assumption that anticipated efforts are consistent (i.e., $a_j^e = \hat{a}_j^e$), this reduces to

$$Pr(Promoted|q_i) := Pr\left(q_i - a_i^e \geq \max_{j \neq i} (\theta_j + \eta_j)\right) \quad (4)$$

Equation (4) highlights that only the joint effect of θ_j and η_j enters into agent i 's assessment. Since the two random variables are independent, it is convenient to merge the terms into a 'combined noise' $\epsilon_j := \theta_j + \eta_j$ for all j . For each j , the combined noise is i.i.d. over $[-\bar{\epsilon}, \bar{\epsilon}]$ ($\bar{\epsilon} = \bar{\theta} + \bar{\eta}$) according to the CDF $F_\epsilon \in \mathcal{C}^2$, with variance $\sigma_\epsilon^2 = \sigma_\theta^2 + \sigma_\eta^2$ and pdf f_ϵ . Observe that due to the independence of θ_i and η_i , f_ϵ is single-peaked and symmetric. Using the independence of ϵ_j 's, the probability that i is promoted for a given output q_i is thus $F_\epsilon(q_i - a_i^e)^{n-1}$.

Agent i , however, only chooses their effort, not their output q_i , which corresponds to the sum of effort a_i and their own combined noise ϵ_i . Since agent i correctly treats ϵ_i as a random variable, they maximize the following objective function (with $E_{\epsilon_i}(\cdot)$ the conditional operator over ϵ_i):

$$\begin{aligned} K_i(a_i) &:= V E_{\epsilon_i}(Pr(Promoted|a_i, \epsilon_i)) - c(a_i) \\ &= V \int_{-\bar{\epsilon}}^{\bar{\epsilon}} F_\epsilon(\epsilon + a_i - a_i^e)^{n-1} f_\epsilon(\epsilon) d\epsilon - c(a_i) \end{aligned} \quad (5)$$

Through standard maximization (our assumption on variances guarantees that the second-order condition holds), we can then characterize the agents' effort.

Proposition 1. *There exists a unique equilibrium effort in pure strategies. The equilibrium effort is symmetric—denoted as $\bar{a}(n, \sigma_\epsilon, V)$ —and characterized by the following equation:*

$$c'(\bar{a}(n, \sigma_\epsilon, V)) = V \int_{-\bar{\epsilon}}^{\bar{\epsilon}} f_\epsilon(\epsilon) dF_\epsilon(\epsilon)^{n-1} \quad (6)$$

Proof. Proofs for this section are collected in Appendix A. \square

Our next set of results establishes how the tournament structure—the value of the prize (V), the variance of the combined noise (σ_ϵ), and the size of the tournament (n)—affects the equilibrium effort.

Proposition 2. *The equilibrium level of effort $\bar{a}(n, \sigma_\epsilon, V)$ is as follows:*

- (i) *Decreasing in n (i.e., $\bar{a}(n'', \sigma_\epsilon, V) \leq \bar{a}(n', \sigma_\epsilon, V)$ for $n'' > n' \geq 2$);*
- (ii) *Strictly decreasing in σ_ϵ ;*
- (iii) *Strictly increasing in V .*

Points (ii) and (iii) of Proposition 2 are intuitive. As the prize becomes more attractive (the value of promotion V increases), agents have more incentive to exert effort. As the variance of the combined noise grows, effort is less and less likely to determine the outcome of the tournament. As a consequence, all agents have reduced incentive to exert effort.

The intuition for point (i) is more subtle. As the size of the tournament grows, competition to be promoted becomes more intense. Naive economic reasoning then suggests that effort should increase. This intuition fails to consider the effect of the combined noises on agents' effort choices. When choosing their effort, agent i does not compare their output with every other agent's. Rather, as explained above, they only take into account their best-performing competitor. As the number of competing agents increases, the combined noise associated with agent i 's best competitor becomes increasingly likely to take high values (more precisely, $F_\epsilon(\epsilon)^{n''-1}$ strictly first-order stochastically dominates $F_\epsilon(\epsilon)^{n'-1}$ for all $n'' > n' \geq 2$).

Better performance by the highest competitor alone, however, is not sufficient to discourage effort by agent $i \in \{1, \dots, n\}$. Indeed, agent i 's effort equates the marginal cost with the *marginal* gain in promotion probability. Their total winning probability has little effect on their effort choice. In turn, the marginal gain in promotion probability depends on the distribution of their own combined noise ϵ_i . Since $f_\epsilon(\cdot)$ is single-peaked and symmetric, agent i 's effort is more likely to be complemented by low (in absolute value) rather than large combined noise. Consequently, as the size of the tournament grows, the probability that their output is close to their highest competitor's decreases, as does the likelihood that a marginal increase in effort significantly increases their promotion probability.

The reasoning above thus indicates that an agent's performance decreases with the size of the tournament due to the combination of two effects. First, their highest competitor's performance increases. Second, more effort is less and less likely to significantly impact their promotion probability. Absent either of these two effects, the size of the tournament has no impact on an agent's effort choice.⁷ In particular, when ϵ_i is uniformly distributed so that all levels of combined noise are equally likely, the size of the tournament has no effect on agents' effort.⁸

Figure 1 illustrates these scenarios. We display three pdfs of the combined noise ϵ : the pdf for agent i (in red, dotted), the pdf for the best competitor in a tournament of size 4 (blue, dashed), and the pdf for the highest competitor in a tournament of size 10 (black, plain). In Figure 1a, the ϵ_j of tournament participants is normally distributed (mean 0, variance 0.25). In Figure 1b, the ϵ_j of participants is uniformly distributed (over $[-2, 2]$). For both distributions, agent i 's chances of winning are lower when there are 10 participants than when there are 4 participants in the tournament: The red curve crosses the blue dashed curve before the black curve and, as we noted, the best competitor's pdf when $n = 10$ is to the left of the best competitor's pdf when $n = 4$ (i.e., $F_\epsilon(\epsilon)^9$ strictly first-order stochastically dominates $F_\epsilon(\epsilon)^3$). Yet, agent i does not consider his total chances of winning when deciding their effort but the marginal change in winning probabilities. The increase in their chances of winning following greater effort is illustrated by the yellow filled zone

when $n = 4$ and the orange filled zone when $n = 10$ in Figure 1a,b. With a singled-peaked distribution like the normal distribution, we can visually see that increased effort has a smaller effect on i 's winning chances when $n = 10$ than when $n = 4$ (the orange zone is smaller than the yellow zone). Hence, agent i has less incentives to exert effort in a larger tournament. In contrast, with the uniform distribution, the effect of increased effort is the same for both tournament sizes (the yellow zone and the orange zone have the same area). In this case, the incentive to exert effort is independent of the tournament size.

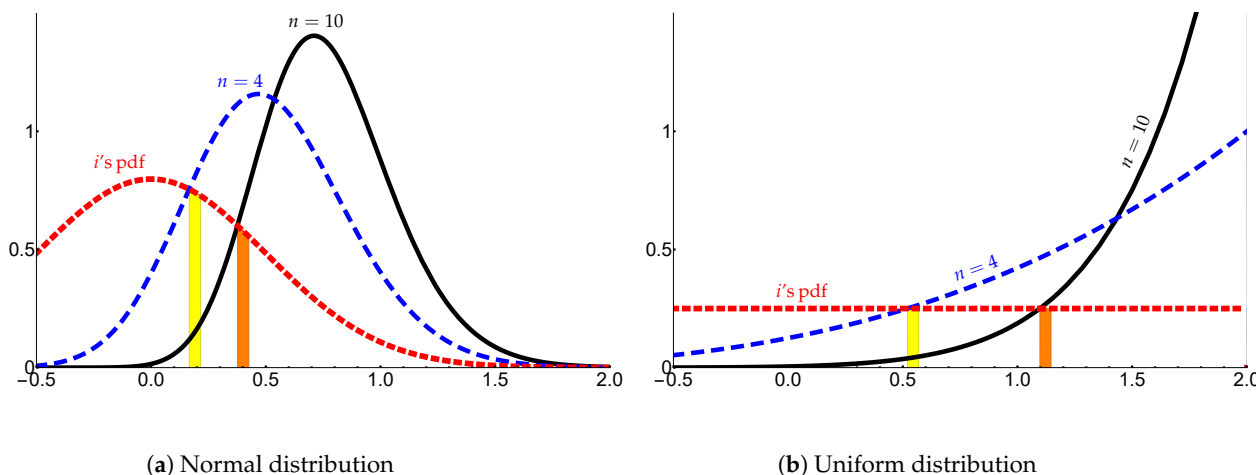


Figure 1. Comparing distributions. The graphs display the pdf of ϵ_i for agent i and the pdf of ϵ for the best competitor for tournament of size 4 (dashed blue line) and size 10 (plain black line). (a) ϵ_i for all participants is normally distributed with mean 0 and variance 0.25. (b) ϵ_i for all participants is uniformly distributed over $[-2, 2]$.

As Corollary 1 also shows, 2-agent and 3-agent tournaments yield the same equilibrium effort, as the negative effect of the highest competitor's increased performance is counterbalanced by a positive effect of lower variance, which increases the marginal value of effort.

Corollary 1. *The equilibrium effort satisfies $\bar{a}(n+1, \sigma_\epsilon, V) < \bar{a}(n, \sigma_\epsilon, V)$ if and only if $n > 2$ and there exists $\epsilon \in [0, \bar{\epsilon}]$ such that $f'_\epsilon(\epsilon) < 0$.*

Without imposing additional conditions on the shape of agents' cost function $c(\cdot)$, we cannot characterize further features of equilibrium effort. Since effort is determined by the marginal cost, any second-order prediction depends on the third derivative of the cost function. In the next corollary, we instead examine certain properties of the marginal benefit of effort, which we denote as $b(n, \sigma_\epsilon, V) := V \int_{-\bar{\epsilon}}^{\bar{\epsilon}} f_\epsilon(\epsilon) dF_\epsilon(\epsilon)^{n-1}$ (see the right-hand side of Equation (6)).

Corollary 2. *The marginal benefit of effort satisfies the following properties:*

- (i) *Convex in n (i.e., $b(n'' + k, \sigma_\epsilon, V) - b(n'', \sigma_\epsilon, V) \geq b(n' + k, \sigma_\epsilon, V) - b(n', \sigma_\epsilon, V)$ for $n'' > n' \geq 2$ and $k \geq 1$), strictly if $f'_\epsilon(\epsilon) > 0$ for some $\epsilon \in [0, \bar{\epsilon}]$;*
- (ii) *Strictly convex in σ_ϵ ;*
- (iii) *Neither concave nor convex in V .*

Point (iii) follows from the risk neutrality of the agent. Any increase in the prize linearly increases the benefit from winning the tournament and, in turn, the marginal benefit of effort. Point (ii) follows from the effect of higher variance on $f_\epsilon(\cdot)$. As σ_ϵ increases, the pdf of the combined noise flattens, and an increase in effort becomes marginally more likely

to affect the tournament outcome. This second-order effect attenuates the direct effect of variance identified above.

Point (i) indicates that the negative effect on incentives of increased tournament size on effort attenuates as the tournament grows larger. This arises because of the effect of adding one more agent on the highest performance (in expectation). The direct effect is that expected highest performance increases. However, this also implies, as a second-order effect, that higher and higher realizations of the combined noise are necessary to improve the highest performance. Since $f_\epsilon(\cdot)$ is single-peaked (i.e., low values of ϵ in absolute terms are more likely), such events are increasingly unlikely. Consequently, the expected highest performance increases at a decreasing rate with the number of participants. Since the marginal benefit of effort is inversely related to the highest performance of other agents, it is convex in n .

Figure 2 summarizes the results of this section. There, we display how the marginal benefit of effort ($V \int_{-\bar{\epsilon}}^{\bar{\epsilon}} f_\epsilon(\epsilon) dF_\epsilon(\epsilon)^{n-1}$) varies with the size of the tournament (left panel), the value of the prize V (middle panel), and the variance of the combined noise σ_ϵ^2 (right panel), assuming that both ability and luck are normally distributed. Figure 2a highlights that (i) the marginal benefit of effort (and thus the equilibrium effort itself) is the same for $n = 2$ and $n = 3$ tournaments, and (ii) the marginal benefit of effort is decreasing and convex nature in the number of participants. In Figure 2b, we see how the marginal benefit increases linearly with the value of the prize, whereas the marginal benefit of effort is decreasing and convex with the variance of $\epsilon = \theta + \eta$, as shown in Figure 2c. In these last two figures, we also observe how the marginal benefit of effort is lower for large than for small tournaments.

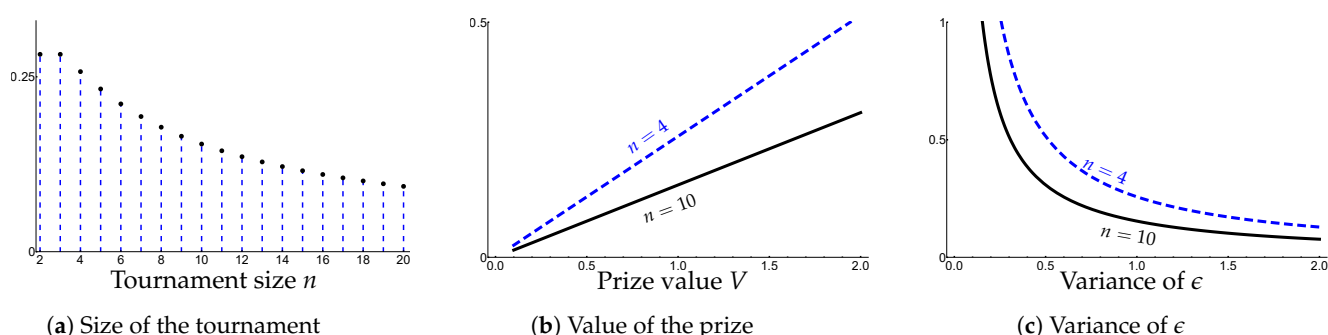


Figure 2. Comparative statics on the marginal benefit. The graphs display the marginal benefit as a function of the tournament size (a), the prize (b), and the variance of ϵ (c). (b,c) we display the marginal benefit for a tournament of size 4 (blue dashed curve) and of size 10 (dark plain curve). (a,b) σ_ϵ^2 takes value one. (a,c) V takes value one. In all panels, F_ϵ is normally distributed with mean zero.

Having established the general properties of the marginal benefit of effort, we can provide some predictions for the effect of the tournament parameters on effort. The second-order effect of V on effort depends on the sign of $c'''(\cdot)$. In turn, the equilibrium effort is convex for both σ_ϵ and the number of agents n whenever $c'''(\cdot)$ is not too large.

4. Tournament and Selection

As highlighted by Equation (3), the principal cares not only about the agents' effort (through $g(a_1, \dots, a_{\bar{n}})$) but also about the ability of the promoted agent. Because all agents exert the same level of effort $\bar{a}(\cdot)$, the principal simply promotes the best performer. Further, since she correctly anticipates effort, she can recover the combined noise of the best performer, which is denoted as $\epsilon^{\max}(n, \sigma_\theta, \sigma_\eta)$ (i.e., $\epsilon^{\max}(n, \sigma_\theta, \sigma_\eta) = \max_{i \in \{1, \dots, n\}} q_i - \bar{a}$). From the

principal's perspective, the promoted agent's expected ability—denoted as $\theta^{max}(n, \sigma_\theta, \sigma_\eta)$ —is then, for a given output,

$$\theta^{max}(n, \sigma_\theta, \sigma_\eta) := E_\theta(\theta | \epsilon = \epsilon^{max}(n, \sigma_\theta, \sigma_\eta)) \quad (7)$$

When deciding the tournament size, the principal can only condition her decision to select n agents on the expected realizations of agents' ability and luck. From an ex ante perspective, the important statistic for the principal is

$$\theta^W(n, \sigma_\theta, \sigma_\eta) := E_{\epsilon^{max}}(E_\theta(\theta | \epsilon = \epsilon^{max}(n, \sigma_\theta, \sigma_\eta))) \quad (8)$$

Our first result is quite intuitive. As the tournament size grows, the (ex ante) expected ability of the promoted agent $\theta^W(\cdot)$ increases. Given that effort does not affect the principal's promotion decision, adding agents can only increase the probability that a large combined noise is realized. While effort declines in a larger tournament, selection strictly improves.

Proposition 3. *The ex ante expected ability of the promoted agent increases with the size of the tournament.*

Additional conditions on the distribution of ability θ_i and luck η_i are required to establish further properties of $\theta^W(\cdot)$. As is common in the literature on career concerns, we impose that θ_i and η_i are both normally distributed (i.e., $\theta_i \sim \mathcal{N}(0, \sigma_\theta^2)$ and $\eta_i \sim \mathcal{N}(0, \sigma_\eta^2)$). Under these assumptions, $\theta^{max}(n, \sigma_\theta, \sigma_\eta)$ takes a very simple form: $\theta^{max}(n, \sigma_\theta, \sigma_\eta) = \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\eta^2} \epsilon^{max}(n, \sigma_\theta, \sigma_\eta)$. Denoting $\epsilon^W(n, \sigma_\theta, \sigma_\eta) = E_{\epsilon^{max}}(\epsilon^{max}(n, \sigma_\theta, \sigma_\eta))$, we obtain

$$\theta^W(n, \sigma_\theta, \sigma_\eta) := \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\eta^2} \epsilon^W(n, \sigma_\theta, \sigma_\eta) \quad (9)$$

Under the assumption of normality, we can provide a fuller picture of the impact of tournament characteristics on the ex ante ability of the promoted agent, which is defined as $\theta^W(\cdot)$.

Proposition 4. *Suppose θ_i and η_i are normally distributed. The ex ante expected ability of the promoted agent $\theta^W(n, \sigma_\theta, \sigma_\eta)$ is defined as follows:*

- (i) *Strictly increasing in σ_θ ;*
- (ii) *Strictly decreasing in σ_η .*
- (iii) *Strictly concave in n .*

As the variance in abilities increases, extreme realizations of ability become more likely. Since the principal only cares about the highest realization, not the average of realizations, at the time of promotion, this always benefits the principal. In turn, when the variance of the noise component η increases, high output is more likely to be due to luck than ability. This, therefore, tends to depress the expected ability of the promoted agent. Finally, the last point of the proposition indicates that the gain from adding contestants decreases with the size of the tournament. The logic is the same as for the convexity of effort in n . As the expected highest ability rises (i.e., $\theta^W(n, \sigma_\theta, \sigma_\eta)$ increases with n), it becomes less and less likely that the ability of an additional agent will surpass $\theta^W(n, \sigma_\theta, \sigma_\eta)$.

Figure 3 summarizes the results of this section. There, we display the expected ability of the tournament winner $\theta^W(\cdot)$ as a function of the size of the tournament (left panel), the variance of ability σ_θ^2 (middle panel), and the variance of the noise σ_η^2 (right panel), assuming that both ability and noise are normally distributed. As indicated above, the

selection improves with the size of the tournament, though in a concave fashion (Figure 3a). Increased variance in ability also benefits selection (Figure 3b), whereas more noise is detrimental to selection (Figure 3c). It is interesting (though not so surprising) to contrast this last point with the effect of σ_θ^2 and σ_η^2 on effort. Increasing randomness (from the perspective of an agent i) is always detrimental for effort, since it implies that an agent's effort is less likely to matter for their success. As such, we did not need to distinguish between the effect of greater variance for ability and noise above (see e.g., Figure 2c). In turn, for selection, only one type of randomness has negative consequences: pure luck. The principal actually cares for a high degree of randomness in ability, since it increases the expected ability of the tournament winner, especially when the tournament is of a large size (see how the black line increases faster than the blue dashed line in Figure 3b).

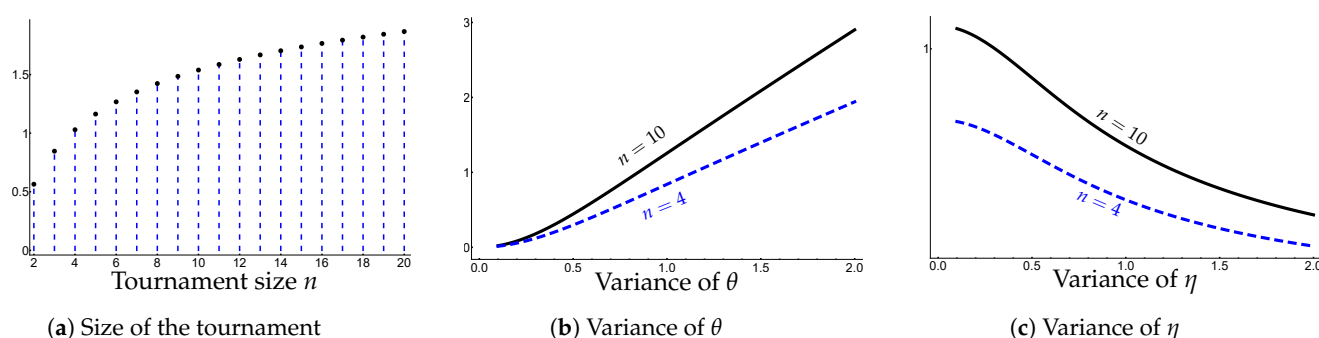


Figure 3. Comparative statics on the expected ability of tournament winner. The graphs display the expected ability of the tournament winner as a function of the tournament size (a), the variance of ability θ (b), and the variance of the noise η (c). (a,c) σ_θ^2 takes value $1/2$. (a,b) σ_η^2 takes value $1/2$. In all panels, ability and noise are normally distributed with mean zero.

5. Tournament Design

Having discussed both the effects of tournament size on effort (Section 3) and selection (Section 4), we can now consider how many agents the principal selects to compete in the promotion tournament. In what follows, we refer to ‘optimal tournament size’ as the number of contestants that maximize the principal’s expected utility given agents’ unique equilibrium pure strategies.

When choosing the number of competitors, the principal maximizes her ex ante expected utility, taking into account the effect of tournament size on effort and selection. That is, her maximization problem is

$$\max_{n \in \{1, \dots, \bar{n}\}} \omega g(\bar{a}(n, \cdot), \dots, \bar{a}(n, \cdot), 0, \dots, 0) + (1 - \omega) \theta^W(n, \cdot) - C_P(n) \quad (10)$$

Our first result establishes that tournaments of size 2 are generally suboptimal from the principal’s perspective.

Remark 1. The optimal tournament size is strictly less than 3 only if $C_P(3) - C_P(2) > (1 - \omega)(\theta^W(3, \sigma_\theta, \sigma_\eta) - \theta^W(2, \sigma_\theta, \sigma_\eta))$.

Proof. Proofs for this section are collected in Appendix C. \square

Note that this result does not require any restriction on $g(a_1, \dots, a_{\bar{n}})$ (besides the assumption that it is weakly increasing in its arguments) or on the distributions of θ_i and η_i . It is a direct consequence of Corollary 1 and Proposition 3. Increasing the size of the tournament from two to three agents has no effect on an agent’s effort. Since the principal always gains in terms of selection, she prefers not to restrict the set of contestants to two unless the cost of adding a third competitor is very large.

This result has important implications when one interprets elections as a tournament. Effort in this case can be understood as campaign spending meant to manipulate the voter's evaluation of candidates. In this case, $g(a_1, \dots, a_n)$ is constant in agents' output. Alternatively, effort may correspond to commitment to targeted spending, with $g(a_1, \dots, a_n) = \mathbb{I}_{\{i=\text{promoted}\}} a_i$, and $c(\cdot)$ as the cost of advertising campaign promises (see [Prato & Wolton, 2016](#)). Alternatively, candidates may actually invest in policy quality, with this investment depending on the candidates' ability and being imperfectly observable by voters (a setting adapted from [Hirsch & Shotts, 2015](#)). Finally, stretching the analogy slightly, one could think of effort in a more Downsian fashion. Under this perspective, the principal is the median voter, and effort is the cost of moving away from one's ideal point (in this mapping we briefly describe, voters need to have linear utility in policy position, whereas candidates have non-linear utility functions). Then, ability θ_i can be thought as a common-value valence shock. The policy promises and valence of each candidate need then to be imperfectly observable due to the noise of the campaign (captured by η_i). Remark 1 indicates that, independently of the interpretation of agents' output, two-candidate elections are generally inefficient from the electorate's point of view. It suggests that there may be some welfare gain for voters from encouraging at least one more candidate to participate in elections (as long as the marginal cost of increasing the number of candidates is relatively small).

To provide more details on the optimal tournament size, we return throughout the remainder of this section to the case of normally distributed ability (θ) and luck (η). We further impose that the agents' cost function is quadratic: $c(a) = \frac{a^2}{2\beta}$, with β sufficiently small so that the second-order conditions hold. This second assumption allows us to identify the direct effect of size on effort, not the effect mediated through the shape of the cost function.⁹

We focus on two natural situations. First, we suppose that the principal's preference for effort depends only on the output of a single agent; that is, $g(\bar{a}(n, \cdot), \dots, \bar{a}(n, \cdot), 0, \dots, 0) = \bar{a}(n, \cdot)$. Second, we assume that the principal cares about total effort (equivalently, total output): $g(\bar{a}(n, \cdot), \dots, \bar{a}(n, \cdot), 0, \dots, 0) = n\bar{a}(n, \cdot)$.

5.1. Optimal Tournament Size When the Principal Cares About Single Effort

This first situation subsumes several distinct objectives that a principal might have. It corresponds to a principal valuing average output (since agents play a symmetric strategy). This naturally occurs when the principal delegates policies to subordinate units and is concerned about all entities receiving similar treatment—either to avoid inconsistencies, as in the case of judicial decisions, or to ensure some common standards, as in the case of education policies. Such an objective is also consistent with the principal seeking to maximize minimum output. In domains where negative spillovers are present (e.g., environmental policy), a principal may want to minimize the damage caused by the worst performer. In the case of antiterrorism policies, a principal wants to strengthen the weakest link ([De Mesquita, 2007](#); [Powell, 2007](#)). Finally, the function $g(\bar{a}(n, \cdot), \dots, \bar{a}(n, \cdot), 0, \dots, 0) = \bar{a}(n, \cdot)$ is also consistent with a principal only taking into account the performance of the promoted agent (or, equivalently, the best performance). Examples of such objectives include elections (as discussed above), maximizing the quality of implemented policies ([Hirsch & Shotts, 2015](#)), or innovation.¹⁰

Under our conditions on the cost function, there exists a unique size maximizing the principal's expected welfare, which we denote as n_A^* . The next proposition describes how n_A^* varies with the model parameters.

Proposition 5. When $g(\bar{a}(n, \cdot), \dots, \bar{a}(n, \cdot), 0, \dots, 0) = \bar{a}(n, \cdot)$, Equation (10) has a unique maximum n_A^* with the following properties:

- (i) n_A^* is decreasing with V ;
- (ii) n_A^* is increasing with σ_θ .

Recall that the expected ability of the promoted agent does not depend on the tournament prize V , whereas an agent's effort is strictly *increasing* in V . Nonetheless, Proposition 5 indicates that the optimal size is decreasing in the value of promotion. To understand this result, observe that when choosing n_A^* , the principal compares the *marginal gain* from increasing the size of the tournament by one agent with the marginal cost. As highlighted in Proposition 2, effort is decreasing with size. Further, since effort is linear in V (under our assumption of quadratic cost of effort), as the value of the prize increases, so does the difference between $\bar{a}(n, \sigma_\theta, V)$ and $\bar{a}(n+1, \sigma_\theta, V)$ (see in Figure 2b how the gap between the marginal benefits of effort for tournament of size 4 and tournament of size 10 increases with V). Therefore, as V increases, it becomes more costly, in terms of reduced effort, to increase the size of the tournament by one agent. Hence, n_A^* is decreasing in V .

A similar logic explains why the optimal size is increasing in the variance of ability even though σ_θ tends to depress effort. An increase in variance reduces the loss in effort terms from increasing the size of the tournament (see Figure 2c for an illustration of this reduced difference). Further, greater variance in abilities not only increases the (expected) quality of the promoted agent, but it also enhances the selection gain from adding one agent, as the chance of getting a high-ability agent increases (see Figure 3b, where the distance between the expected ability of the winner of a tournament of size 10 and the expected ability of the winner of a tournament of size 4 is much larger for high than for low σ_θ).

5.2. Optimal Tournament Size When the Principal Cares About Total Effort

Situations when the principal values total effort are quite common. For example, a profit-maximizing firm seeks to maximize the performance of its semi-independent divisions competing in different markets. Similarly, an editor wants to increase the quality of reporting among all journalists. Totalitarian leaders aim to maximize the performance of all party members—whether regional heads (Li & Zhou, 2005; Lu & Landy, 2014; Maskin et al., 2000) or rank-and-file members (Montagnes & Wolton, 2019).

The next Lemma establishes that while individual effort is decreasing and convex in n , total effort is increasing and concave in n .

Lemma 1. *Total equilibrium effort $n\bar{a}(n, \sigma_\theta, V)$ is strictly increasing (i.e., $(n+1)\bar{a}(n+1, \sigma_\theta, V) > n\bar{a}(n, \sigma_\theta, V)$) and strictly concave (i.e., $(n''+1)\bar{a}(n''+1, \sigma_\theta, V) - n''\bar{a}(n'', \sigma_\theta, V) < (n'+1)\bar{a}(n'+1, \sigma_\theta, V) - n'\bar{a}(n', \sigma_\theta, V)$ for all $n'' > n'$).*

We can then use Lemma 1 to establish the properties of the optimal tournament size—denoted as n_T^* —when the principal cares about total effort.

Proposition 6. *When $g(\bar{a}(n, \cdot), \dots, \bar{a}(n, \cdot)) = n\bar{a}(n, \cdot)$, Equation (10) has a unique maximum n_T^* with the following properties:*

- (i) $n_T^* \geq n_A^*$;
- (ii) n_T^* is increasing with V ;
- (iii) n_T^* is decreasing with σ_η .

When the principal cares about individual effort, an increase in tournament size only improves selection. In contrast, when the principal values total effort, increasing the size of the tournament has a positive effect on the effort ($n\bar{a}(n, \cdot)$) and selection ($\theta^W(n, \cdot)$) components of her objective function. As a direct consequence, the principal has more

incentive to increase the size of the tournament in the latter case, and this is reflected in point (i) of Proposition 6.

Points (ii) and (iii) follow from similar reasoning as for Proposition 5. Observe, however, that since total effort is increasing with n , an increase in V tends to magnify the benefit of adding contestants. Consequently, n_T^* is increasing with V . The impact of σ_η , in turn, is due to its negative impact on effort (which attenuates the positive effect of increasing tournament size) and selection (as higher performances are more likely to be due to luck than ability).

Notice that Proposition 6 says nothing about the effect of an increase in the variance of abilities. This results from the conflicting effects of increased σ_θ on the marginal benefit of adding a participant. Greater variance in abilities, by depressing effort (Proposition 2), mitigates the gain in total effort, whereas by improving $\theta^W(n, \sigma_\theta, \sigma_\eta)$ (Proposition 4), it enhances the selection gain.

6. Discussion

Any hierarchical structure presupposes promotion—the selection of the agent(s) most able to perform at the higher level. In turn, promotion generates incentives akin to tournaments. Consequently, tournaments should primarily be understood as a selection mechanism rather than a scheme to promote effort.

Not surprisingly, a large body of work on tournaments in economics has long established that (sufficiently complete) contracts always generate higher output than tournaments (Green & Stokey, 1983; Nalebuff & Stiglitz, 1983; Prendergast, 1999). But this literature misses much of the story. As in elections (Fearon, 1999), effort in promotion-based hierarchies is a by-product of selection incentives, as competing agents attempt to manipulate the principal's evaluation of their ability.

Because promotions are intrinsically a selection mechanism, the principal cannot commit to any fixed promotion standards other than promoting the agent with the highest expected ability. Any setting that presupposes other promotion rules falls short of Fearon's (1999) critique.

Agents' efforts do not matter at the time of promotion, but they do affect the design of the tournament. Importantly, we cannot evaluate the effectiveness of a tournament without a proper understanding of the principal's objective. When she values total effort, tournaments tend to be large and generate low individual effort. When she values individual effort, tournaments are relatively smaller and generate relatively high effort. When the principal's objective is unobservable to researchers, the relationship between size and the value of promotion can help to recover it. Indeed, the correlation is positive if the principal values total effort.

We derived our results under additive separability of ability and effort. The qualitative story, however, carries over if output is instead multiplicative, $q_i = \theta_i a_i + \eta_i$, where each θ_i is drawn from the same distribution and remains unknown to the agent when choosing effort. In that setting, a higher realization of θ_i still makes effort more valuable, so equilibrium effort again hinges on the probability that an extra unit of effort shifts the promotion ranking. Although the algebra becomes less transparent—agents must reason about the product θa rather than a simple sum—the central comparative statics insight survives: once the tournament has more than two contestants, the expected output of the leading rival continues to rise with tournament size, preserving the accountability–selection tradeoff highlighted in the paper.

We also assume that the principal can perfectly control the number of tournament participants. This can easily be accomplished by an autocrat (who can include agents and exclude others, without too much concern for participation constraints). This may be more

difficult in democratic settings. We believe that democratic leaders have other tools at their disposal to obtain tournaments of optimal size. For example, they can manipulate rewards or costs to participate in a tournament.¹¹ Yet, the prevalence of two-candidate elections (or rather elections with only two viable candidates), despite their inefficiency for voters (Remark 1), suggests that the incentives of those who design tournaments—office-holders—may be distinct from the incentives of the principal within a tournament—the electorate. We leave the analysis of this question for future research.

Author Contributions: Conceptualisation: B.P.M. and J.J.; Formal analysis, B.P.M. and S.W.; Writing: B.P.M. and S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Properties of Agents' Effort

Throughout the Appendix, we assume that ϵ is distributed over $[-\bar{\epsilon}, \bar{\epsilon}]$ (with $\bar{\epsilon}$ not necessarily finite). Denote $\hat{F}_\epsilon(\cdot)$ and $\hat{f}_\epsilon(\cdot)$ as the associated standardized (resp.) cumulative distribution function (CDF) and probability density function (pdf). Since $f_\eta(\cdot)$ and $f_\theta(\cdot)$ are symmetric and unimodal around 0, and η and θ are independent, $f_\epsilon(\cdot)$ is symmetric and unimodal around 0. Consequently, $f'_\epsilon(\epsilon) = -f'_\epsilon(-\epsilon)$ over the relevant domain.

In this section, we establish properties of the agents' level of effort. In doing so, we prove Propositions 1 and 2.

Proof of Proposition 1. As explained in the text, agent i 's objective is

$$\max_{a_i} K_i(a_i) = VPr(Promoted|a_i) - c(a_i)$$

Denote \bar{a} as the (symmetric) effort of all other agents. The probability that i is promoted is as follows (recall that $\hat{\cdot}$ indicates the standardized distribution): $Pr(Promoted|a_i) = \int_{-\bar{\epsilon}}^{\bar{\epsilon}} \frac{1}{\sigma_\epsilon} \hat{F}_\epsilon\left(\frac{a_i - \bar{a} + \epsilon}{\sigma_\epsilon}\right)^{n-1} \hat{f}_\epsilon\left(\frac{\epsilon}{\sigma_\epsilon}\right) d\epsilon$.

The first order condition is

$$K'_i(a_i) = V \int_{-\bar{\epsilon}}^{\bar{\epsilon}} \frac{n-1}{\sigma_\epsilon^2} \hat{F}_\epsilon\left(\frac{a_i - \bar{a} + \epsilon}{\sigma_\epsilon}\right)^{n-2} \hat{f}_\epsilon\left(\frac{a_i - \bar{a} + \epsilon}{\sigma_\epsilon}\right) \hat{f}_\epsilon\left(\frac{\epsilon}{\sigma_\epsilon}\right) d\epsilon - c'(a_i) \leq 0 \quad (A1)$$

Imposing symmetry, we obtain that the agent's effort in the symmetric equilibrium— \bar{a} —is implicitly defined by

$$c'(\bar{a}) = V \int_{-\bar{\epsilon}}^{\bar{\epsilon}} f_\epsilon(\epsilon) dF_\epsilon(\epsilon)^{n-1} = \frac{V}{\sigma_\epsilon^2} \int_{-\bar{\epsilon}}^{\bar{\epsilon}} (n-1) \hat{F}\left(\frac{\epsilon}{\sigma_\epsilon}\right)^{n-2} \hat{f}\left(\frac{\epsilon}{\sigma_\epsilon}\right)^2 d\epsilon \quad (A2)$$

Accordingly, we concentrate on the symmetric equilibrium derived in Proposition X. Standard arguments imply that this equilibrium is in fact the only one satisfying the usual regularity and stability criteria; we therefore do not reproduce those arguments here.

In what follows, we look at the property of effort in the symmetric equilibrium. Denote the equilibrium effort $\bar{a}(n, \sigma_\epsilon, V)$ to be implicitly defined by Equation (A2). To establish Proposition 2, it is useful to denote the equilibrium marginal benefit of effort as follows:

$$b(n, \sigma_\epsilon, V) := \frac{V}{\sigma_\epsilon^2} \int_{-\bar{\epsilon}}^{\bar{\epsilon}} (n-1) \hat{F}\left(\frac{\epsilon}{\sigma_\epsilon}\right)^{n-2} \hat{f}\left(\frac{\epsilon}{\sigma_\epsilon}\right)^2 d\epsilon \quad (\text{A3})$$

□

Lemma A1. *The marginal benefit of effort is defined as follows:*

1. Linearly increasing in V ;
2. Decreasing and convex in the variance σ_ϵ ;
3. Decreasing and concave in the number of participants n for $n \geq 3$.

Proof. The proof of point 1 follows directly from an examination of Equation (A3) and is thus omitted.

Regarding points 2 and 3, first we use a change of indices ($z = \epsilon/\sigma_\epsilon$) to rewrite $b(\cdot)$ as

$$b(n, \sigma_\epsilon, V) := \frac{V}{\sigma_\epsilon} \int_{-\bar{\epsilon}/\sigma_\epsilon}^{\bar{\epsilon}/\sigma_\epsilon} (n-1) \hat{F}(z)^{n-2} \hat{f}(z)^2 dz \quad (\text{A4})$$

By integrating by parts, we then obtain

$$\begin{aligned} b(n, \sigma_\epsilon, V) &= \frac{V}{\sigma_\epsilon} \left[\hat{F}(z)^{n-1} \hat{f}(z) \right]_{-\bar{\epsilon}/\sigma_\epsilon}^{\bar{\epsilon}/\sigma_\epsilon} - \frac{V}{\sigma_\epsilon} \int_{-\bar{\epsilon}/\sigma_\epsilon}^{\bar{\epsilon}/\sigma_\epsilon} \hat{F}(z)^{n-1} \hat{f}'(z) dz \\ &= \frac{V}{\sigma_\epsilon} \hat{f}\left(\frac{\bar{\epsilon}}{\sigma_\epsilon}\right) - \frac{V}{\sigma_\epsilon} \int_0^{\bar{\epsilon}/\sigma_\epsilon} \hat{F}(z)^{n-1} \hat{f}'(z) dz - V \int_{-\bar{\epsilon}/\sigma_\epsilon}^0 \hat{F}(z)^{n-1} \hat{f}'(z) dz \\ &= \frac{V}{\sigma_\epsilon} \hat{f}\left(\frac{\bar{\epsilon}}{\sigma_\epsilon}\right) - \frac{V}{\sigma_\epsilon} \int_0^{\bar{\epsilon}/\sigma_\epsilon} \hat{F}(z)^{n-1} \hat{f}'(z) dz + V \int_{\bar{\epsilon}/\sigma_\epsilon}^0 \hat{F}(-z)^{n-1} \hat{f}'(-z) dz \\ &= \frac{V}{\sigma_\epsilon} \hat{f}\left(\frac{\bar{\epsilon}}{\sigma_\epsilon}\right) + \frac{V}{\sigma_\epsilon} \int_0^{\bar{\epsilon}/\sigma_\epsilon} (-\hat{f}'(z)) \left(\hat{F}(z)^{n-1} - (1 - \hat{F}(z))^{n-1} \right) dz \end{aligned} \quad (\text{A5})$$

For point 2, differentiating $b(\cdot)$ then yields

$$\begin{aligned} \frac{\partial b(n, \sigma_\epsilon, V)}{\partial \sigma_\epsilon} &= -\frac{b(n, \sigma_\epsilon, V)}{\sigma_\epsilon} - \frac{V \bar{\epsilon}}{\sigma_\epsilon^3} \hat{f}'\left(\frac{\bar{\epsilon}}{\sigma_\epsilon}\right) + \frac{V}{\sigma_\epsilon} \left(-\frac{\bar{\epsilon}}{\sigma_\epsilon^2} \right) \left(-\hat{f}'\left(\frac{\bar{\epsilon}}{\sigma_\epsilon}\right) \right) \\ &= -\frac{b(n, \sigma_\epsilon, V)}{\sigma_\epsilon} < 0 \end{aligned} \quad (\text{A6})$$

It can then be checked that $\frac{\partial^2 b(n, \sigma_\epsilon, V)}{(\partial \sigma_\epsilon)^2} > 0$.

For point 3, denote $B(n, k, \sigma_\epsilon, V) = b(n+k, \sigma_\epsilon, V) - b(n, \sigma_\epsilon, V)$.¹² First, notice that for all $n \geq 2$,

$$\begin{aligned} B(n, 1, \sigma_\epsilon, V) &= \frac{V}{\sigma_\epsilon} \int_0^{\bar{\epsilon}/\sigma_\epsilon} (-\hat{f}'(z)) \left(\left(\hat{F}(z)^n - (1 - \hat{F}(z))^n \right) - \left(\hat{F}(z)^{n-1} - (1 - \hat{F}(z))^{n-1} \right) \right) dz \\ &= -\frac{V}{\sigma_\epsilon} \int_0^{\bar{\epsilon}/\sigma_\epsilon} (-\hat{f}'(z)) \left(\hat{F}(z)^{n-1} (1 - \hat{F}(z)) - (1 - \hat{F}(z))^{n-1} \hat{F}(z) \right) dz \\ &= -\frac{V}{\sigma_\epsilon} \int_0^{\bar{\epsilon}/\sigma_\epsilon} (-\hat{f}'(z)) \hat{F}(z) (1 - \hat{F}(z)) \left(\hat{F}(z)^{n-2} - (1 - \hat{F}(z))^{n-2} \right) dz \end{aligned} \quad (\text{A7})$$

Clearly, $B(n, 1, \sigma_\epsilon, V) < 0$. Since $B(n, k, \sigma_\epsilon, V) = (b(n+1, \sigma_\epsilon, V) - b(n, \sigma_\epsilon, V)) + (b(n+2, \sigma_\epsilon, V) - b(n+1, \sigma_\epsilon, V)) + \dots + (b(n+k, \sigma_\epsilon, V) - b(n+k-1, \sigma_\epsilon, V))$, or equivalently, $B(n, k, \sigma_\epsilon, V) = \sum_{j=0}^{k-1} B(n+j, 1, \sigma_\epsilon, V)$, we obtain $B(n, k, \sigma_\epsilon, V) < 0$.

To prove the concavity argument, we use the following claim:

Claim A1. $B(\underline{n}, 1, \sigma_\epsilon, V) < B(\bar{n}, 1, \sigma_\epsilon, V)$ for all $\bar{n} > \underline{n} \geq 4$.

Proof. To prove the claim, define the function $Q(x, \alpha) = x^\alpha - (1-x)^\alpha$ for all $x \in [1/2, 1]$ and $\alpha \geq 2$. Notice that $Q(\hat{F}(z), n-2) = \hat{F}(z)^{n-2} - (1-\hat{F}(z))^{n-2}$ for all $z \in [0, \bar{\epsilon}/\sigma_\epsilon]$. It is thus sufficient to show that $Q(x, \alpha)$ is decreasing in α (strictly for some $x \in [1/2, 1]$) to prove the claim.

Since the function is C^∞ in both arguments for $x \in [1/2, 1]$ and $\alpha \geq 2$, we obtain

$$G(x) = \frac{\partial Q(x, \alpha)}{\partial \alpha} = \ln(x)x^\alpha - \ln(1-x)(1-x)^\alpha$$

Observe that $G(1/2) = G(1) = 0$. Further,

$$\begin{aligned} G'(x) &= x^{\alpha-1}(1 + \alpha \ln(x)) + (1-x)^{\alpha-1}(1 + \alpha \ln(1-x)) \\ G''(x) &= (\alpha-1) \left(x^{\alpha-2}(1 + \ln(x)) - (1-x)^{\alpha-2}(1 + \ln(1-x)) \right) + \alpha \left(x^{\alpha-2} - (1-x)^{\alpha-2} \right) \end{aligned}$$

Observe that $1 + \alpha \ln(x) \geq 0 \Leftrightarrow x \geq \exp(-1/\alpha)$ and $1 + \alpha \ln(1-x) \geq 0 \Leftrightarrow x \leq 1 - \exp(-1/\alpha)$. Notice that since $\alpha \geq 2$, $\exp(-1/\alpha) > 1/2$. This implies that $1 + \alpha \ln(1-x) < 0$ for all $x \geq 1/2$. Further, $G'(x) < 0$ for all $x \in [1/2, \exp(-1/\alpha)]$ and $G''(x) > 0$ for all $x \in [\exp(-1/\alpha), 1]$. Putting together all the properties of $G(\cdot)$, we obtain that $G(x) < 0$ for all $x \in (1/2, 1)$, which proves the claim. \square

Using the claim, for all $\bar{n} > \underline{n} \geq 4$ and all $k \geq 1$, we thus have $B(\underline{n}, k, \sigma_\epsilon, V) = \sum_{j=0}^{k-1} B(\underline{n} + j, 1, \sigma_\epsilon, V) < \sum_{j=0}^{k-1} B(\bar{n} + j, 1, \sigma_\epsilon, V) = B(\bar{n}, k, \sigma_\epsilon, V)$. To complete the proof, notice that

$$\begin{aligned} B(4, 1, \sigma_\epsilon, V) - B(3, 1, \sigma_\epsilon, V) &= \frac{V}{\sigma_\epsilon} \int_0^{\bar{\epsilon}/\sigma_\epsilon} (-\hat{f}'(z)) \hat{F}(z) (1 - \hat{F}(z)) \times \\ &\quad \left(\hat{F}(z)^2 - (1 - \hat{F}(z))^2 - (\hat{F}(z) - (1 - \hat{F}(z))) \right) dz = 0 \end{aligned}$$

Thus, point 3 holds for all $n \geq 3$. \square

Remark A1. The marginal benefit of effort is the same in tournaments with two and three participants.

Proof. See Equation (A7), with $n = 2$. \square

Remark A2. For $n \geq 3$, effort is constant in the number of players if and only if ϵ is uniformly distributed over $[-\bar{\epsilon}, \bar{\epsilon}]$.

Proof. If ϵ is uniformly distributed over $[-\bar{\epsilon}, \bar{\epsilon}]$, then $\hat{f}'_\epsilon(z) = 0$ for all $z \in [0, \bar{\epsilon}/\sigma_\epsilon]$ and $B(n, 1, \sigma_\epsilon, V) = 0$ for all n . If ϵ does not follow a uniform distribution, then there exists $z \in (0, \bar{\epsilon}/\sigma_\epsilon)$, such as $f'_\epsilon(z) > 0$. By Equation (A7), then $B(n, 1, \sigma_\epsilon, V) > 0$. \square

Proof of Proposition 2. The proof follows directly from Lemma A1 and the convexity of $c(\cdot)$. \square

Proof of Corollary 1. Observe that (ignoring arguments whenever possible)

$$c''(\bar{a}) \left(\frac{\partial \bar{a}(n, \sigma_\epsilon, V)}{\partial V} \right)^2 + c''(\bar{a}) \frac{\partial^2 \bar{a}(n, \sigma_\epsilon, V)}{\partial V^2} = \frac{\partial^2 b(n, \sigma_\epsilon, V)}{\partial V^2}$$

Since $\frac{\partial^2 b(n, \sigma_\epsilon, V)}{\partial V^2} = 0$ by Lemma A1 and $c'''(\cdot) \leq 0$ by assumption, the result holds.

A similar reasoning holds for σ_ϵ , noting that $\frac{\partial^2 b(n, \sigma_\epsilon, V)}{\partial V^2} < 0$.

For the concavity in n , denote $h(\cdot) := (c')^{-1}(\cdot)$. Notice that $h(\cdot)$ is increasing and convex, given the properties of $c'(\cdot)$. The effort is concave in n if and only if for all $\bar{n} > \underline{n}$ and all $k \geq 1$,

$$\begin{aligned} & \bar{a}(\bar{n} + k, \sigma_\epsilon, V) - \bar{a}(\bar{n}, \sigma_\epsilon, V) - (\bar{a}(\underline{n} + k, \sigma_\epsilon, V) - \bar{a}(\underline{n}, \sigma_\epsilon, V)) \geq 0 \\ \Leftrightarrow & h(b(\bar{n} + k, \sigma_\epsilon, V)) - h(b(\bar{n}, \sigma_\epsilon, V)) - (h(b(\underline{n} + k, \sigma_\epsilon, V)) - h(b(\underline{n}, \sigma_\epsilon, V))) \geq 0 \quad (\text{A8}) \end{aligned}$$

By Lemma A1, $B(\bar{n}, k, \sigma_\epsilon, V) > B(\underline{n}, k, \sigma_\epsilon, V)$. By the convexity of $h(\cdot)$, Equation (A8) holds.¹³ \square

Finally, we establish some additional properties of the marginal benefit of effort.

Lemma A2. *The marginal benefit of effort satisfies, for all $n \geq 2$ and $k \geq 1$, the following:*

- (i) $\frac{\partial^2 b(n, \sigma_\epsilon, V)}{\partial V \partial \sigma_\epsilon} < 0$;
- (ii) $\frac{\partial B(n, k, \sigma_\epsilon, V)}{\partial V} < 0$;
- (iii) $\frac{B(n, k, \sigma_\epsilon, V)}{\partial \sigma_\epsilon} > 0$;
- (iv) $(n + 1)b(n + 1, \sigma_\epsilon, V) - nb(n, \sigma_\epsilon, V) > 0$;
- (v) $(\bar{n} + 1)b(\bar{n} + 1, \sigma_\epsilon, V) - \bar{n}b(\bar{n}, \sigma_\epsilon, V) < (\underline{n} + 1)b(\underline{n} + 1, \sigma_\epsilon, V) - \underline{n}b(\underline{n}, \sigma_\epsilon, V), \bar{n} > \underline{n} \geq 2$.

Proof. The proof of point (i) follows directly from Equation (A6) and $\partial b(n, \sigma_\epsilon, V) / \partial V > 0$. The proof of point (ii) follows directly from examination of Equation (A7) and $B(n, k, \sigma_\epsilon, V) = \sum_{j=0}^{k-1} B(n + j, 1, \sigma_\epsilon, V)$.

To prove point (iii), observe that $\frac{B(n, 1, \sigma_\epsilon, V)}{\partial \sigma_\epsilon} = \frac{V}{\sigma_\epsilon^2} \int_0^{\bar{\epsilon}/\sigma_\epsilon} (-\hat{f}'(z)) \hat{F}(z) (1 - \hat{F}(z)) (\hat{F}(z)^{n-2} - (1 - \hat{F}(z))^{n-2}) dz > 0$ by Equation (A7), so $\frac{B(n, k, \sigma_\epsilon, V)}{\partial \sigma_\epsilon} < 0$.

To prove point (iv), we obtain, using Equation (A5) for $n \geq 2$,

$$\begin{aligned} (n + 1)b(n + 1, \sigma_\epsilon, V) - nb(n, \sigma_\epsilon, V) &= \frac{V}{\sigma_\epsilon} \int_0^{\bar{\epsilon}/\sigma_\epsilon} (-\hat{f}'(z)) \left[(n + 1) \left(\hat{F}(z)^n - (1 - \hat{F}(z))^n \right) \right. \\ &\quad \left. - n \left(\hat{F}(z)^{n-1} - (1 - \hat{F}(z))^{n-1} \right) \right] dz \\ &= \frac{V}{\sigma_\epsilon} \int_0^{\bar{\epsilon}/\sigma_\epsilon} (-\hat{f}'(z)) \left[\left(\hat{F}(z)^n - (1 - \hat{F}(z))^n \right) \right. \\ &\quad \left. - n \hat{F}(z) (1 - \hat{F}(z)) (\hat{F}(z)^{n-1} - (1 - \hat{F}(z))^{n-1}) \right] dz \quad (\text{A9}) \end{aligned}$$

Consider for $x \in [1/2, 1]$ and $n \geq 1$, $n \in \mathbb{N}$ the function

$$\begin{aligned} \mathcal{H}(x, n) &:= (n + 1)(x^n - (1 - x)^n) - n(x^{n-1} - (1 - x)^{n-1}) \\ &:= x^n - (1 - x)^n - nx(1 - x)(x^{n-1} - (1 - x)^{n-1}) \end{aligned}$$

Observe that $\mathcal{H}(1/2, n) = 0$ for all $n \geq 1$. We prove by induction that $\mathcal{H}(x, n) > 0$ for all $x \in (1/2, 1]$ and all $n \geq 1$. For $n = 1$, observe that $\mathcal{H}(x, 1) = 2x - 1 > 0$ for all $x > 1/2$. Suppose that the claim holds for $n = d \geq 1$. Denote $\mathcal{H}_x(x, n)$ as the (partial) derivative of $\mathcal{H}(\cdot)$ with respect to its x . Observe that

$$\begin{aligned} \mathcal{H}_x(x, d + 1) &= (d + 1)(x^d - (1 - x)^d) - (d + 1)dx(1 - x)(x^{d-1} - (1 - x)^{d-1}) \\ &\quad - (d + 1)(1 - 2x)(x^d - (1 - x)^d) \\ &= (d + 1)\mathcal{H}(x, d) + (d + 1)(2x - 1)(x^d - (1 - x)^d) > 0, \end{aligned}$$

where the inequality follows from $x > 1/2$ and $\mathcal{H}(x, d) > 0$ (by induction). Given $\mathcal{H}(1/2, d+1) = 0$, this implies that $\mathcal{H}(x, d+1) > 0$ for all $x \in (1/2, 1]$.

Observe that $\mathcal{H}(\hat{F}(z), n) = (\hat{F}(z)^n - (1 - \hat{F}(z))^n) - n\hat{F}(z)(1 - \hat{F}(z))(\hat{F}(z)^{n-1} - (1 - \hat{F}(z))^{n-1}) > 0$ for all $z > 0$, so $(n+1)b(n+1, \sigma_\epsilon, V) - nb(n, \sigma_\epsilon, V) > 0$ by Equation (A9).

For point (iii), notice that for all $n \geq 2$,

$$\begin{aligned} \mathcal{H}(x, n+1) - \mathcal{H}(x, n) &= (x^{n+1} - (1-x)^{n+1}) - (x^n - (1-x)^n) \\ &\quad - x(1-x) \left[(n+1)(x^n - (1-x)^n) - n(x^{n-1} - (1-x)^{n-1}) \right] \\ &= -x(1-x)(x^{n-1} - (1-x)^{n-1}) - x(1-x)\mathcal{H}(x, n) < 0 \end{aligned}$$

From Equation (A8), $(n+1)b(n+1, \sigma_\epsilon, V) - nb(n, \sigma_\epsilon, V) = \frac{V}{\sigma_\epsilon} \int_0^{\bar{\epsilon}/\sigma_\epsilon} (-\hat{f}'(z))\mathcal{H}(\hat{F}(z), n)dz$ for all $n \geq 2$. Using the properties of $\mathcal{H}(x, n)$, $(n+1)b(n+1, \sigma_\epsilon, V) - nb(n, \sigma_\epsilon, V) > (n+2)b(n+2, \sigma_\epsilon, V) - (n+1)b(n+1, \sigma_\epsilon, V)$ for all $n \geq 2$. The claim follows by iterating the inequality. \square

Proof of Corollary 2. This is directly determined from Lemma A2. \square

Appendix B. Properties of Tournament Winner

As explained in the text, the principal selects (in the symmetric equilibrium) the agent with the highest performance. Since the principal correctly anticipates all agents' efforts, this is equivalent to selecting the highest realization of the uncertainty component $\epsilon = \theta + \eta$. Denote $\epsilon^W(n, \sigma_\theta, \sigma_\eta) = E(\max_i \epsilon_i)$. We thus have $\epsilon^W(n, \sigma_\theta, \sigma_\eta) = \int_{-\infty}^{\infty} \epsilon dF_\epsilon(\epsilon)^n$. Using the standardized pdf and CDF, this is equivalent to $\epsilon^W(n, \sigma_\theta, \sigma_\eta) = \int_{-\infty}^{\infty} n \frac{\epsilon}{\sigma_\epsilon} \hat{F}_\epsilon(\frac{\epsilon}{\sigma_\epsilon})^{n-1} \hat{f}_\epsilon(\frac{\epsilon}{\sigma_\epsilon}) d\epsilon = \sigma_\epsilon \int_{-\infty}^{\infty} n z \hat{F}_\epsilon(z)^{n-1} \hat{f}_\epsilon(z) dz$ (the second equality follows from change of indices). Hence, we have

$$\epsilon^W(n, \sigma_\theta, \sigma_\eta) = \sigma_\epsilon \int_{-\infty}^{\infty} z d\hat{F}_\epsilon(z)^n \quad (\text{A10})$$

The expected competence of the tournament winner (using $\epsilon = \theta + \eta$) is then

$$\theta^W(n, \sigma_\epsilon, \sigma_\eta) = E(\theta | \theta + \eta = \epsilon^W(n, \sigma_\theta, \sigma_\eta)) \quad (\text{A11})$$

Proof of Proposition 3. To prove the proposition, we just need to show that $\epsilon^W(n, \sigma_\theta, \sigma_\eta)$ is increasing with n (since θ and $\theta + \eta$ are positively correlated). Notice that for all $\epsilon \in \mathbb{R}$, $\hat{F}(\epsilon)^n \geq \hat{F}(\epsilon)^{n+1}$ (with strict inequality for all finite ϵ). Hence, $\hat{F}(\cdot)^{n+1}$ strictly first-order stochastically dominates $\hat{F}(\cdot)^n$. Hence, $\epsilon^W(\bar{n}, \sigma_\theta, \sigma_\eta) > \epsilon^W(\underline{n}, \sigma_\theta, \sigma_\eta)$ for all $\bar{n} > \underline{n}$.

To prove concavity, it is useful to denote $S(n, k) = \epsilon^W(n+k, \sigma_\theta, \sigma_\eta) - \epsilon^W(n, \sigma_\theta, \sigma_\eta)$. Using Equation (A10),

$$\begin{aligned} S(n, k) &= \sigma_\epsilon \int_{-\infty}^{+\infty} x dF(x)^{n+k} - \sigma_\epsilon \int_{-\infty}^{+\infty} x dF(x)^n \\ &= \sigma_\epsilon \int_{-\infty}^{+\infty} x d(F(x)^{n+k} - F(x)^n) \\ &= \sigma_\epsilon [x(F(x)^{n+k} - F(x)^n)]|_{-\infty}^{+\infty} - \sigma_\epsilon \int_{-\infty}^{+\infty} (F(x)^{n+k} - F(x)^n) dx \\ &= \sigma_\epsilon \int_{-\infty}^{+\infty} (F(x)^n - F(x)^{n+k}) dx \\ &= \sigma_\epsilon \int_{-\infty}^{+\infty} F(x)^n (1 - F(x)^k) dx \end{aligned} \quad (\text{A12})$$

The third equality follows from integration by parts. This implies that

$$S(\bar{n}, k) - S(\underline{n}, k) = \sigma_\epsilon \int_{-\infty}^{+\infty} [F(x)^{\bar{n}} - F(x)^{\underline{n}}](1 - F(x)^k) dx \quad (\text{A13})$$

For all $\bar{n} > \underline{n}$ and $k \geq 1$, $S(\bar{n}, k) - S(\underline{n}, k) < 0$. \square

Imposing the normal distributions on θ and η , $\theta^W(\cdot)$ can be expressed as

$$\begin{aligned} \theta^W(n, \sigma_\epsilon, \sigma_\eta) &= \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\eta^2} \epsilon^W(n, \sigma_\epsilon, \sigma_\eta) \\ \theta^W(n, \sigma_\epsilon, \sigma_\eta) &= \frac{\sigma_\theta^2}{\sqrt{\sigma_\theta^2 + \sigma_\eta^2}} \int_{-\infty}^{\infty} z d\hat{F}_\epsilon(z)^n, \end{aligned} \quad (\text{A14})$$

where the second line uses $\sigma_\epsilon = \sqrt{\sigma_\theta^2 + \sigma_\eta^2}$.

Proof of Proposition 4. Points 1 and 2 follow from taking the derivative of $\frac{\sigma_\theta^2}{\sqrt{\sigma_\theta^2 + \sigma_\eta^2}}$ with respect to σ_θ^2 and σ_η^2 , respectively. \square

We now prove some additional properties of selection in tournament. To do so, it is useful to define

$$\Delta^W(n, k, \sigma_\epsilon, \sigma_\eta) := \theta^W(n + k, \sigma_\epsilon, \sigma_\eta) - \theta^W(n, \sigma_\epsilon, \sigma_\eta) \quad (\text{A15})$$

Lemma A3. For all $n \geq 2$ and all $k \geq 1$, $\Delta^W(n, k, \sigma_\epsilon, \sigma_\eta)$ is increasing with σ_θ and decreasing with σ_η .

For all $\bar{n} > \underline{n}$ and $k \geq 1$, $\Delta^W(\bar{n}, k, \sigma_\epsilon, \sigma_\eta) - \Delta^W(\underline{n}, k, \sigma_\epsilon, \sigma_\eta)$ is decreasing with σ_θ and increasing with σ_η .

Proof. Observe that, using Equations (A11) and (A12),

$$\Delta^W(n, k, \sigma_\epsilon, \sigma_\eta) = \frac{\sigma_\theta^2}{\sqrt{\sigma_\theta^2 + \sigma_\eta^2}} S(n, k),$$

with $S(n, k) > 0$ defined in the proof of Proposition 3. The claim then follows from a similar reasoning as in the proof of Proposition 4.

The comparative statics are reversed for $\Delta^W(\bar{n}, k, \sigma_\epsilon, \sigma_\eta) - \Delta^W(\underline{n}, k, \sigma_\epsilon, \sigma_\eta)$, since $S(\bar{n}, k) - S(\underline{n}, k) < 0$ by Proposition 3. \square

Appendix C. Optimal Tournament Design

Throughout, we use the notation $h(\cdot) := (c')^{-1}(\cdot)$. We provide general conditions for our comparative results to hold before imposing the assumption $c(a) = \frac{a^2}{\beta}$. Further, we assume that θ and η are normally distributed with respective variances σ_θ^2 and σ_η^2 .

The existence of n_A^* and n_T^* is guaranteed by the properties of $C_P(n)$ detailed in the text and the bounded nature of the choice set available to the principal (i.e., $n \in \{2, \dots, n^{\max}\}$).

Proof of Remark 1. By Remark A1, effort is the same in tournaments with two or three competitors. When the principal cares about average effort, the marginal gain from increasing the tournament from two to three players is thus $\theta^W(3, \sigma_\theta, \sigma_\eta) - \theta^W(2, \sigma_\theta, \sigma_\eta)$. When the principal cares about total effort, the marginal gain from increasing the tournament from two to three players is thus $h(b(3, \sigma_\epsilon, V)) + \theta^W(3, \sigma_\theta, \sigma_\eta) - \theta^W(2, \sigma_\theta, \sigma_\eta)$. In both cases, the principal is worse off with a tournament with more than two competitors only if the

condition stated in the text of the remark holds (note that the condition is also sufficient for average-effort tournament). \square

Proof of Proposition 5. By Milgrom and Shannon (1994), the solution is increasing (decreasing) in a parameter if the objective function exhibits the single-crossing property which is implied by showing increasing (decreasing) differences in the parameter.

Denote the benefit of increasing the size of a tournament as

$$K_A(n, \sigma_\theta, \sigma_\eta, V) = \omega a(n, \sigma_\eta, \sigma_\theta, V) + (1 - \omega) \theta^W(n, \sigma_\eta, \sigma_\theta) \quad (\text{A16})$$

Further, let, for $\bar{n} > \underline{n} \geq 2$,

$$\mathcal{K}_A(\sigma_\theta, \sigma_\eta, V) = K_A(\bar{n}, \sigma_\theta, \sigma_\eta, V) - K_A(\underline{n}, \sigma_\theta, \sigma_\eta, V). \quad (\text{A17})$$

The objective function exhibits increasing (decreasing) differences in σ_θ , σ_η , or V whenever $\mathcal{K}_A(\sigma_\theta, \sigma_\eta, V)$ is increasing (decreasing) in the variable.

To prove point (i), observe that, using Equations (A6) and (A14), and $\sigma_\epsilon = \sqrt{\sigma_\theta^2 + \sigma_\eta^2}$,

$$\begin{aligned} \frac{\partial \mathcal{K}_A(\sigma_\theta, \sigma_\eta, V)}{\partial \sigma_\theta^2} &= -\omega \frac{1}{2\sigma_\epsilon^2} (b(\bar{n}, \sigma_\epsilon, V) h'(b(\bar{n}, \sigma_\epsilon, V)) - b(\underline{n}, \sigma_\epsilon, V) h'(b(\underline{n}, \sigma_\epsilon, V))) \\ &\quad + (1 - \omega) \frac{\frac{\sigma_\theta^2}{2} + \sigma_\eta^2}{\sigma_\theta^2 \sigma_\epsilon} (\theta^W(\bar{n}, \sigma_\theta, \sigma_\eta) - \theta^W(\underline{n}, \sigma_\theta, \sigma_\eta)) \end{aligned}$$

Imposing the quadratic cost assumption (i.e., $h'(a) = \beta$), $\frac{\partial \mathcal{K}_A(\sigma_\theta, \sigma_\eta, V)}{\partial \sigma_\theta^2} > 0$ by Corollary 2 and Proposition 3.¹⁴

The proof of points (ii) and (iii) follows from a similar reasoning in noting that $\theta^W(\cdot)$ does not depend on V . \square

Proof of Proposition 6. The proof again uses the increasing (decreasing) differences property of the objective function. Denote the benefit of a tournament as

$$K_T(n, \sigma_\theta, \sigma_\eta, V) = \omega n a(n, \sigma_\eta, \sigma_\theta, V) + (1 - \omega) \theta^W(n, \sigma_\eta, \sigma_\theta) \quad (\text{A18})$$

Further, let, for $\bar{n} > \underline{n} \geq 2$,

$$\mathcal{K}_T(\sigma_\theta, \sigma_\eta, V) = K_T(\bar{n}, \sigma_\theta, \sigma_\eta, V) - K_T(\underline{n}, \sigma_\theta, \sigma_\eta, V). \quad (\text{A19})$$

To prove point (i), observe that, using Equations (A9) and (A14), and $\sigma_\epsilon = \sqrt{\sigma_\theta^2 + \sigma_\eta^2}$,

$$\begin{aligned} \frac{\partial \mathcal{K}_T(\sigma_\theta, \sigma_\eta, V)}{\partial \sigma_\eta^2} &= -\omega \frac{1}{2\sigma_\epsilon^2} (\bar{n} b(\bar{n}, \sigma_\epsilon, V) h'(b(\bar{n}, \sigma_\epsilon, V)) - \underline{n} b(\underline{n}, \sigma_\epsilon, V) h'(b(\underline{n}, \sigma_\epsilon, V))) \\ &\quad - (1 - \omega) \frac{1}{\sigma_\epsilon^2} (\theta^W(\bar{n}, \sigma_\theta, \sigma_\eta) - \theta^W(\underline{n}, \sigma_\theta, \sigma_\eta)) \end{aligned}$$

Imposing the quadratic cost assumption (i.e., $h'(a) = \beta$), $\frac{\partial \mathcal{K}_T(\sigma_\theta, \sigma_\eta, V)}{\partial \sigma_\eta^2} < 0$ by Lemma A1 and Proposition 3.¹⁵

The proof of point (ii) follows from a similar reasoning noting that $\theta^W(\cdot)$ does not depend on V . \square

Appendix D. Tournaments with Proportional Prize Price

In this last Appendix, we study some properties of tournament when the prize increases with the size of the tournament. Throughout, we focus on the case when luck and ability are both normally distributed ($\tilde{\eta} \in \mathcal{N}(0, \sigma_\eta^2)$ and $\tilde{\theta} \sim \mathcal{N}(0, \sigma_\theta^2)$). Recall that $\epsilon = \theta + \eta$.

Remark A3. Consider a proportional prize increase in a tournament, with a prize equal to nv . As the size of the tournament increases, the effort (a_i) of each agent increases.

Proof of Remark A3. We want to show that effort is increasing in n .

The new equilibrium effort for agent i is defined by

$$c'(a_i) = nv \int f_\epsilon(x) dF_\epsilon(x)^{n-1}. \quad (\text{A20})$$

Adapting from Proposition 2, it is sufficient to show that $b(n+1, \sigma_\epsilon, nv) > b(n, \sigma_\epsilon, nv)$ to prove increasing effort. By integration-by-parts, $b(n, \sigma_\epsilon, nv) = nv(f_\epsilon(x)F_\epsilon(x)^{n-1})|_{-\infty}^{+\infty} - nv \int_{-\infty}^{+\infty} F_\epsilon(x)^{n-1} f'_\epsilon(x) dx$, which yields $b(n, \sigma_\epsilon, nv) = -v \int_{-\infty}^{+\infty} nF_\epsilon(x)^{n-1} f'_\epsilon(x) dx$. We exploit the integration-by-parts technique again, which yields

$$b(n, 1, v) = -v \int_{-\infty}^{+\infty} \frac{f'_\epsilon(x)}{f_\epsilon(x)} dF_\epsilon(x)^n \quad (\text{A21})$$

$$= -v \left(\frac{f'_\epsilon(x)}{f_\epsilon(x)} F_\epsilon(x)^n \right) \Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} F_\epsilon(x)^n d \frac{f'_\epsilon(x)}{f_\epsilon(x)} \quad (\text{A22})$$

We can then write the difference in effort between n - and $n+1$ -person tournament as follows:

$$\begin{aligned} & b(n+1, 1, v) - b(n, 1, v) \\ &= \frac{f'_\epsilon(x)}{f_\epsilon(x)} [F_\epsilon(x)^n - F_\epsilon(x)^{n+1}] \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} F_\epsilon(x)^{n+1} d \frac{f'_\epsilon(x)}{f_\epsilon(x)} - \int_{-\infty}^{+\infty} F_\epsilon(x)^n d \frac{f'_\epsilon(x)}{f_\epsilon(x)} \\ &= \int_{-\infty}^{+\infty} (F_\epsilon(x)^{n+1} - F_\epsilon(x)^n) \frac{f''_\epsilon(x) f_\epsilon(x) - f'_\epsilon(x)^2}{f_\epsilon(x)^2} dx \\ &= \int_{-\infty}^{+\infty} (F_\epsilon(x)^{n+1} - F_\epsilon(x)^n) \left(-\frac{1}{\sigma_\epsilon^2} \right) dx \\ &= -\frac{1}{\sigma_\epsilon^2} \int_{-\infty}^{+\infty} (F_\epsilon(x)^{n+1} - F_\epsilon(x)^n) dx \\ &= -\frac{1}{\sigma_\epsilon^2} \int_0^1 (F_\epsilon(x)^{n+1} - F_\epsilon(x)^n) + [(1 - F_\epsilon(x))^{n+1} - (1 - F_\epsilon(x))^n] dx \end{aligned} \quad (\text{A23})$$

To see how we reach (A23), notice that in the second equality, we have $\frac{f'_\epsilon(x)}{f_\epsilon(x)} [F_\epsilon(x)^n - F_\epsilon(x)^{n+1}] \Big|_{-\infty}^{+\infty} = 0$ given the assumption of normal distributions. Also note that for any normal distribution with mean 0 and variance σ^2 , we have $\frac{f''_\epsilon(x) f_\epsilon(x) - f'_\epsilon(x)^2}{f_\epsilon(x)^2} = \frac{(-\frac{\sigma^2 + x^2}{\sigma^4}) f_\epsilon(x)^2 - \frac{x^2}{\sigma^4} f_\epsilon(x)^2}{f_\epsilon(x)^2} = \frac{-\frac{\sigma^2}{\sigma^4} f_\epsilon(x)^2}{f_\epsilon(x)^2} = -\frac{1}{\sigma^2} < 0$.

To finish the proof, it is enough to note that the first and second bracketed terms in line (A23) are negative, and so the whole expression is positive as required. \square

Remark A4. There exists \bar{v} such that, for prizes $v > \bar{v}$, the ex ante utility of agents is decreasing.

Proof of Remark A4. Consider the agent's FOC as

$$c'(a_i) = V \int F_\epsilon^{n-1}(x) f_\epsilon(x) dx.$$

Let $V \int F_\epsilon^{n-1}(x) f_\epsilon(x) = V\bar{P}$, and then the equilibrium effort a^* is

$$a^* = (c'^{-1})(V\bar{P}).$$

The ex ante total cost for the agent is thus $c((c'^{-1})(V\bar{P}))$, which follows as

$$\frac{\partial c((c'^{-1})(V\bar{P}))}{\partial V} = \frac{c'(\cdot)\bar{P}}{c''(c^{-1}(V\bar{P}))} > 0 \quad (\text{A24})$$

And

$$\frac{\partial^2 c((c'^{-1})(V\bar{P}))}{\partial^2 V} = \frac{(c'')^2 - c'c'''}{(c'')^2} \quad (\text{A25})$$

If we assume that $(c'')^2 - c'c''' > 0$, then the second order of the total cost function (Equation (A25)) is strictly convex; hence, the total benefit is concave. There would thus exist a \bar{V} beyond which the total benefit starts declining. \square

Notes

- ¹ Note that throughout the paper, we use a *forward-looking* notion of accountability: An institution is more accountable when it gives agents stronger incentives to exert costly effort *before* outcomes are realized.
- ² In our model, effort is always beneficial for the principal. This does not have to be the case. For example, [Hvide and Kristiansen \(2003\)](#) and [Fang and Noe \(2022\)](#) highlight that tournaments can lead to excessive risk-taking from the principal's perspective.
- ³ This assumption is consistent with evidence on politicians' skills (albeit in Sweden) collected in [Dal Bó et al. \(2017\)](#).
- ⁴ Our assumption that output is a linear sum of effort, ability, and noise is closest to the model of [Lazear and Rosen \(1981\)](#), but see [Nalebuff and Stiglitz \(1983\)](#) for a model where effort and shocks are not linearly separable.
- ⁵ Equivalently, the principal may care about agent output rather than their effort. This generally does not affect our results (recall agents not considered for promotion produce no output) but only complicates the exposition (see note 10 for more details). We define $g(\cdot)$ over \bar{n} and not n so that the function is well defined for all tournament sizes.
- ⁶ Absent this assumption on variances, selection plays a limited role. The competition between agents resembles an all-pay contest with equilibrium supported by mixed strategies over a support that includes zero effort. We implicitly assume that the second-order conditions for the maximization problems are well defined and hold in what follows.
- ⁷ Furthermore, as the proof of Proposition 2 shows, if the single-peaked assumption is reversed and $f'_\epsilon(\epsilon) > 0$ for all $\epsilon > 0$, the result is reversed. Assuming the existence of a symmetric equilibrium, effort is increasing with the size of the tournament.
- ⁸ Hence, the null effect of tournament size on effort found in [Orrison et al. \(2004\)](#) is a knife-edge result that relies on their distributional assumption.
- ⁹ See the discussion around Corollary 2. All our results hold as long as $|c'''(a)| < \bar{c}$ for all $a \geq 0$, with \bar{c} strictly positive but not too large.
- ¹⁰ If the principal cares about agents' outputs rather than effort, her objective function becomes $\omega E_{\bar{\epsilon}} g(\bar{a}(n, \cdot) + \epsilon_1, \dots, \bar{a}(n, \cdot) + \epsilon_n, 0, \dots, 0) + (1 - \omega)\theta^W(n, \cdot) - C_P(n)$, with $\bar{\epsilon} = (\epsilon_1, \dots, \epsilon_n)$. In this case, when the principal seeks to maximize the lowest (highest) performance, and the objective function reduces—using the symmetry—to $\omega \bar{a}(n, \cdot) + (1 - \omega)\theta^W(n, \cdot) - \omega \epsilon^W(n, \cdot)$ ($\omega \bar{a}(n, \cdot) + (1 - \omega)\theta^W(n, \cdot) + \omega \epsilon^W(n, \cdot)$). All our results remain unchanged for the case of maximizing the best performer and require ω sufficiently small for the case of maximizing the lowest performer. Focusing on agents' effort allows us to keep the objective function the same in all three cases described in the text.
- ¹¹ An alternative approach would be to vary the degree of centralization in a given state. Greater centralization can be understood as greater tournament. Greater centralization would, however, also increase the value of the prize. Hence, it is different from simply varying the reward of the tournament winner or the cost of participation.

- ¹² In the text, we consider the slightly different form of differences $b(n'', \sigma_\epsilon, V) - b(n', \sigma_\epsilon, V)$, where $n'' > n' \geq 2$. Considering fixed differences of size k allows for clearer proofs.
- ¹³ To see this, assume that n is a continuous variable (observe that $b(\cdot)$ is well defined for n continuous), and note that Equation (A8) is equivalent to $\int_0^k \int_{\underline{n}}^{\bar{n}} \frac{\partial^2 b(n, \sigma_\epsilon, V)}{\partial n^2} h''(b(n, \sigma_\epsilon, V)) dn dn$, with $\frac{\partial^2 b(n, \sigma_\epsilon, V)}{\partial n^2} > 0$ and $h''(\cdot) \geq 0$.
- ¹⁴ Observe that the result still holds true as long as $h''(a) < \bar{h}$ for all $a \geq 0$ and \bar{h} is strictly positive.
- ¹⁵ Observe that the result still holds true as long as $h''(a) > -\underline{h}$ for all $a \geq 0$ and \underline{h} is strictly positive.

References

- Ashworth, S. (2005). Reputational dynamics and political careers. *Journal of Law, Economics, and Organization*, 21(2), 441–466. [\[CrossRef\]](#)
- Ashworth, S. (2012). Electoral accountability: Recent theoretical and empirical work. *Annual Review of Political Science*, 15, 183–201. [\[CrossRef\]](#)
- Cameron, C. M., de Figueiredo, J. M., & Lewis, D. E. (2013). *Public sector personnel economics: Wages, promotions, and the competence-control trade-off*. National Bureau of Economic Research.
- Clark, D. J., & Riis, C. (2001). Rank-order tournaments and selection. *Journal of Economics*, 73, 167–191. [\[CrossRef\]](#)
- Connelly, B. L., Tihanyi, L., Crook, T. R., & Gangloff, K. A. (2014). Tournament theory: Thirty years of contests and competitions. *Journal of Management*, 40(1), 16–47. [\[CrossRef\]](#)
- Dal Bó, E., Finan, F., Folke, O., Persson, T., & Rickne, J. (2017). Who becomes a politician? *The Quarterly Journal of Economics*, 132(4), 1877–1914. [\[CrossRef\]](#)
- De Mesquita, E. B. (2007). Politics and the suboptimal provision of counterterrorism. *International Organization*, 61(01), 9–36. [\[CrossRef\]](#)
- Eggers, A. C., & Spirling, A. (2016). Party cohesion in Westminster systems: Inducements, replacement and discipline in the house of commons, 1836–1910. *British Journal of Political Science*, 46(03), 567–589. [\[CrossRef\]](#)
- Epstein, L., Knight, J., & Martin, A. D. (2003). The norm of prior judicial experience and its consequences for career diversity on the US Supreme Court. *California Law Review*, 91, 903–965. [\[CrossRef\]](#)
- Fang, D., & Noe, T. (2022). Less competition, more meritocracy? *Journal of Labor Economics*, 40(3), 669–701. [\[CrossRef\]](#)
- Fearon, J. D. (1999). Electoral accountability and the control of politicians: Selecting good types versus sanctioning poor performance. *Democracy, Accountability, and Representation*, 55, 61.
- Fong, C., & McCrain, J. (2024). A tournament theory of congressional committee leadership. *Public Choice*, 202, 193–215. [\[CrossRef\]](#)
- Ghosh, S., & Waldman, M. (2010). Standard promotion practices versus up-or-out contracts. *The RAND Journal of Economics*, 41(2), 301–325. [\[CrossRef\]](#)
- Green, J. R., & Stokey, N. L. (1983). A comparison of tournaments and contracts. *The Journal of Political Economy*, 91(3), 349–364. [\[CrossRef\]](#)
- Groseclose, T., & Stewart, C., III. (1998). The value of committee seats in the House, 1947–91. *American Journal of Political Science*, 42, 453–474. [\[CrossRef\]](#)
- Gürtler, M., & Gürtler, O. (2015). The optimality of heterogeneous tournaments. *Journal of Labor Economics*, 33(4), 1007–1042. [\[CrossRef\]](#)
- Hirsch, A. V., & Shotts, K. W. (2015). Competitive policy development. *The American Economic Review*, 105(4), 1646–1664. [\[CrossRef\]](#)
- Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies*, 66(1), 169–182. [\[CrossRef\]](#)
- Höchtel, W., Kerschbamer, R., Stracke, R., & Sunde, U. (n.d.). Incentives vs. selection in promotion tournaments: Can a designer kill two birds with one stone? *SSRN Electronic Journal*. [\[CrossRef\]](#)
- Hvide, H. K., & Kristiansen, E. G. (2003). Risk taking in selection contests. *Games and Economic Behavior*, 42(1), 172–179. [\[CrossRef\]](#)
- Krishna, V., & Morgan, J. (1998). The winner-take-all principle in small tournaments. *Advances in Applied Microeconomics*, 7, 61–74.
- Lazear, E. P., & Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *The Journal of Political Economy*, 89(5), 841–864. [\[CrossRef\]](#)
- Levy, J., & Zhang, J. (2024). Promotion and demotion contests. *Journal of Economic Behavior & Organization*, 219, 124–151.
- Li, H., & Zhou, L.-A. (2005). Political turnover and economic performance: The incentive role of personnel control in China. *Journal of Public Economics*, 89(9), 1743–1762. [\[CrossRef\]](#)
- Lu, X., & Landy, P. F. (2014). Show me the money: Interjurisdiction political competition and fiscal extraction in China. *American Political Science Review*, 108, 706–722. [\[CrossRef\]](#)
- Malcomson, J. M. (1986). Rank-order contracts for a principal with many agents. *The Review of Economic Studies*, 53(5), 807–817. [\[CrossRef\]](#)
- Maskin, E., Qian, Y., & Xu, C. (2000). Incentives, information, and organizational form. *The Review of Economic Studies*, 67(2), 359–378. [\[CrossRef\]](#)
- Michels, R. (1915). *Political parties: A sociological study of the oligarchical tendencies of modern democracy*. Hearst's International Library Company.
- Milgrom, P., & Shannon, C. (1994). Monotone comparative statics. *Econometrica*, 62(1), 157–180. [\[CrossRef\]](#)

- Montagnes, B. P., & Wolton, S. (2019). Mass purges: Top-down accountability in autocracy. *American Political Science Review*, 113(4), 1045–1059. [\[CrossRef\]](#)
- Myerson, R. B. (2006). Federalism and incentives for success of democracy. *Quarterly Journal of Political Science*, 1, 3–23. [\[CrossRef\]](#)
- Nalebuff, B. J., & Stiglitz, J. E. (1983). Prizes and incentives: Towards a general theory of compensation and competition. *The Bell Journal of Economics*, 14, 21–43. [\[CrossRef\]](#)
- Orrison, A., Schotter, A., & Weigelt, K. (2004). Multiperson tournaments: An experimental examination. *Management Science*, 50(2), 268–279. [\[CrossRef\]](#)
- Powell, R. (2007). Defending against terrorist attacks with limited resources. *American Political Science Review*, 101(03), 527–541. [\[CrossRef\]](#)
- Prato, C., & Wolton, S. (2016). The voters' curses: Why we need goldilocks voters. *American Journal of Political Science*, 60(3), 726–737. [\[CrossRef\]](#)
- Prendergast, C. (1999). The provision of incentives in firms. *Journal of Economic Literature*, 37(1), 7–63. [\[CrossRef\]](#)
- Ryvkin, D. (2010). The selection efficiency of tournaments. *European Journal of Operational Research*, 206(3), 667–675. [\[CrossRef\]](#)
- Waldman, M. (2013). Classic promotion tournaments versus market-based tournaments. *International Journal of Industrial Organization*, 31(3), 198–210. [\[CrossRef\]](#)
- Wang, R., Lützenrath, J., & Wolton, S. (2022). Can mass purges root out corruption? SSRN. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.