In this online appendix, we collect the analyses and discussions omitted from the main text. Online Appendix A provides sufficient condition under which a flexible cap or no cap can be optimal in a two-player Tullock contest setting. Online Appendix B characterizes the optimal cap schemes in a multi-player contest with two player types. Online Appendix C collects the proofs of propositions.

A Optimal Cap Schemes in Two-Player Contests

Proposition A1 (Flexible Cap vs. No Cap in Two-player Tullock Contests)
Suppose that $n = 2$, $\lambda \in [0, 1]$, and $r \in (0, 1]$. The following statements hold.

(i) If
\[
\frac{r(1 - v^r)}{1 + v^r} + \frac{(1 - v)\lambda - 1}{1 + v} > 0,
\] (A1)
then the optimal contest imposes a flexible cap.

(ii) If
\[
v [(2 + r)v^r - r] > \lambda(1 - v^r)(r - v),
\] (A2)
then the optimal contest imposes no cap.

Remark follows immediately from Proposition A1.
B Optimal Cap Schemes in Multi-player Contests with Two Player Types

The two-player example in Section 3.4 and Figure 1 provide an intuitive account of the fundamental trade-off between the cost and competition effects in asymmetric contests, as well as how the optimum depends on players’ type differential and the noisiness of the winner-selection mechanism. However, a multi-player contest differs substantially from its bilateral counterpart. In a two-player contest, player heterogeneity can be captured by a single parameter, $v \equiv v_2/v_1$. In contrast, heterogeneity is inherently multidimensional with three or more players, which cannot readily be defined or measured without imposing a specific structure on the profile of prize valuations $(v_1, \ldots, v_n)$. This nuance prevents handy comparative statics.

We consider a simple Tullock contest setting with a two-type distribution—i.e., stronger and weaker—to demonstrate the complications. There are $n_s$ stronger players and $n_w$ weaker players, with $n_s + n_w = n \geq 3$. The former type values the prize at $v_s$, while the latter values it at $v_w$, with $v_s \geq v_w > 0$. Despite the vast simplification, it is difficult to provide a simple account of the heterogeneity between players, as in the previous section: This depends on prize valuations across types—i.e., the ratio between $v_s$ and $v_w$—and also the composition of types within the pool, i.e., $(n_s, n_w)$. We analyze two simple cases, which demonstrate that a variation in either dimension may change the optimum fundamentally.

Case I: $n_s = 1$. We first assume one stronger player vs. $n - 1$ weaker opponents. The following result can be obtained.

**Proposition A2 (Optimal Contest with One Strong Player)** Suppose that $n_s = 1$, $n_w \geq 2$, and $\lambda + r > 1$. There exist two cutoffs $\hat{v}_h(\lambda, r) \in (0, 1)$ and $\hat{v}_l(\lambda, r) \in (0, 1)$ such that a flexible cap is optimal if $v_w/v_s < \hat{v}_l(\lambda, r)$ and no cap is optimal if $v_w/v_s > \hat{v}_h(\lambda, r)$.

The prediction is largely in line with that of Proposition A1 in a two-player setting. When $v_w/v_s$ is sufficiently small, a flexible cap plays a more significant equalizing role. Conversely, the optimum requires no cap when $v_w/v_s$ is sufficiently large: The direct discount on bidding incentives outweighs the limited equalizing role of a bid cap; as a result, the contest needs no intervention.

Case II: $n_s \geq 2$. The prediction drastically differs in the case of two or more stronger players, and the optimum with respect to the ratio $v_w/v_s$ can be nonmonotone.

**Proposition A3 (Optimal Contest with Two or More Strong Players)** Suppose that $n_s \geq 2$ and $n_w \geq 1$. Fixing $\lambda < 1$ and $r < 1$, there exists a lower threshold $\tilde{v}(\lambda, r) \in (0, 1)$
and an upper threshold $\bar{v}(\lambda, r) \in (0, 1)$, with $\bar{v}(\lambda, r) \geq v(\lambda, r)$, such that no cap is optimal if $v_w/v_s < \bar{v}(\lambda, r)$ or $v_w/v_s > \bar{v}(\lambda, r)$.

Although a sufficiently large ratio of $v_w/v_s$—i.e., $v_w/v_s > \bar{v}(\lambda, r)$—implies no policy intervention, as in Propositions A1 and A2, no cap also emerges as the optimum when $v_w/v_s$ is sufficiently small, i.e., $v_w/v_s < \bar{v}(\lambda, r)$, which overturns the predictions of Propositions A1 and A2. Proposition A3 suggests that a flexible cap can be optimal only if $v_w/v_s$ is in an intermediate range. This result reveals the complexity involved in a multi-player setting.

The competition effect loses its appeal when multiple stronger players are present. Suppose that $(n_s, n_w) = (2, 1)$. In this case, a stronger player has to outperform his equally competent peer to secure the prize, which may help discipline him from shirking regardless of the prevailing cap scheme. Meanwhile, a cap that handicaps the stronger may not effectively revive the weaker’s momentum, as a win is difficult regardless when outnumbered by more competent opponents. A smaller $v_w/v_s$ turns out to elevate the cost of a flexible cap: To level the playing field and incentivize the single underdog, a sufficiently high marginal tax rate is required to offset the initial asymmetry, which may cause excessive incentive loss from the two stronger players. In this scenario, contest design involves a hidden selection problem: The designer may simply “abandon” the weaker, while sustaining the competition between the stronger. This effect would not come into play in a bilateral contest.

C Proofs

Proof of Proposition A1

Proof. Clearly, with $n = 2$, both players are active in equilibrium and the set $\mathcal{P}$ defined in (23) can be simplified as

$$\mathcal{P} = \left\{(p_1^*, p_2^*) : p_1^* + p_2^* = 1, \frac{1}{2} \leq p_1^* \leq \frac{1}{1 + v^r}\right\}.$$

For notational convenience, define $p_1^\dagger := 1/(1 + v^r)$. Substituting $p_2^* = 1 - p_1^*$ into the contest objective (22), the maximization problem degenerates to a single-variable optimization problem as follows:

$$\max_{p_1^* \in [1/2, p_1^\dagger]} \mathcal{F}(p_1^*),$$

where

$$\mathcal{F}(p_1^*) = r \left\{ (1 - \lambda)vp_1^*(1 - p_1^*)^{1 - \frac{1}{h}} \left[ (p_1^*)^{\frac{1}{h}} + (1 - p_1^*)^{\frac{1}{h}} \right] \right\}.$$
\[ + \lambda \left[ 2vp_1^*(1-p_1^*) + (1-p_1^*) \left[ p_1^* - (p_1^*)^{1+\frac{1}{r}}(1-p_1^*)^\frac{1}{r} \right] \right]. \]

Carrying out the algebra, we can obtain that
\[ \mathcal{F}'(p_1^*) = (1-p_1^*) \mathcal{G}(\eta), \]
where \( \eta := p_1^*/(1-p_1^*) \in [1, v^{-r}] \) and
\[ \mathcal{G}(\eta) := r \left\{ (1-\lambda) v \left[ \left( 1 + \frac{1}{r} \right) \eta^\frac{1}{r} + \left( \frac{1}{r} - 1 \right) \eta^{1+\frac{1}{r}} + 1 - \eta \right] \right. \]
\[ + \lambda \left. \left[ 2v (1-\eta) + \left( 1-\eta + \left( \frac{1}{r} - 1 \right) \left( \frac{1}{\eta} \right)^{\frac{1}{r}} + \left( \frac{1}{r} + 1 \right) \left( \frac{1}{\eta} \right)^{\frac{1}{r}+1} \right) \right] \right\}. \]

It can be verified that \( p_1^* = p_1^\dagger = 1/(1+v^r) \), or equivalently, \( \eta = v^{-r} \), in a two-player contest without a cap. Therefore, a sufficient condition for a flexible cap to be optimal is \( \mathcal{F}'(p_1^\dagger) < 0 \), or equivalently, \( \mathcal{G}(v^{-r}) < 0 \). Carrying out the algebra, we can obtain that
\[ \mathcal{G}(v^{-r}) = v^{-r} \times \left\{ (1-\lambda) \left[ (r+1)v^r + 1 - r + rv^{r+1} - rv \right] \right. \]
\[ + \lambda \times \left. \left[ (r+1)v^{r+1} + rv^r + (1-r)v - r \right] \right\} \]
\[ = v^{-r} \times \left[ \lambda(v^r + 1)(v-1) + r(v+1)(v^r - 1) + (v^r + 1) \right] \]
\[ = -(1+v^{-r})(v+1) \times \left[ \frac{r(1-v^r)}{1+v^r} + \frac{(1-v)\lambda - 1}{1+v} \right]. \]

It is evident that \( \mathcal{G}(v^{-r}) < 0 \) if
\[ \frac{r(1-v^r)}{1+v^r} + \frac{(1-v)\lambda - 1}{1+v} > 0, \]
which corresponds to (A1(i) in Proposition A1(i)).

Next, note that \( \mathcal{G}(\eta) \) can be bounded from below by
\[ \mathcal{G}(\eta) = (1-\lambda)v \left[ \left( 1 + \frac{1}{r} \right) \eta^\frac{1}{r} + \left( \frac{1}{r} - 1 \right) \eta^{1+\frac{1}{r}} + 1 - \eta \right] \]
\[ + \lambda \left[ 2v(1-\eta) + 1 - \eta + \left( \frac{1}{r} - 1 \right) \eta^{-\frac{1}{r}} + \left( \frac{1}{r} + 1 \right) \eta^{1-\frac{1}{r}} \right] \]

\[ \geq (1 - \lambda) v \left[ \left( 1 + \frac{1}{r} \right) + \left( \frac{1}{r} - 1 \right) + 1 - v^{-r} \right] \]

\[ + \lambda \left[ 2v(1 - v^{-r}) + 1 - v^{-r} + \left( \frac{1}{r} - 1 \right) v + \left( \frac{1}{r} + 1 \right) v^{1-r} \right] \]

\[ = \frac{v^{-r}}{r} \left\{ v \left[ (2 + r)v^r - r \right] + \lambda (v^r - 1)(r - v) \right\} , \]

where the inequality follows from \( \eta \in [1, v^{-r}] \). Clearly, \( G(\eta) > 0 \) for all \( \eta \in [1, v^{-r}] \), or equivalently, \( F'(p^*_1) > 0 \) for all \( p^*_1 \in [\frac{1}{2}, p^*_1] \), if

\[ v \left[ (2 + r)v^r - r \right] > \lambda (1 - v^r)(r - v), \]

which implies that \( F(p^*_1) \) is uniquely maximized at \( p^*_1 = p^*_1 \) on \([\frac{1}{2}, p^*_1]\) and it is optimal to have no cap. Note that the above inequality corresponds to \( \text{(A2)} \) in Proposition \( \text{A1}(\text{ii}) \). This completes the proof. \( \blacksquare \)

**Proof of Proposition A2**

**Proof.** Note that players of the same type must win with equal probabilities in equilibrium. Therefore, the winning probability distribution \( \mathbf{p}^* \equiv (p^*_1, \ldots, p^*_n) \) is fully characterized by \((p^*_s, p^*_w)\), where \( p^*_s \) and \( p^*_w \) respectively represent the stronger players’ and the weaker players’ equilibrium winning probabilities. With slight abuse of notation, the set \( \mathcal{P} \) defined in (23) can then be simplified as

\[ \mathcal{P} = \left\{ (p^*_s, p^*_w) : n_sp^*_s + n_wp^*_w = 1, 1/n \geq p^*_w \geq p^*_w \right\}, \]

where \( p^*_w \) is the equilibrium winning probability of each weaker player under no cap. Normalizing \( v_s \) to 1 without loss of generality and substituting \( p^*_s = (1 - n_wp^*_w)/n_s \) into the contest objective (22), the designer’s optimization problem boils down to

\[ \max_{p^*_w \in [p^*_w, 1/n]} F(p^*_w), \]

where \( F(\cdot) \) is given by

\[ F(p^*_w) := (1 - \lambda)v_w(p^*_w)^{1-\frac{1}{r}}(1 - p^*_w)^{\frac{1}{r}} \left[ n_s \left( \frac{1 - n_wp^*_w}{n_s} \right)^{\frac{1}{r}} + n_wp^*_w \right]^{\frac{1}{r}} \]
Carrying out the algebra, we can obtain that

\[ \mathcal{F}'(p_w^*) = (1 - \lambda)v_w \times \left\{ \left( 1 - \frac{1}{r} \right) \left( p_w^* \right)^{-\frac{1}{r}} (1 - p_w^*) \left[ n_s(p_s^*)^{\frac{1}{r}} + n_w(p_w^*)^{\frac{1}{r}} \right] \right. \]

\[ - \left. (p_w^*)^{1-\frac{1}{r}} \left[ n_s(p_s^*)^{\frac{1}{r}} + n_w(p_w^*)^{\frac{1}{r}} \right] \right) + \frac{1}{r} n_w(p_s^*)^{\frac{1}{r}} (1 - p_s^*) \left[ (p_s^*)^{\frac{1}{r}} - (p_w^*)^{\frac{1}{r}} \right] \]

\[ - n_s(p_s^*)^{1-\frac{1}{r}} (1 - p_s^*) \left[ \frac{n_w}{n_s} (p_s^*)^{\frac{1}{r}-1} + (p_w^*)^{\frac{1}{r}-1} \right] \]

\[ + \frac{n_w}{n_s} (1 - 2p_w^*) \right\}, \quad \text{(A4)} \]

Recall that \( p_w^\dagger \) is the equilibrium winning probability of each weaker player under no cap. Therefore, for a flexible cap to be optimal, it suffices to show that \( \mathcal{F}'(p_w^\dagger) > 0 \) when \( v_w \) is sufficiently small.

Denote the equilibrium winning probability of each strong player by \( p_s^\dagger \). We first take a closer look at the equilibrium winning probability \( (p_s^\dagger, p_w^\dagger) \) under no cap. From the first-order conditions for each type of players, we have that

\[ (p_s^\dagger)^{1-\frac{1}{r}} (1 - p_s^\dagger) = v_w (p_w^\dagger)^{1-\frac{1}{r}} (1 - p_w^\dagger). \quad \text{(A5)} \]

Note that \( n_s = 1 \) by assumption. Therefore, we have that \( p_s^\dagger = 1 - n_w p_w^\dagger \). Substituting the expression of \( p_s \) into the above condition, for a sufficiently small \( v_w \), we can obtain that

\[ p_w^\dagger = \left( \frac{v_w}{n_w} \right)^r [1 + o(1)]. \]

Carrying out the algebra, for a sufficiently small \( v_w \), we have that

\[ \mathcal{F}'(p_w^\dagger) = (1 - \lambda) \times \left\{ v_w \left( 1 - \frac{1}{r} \right) \left( \frac{v_w}{n_w} \right)^{-1} [1 + o(1)] + o(1) \right\} \]

\[ + \lambda \times \left\{ n_w [1 + o(1)] + o(1) \right\} \]

\[ = \frac{n_w}{r} (\lambda + r - 1) + o(1) > 0, \]
where the strict inequality follows from the condition $\lambda + r > 1$ assumed in Proposition A2.

In other words, there exists a threshold $\hat{v}_l(\lambda, r) > 0$ such that imposing a flexible cap is optimal to the designer for all $v_w/v_s < \hat{v}_l(\lambda, r)$.

Next, we show that having no cap is optimal if $v_w$ is sufficiently large. It is evident that $p^+_s = 1/n + o(1)$ and $p^+_w = 1/n + o(1)$ in this case. Therefore, $\mathcal{F}(p^*_w)$ in (A4) can be bounded from above by

$$
\mathcal{F}(p^*_w) = (1 - \lambda) \times n \times \left( 1 - \frac{1}{r} \right) \left( 1 - \frac{1}{n} \right) - n \times \frac{1}{n} + o(1)
+ \lambda \times \left[ -n \times \frac{1}{r} \left( 1 - \frac{1}{n} \right) + n \times \left( 1 - \frac{2}{n} \right) + o(1) \right] < 0, \text{ for all } p^*_w \in [p^+_w, 1/n].
$$

Therefore, there exists a threshold $\hat{v}_h(\lambda, r) > 0$ such that having no cap is optimal for all $v_w/v_s > \hat{v}_h(\lambda, r)$. This concludes the proof. ■

**Proof of Proposition A3**

**Proof.** Similar to the proof of Proposition A2, we normalize $v_s$ to 1 without loss of generality.

We first consider the case in which $v_w$ is sufficiently small. It is evident that $p^+_w = o(1)$ and $p^+_s = 1/n_s + o(1)$. It follows from the first-order conditions (A5) that

$$
p^+_w = \frac{1}{n_s} \left( \frac{v_w n_s}{n_s - 1} \right)^{\frac{r}{1-r}} \left[ 1 + o(1) \right].
$$

By the above equation and (A3), when $v_w$ is sufficiently small, we can obtain that

$$
\mathcal{F}(p^+_w) = (1 - \lambda) v_w \left\{ \frac{1}{n_s} \left( \frac{v_w n_s}{n_s - 1} \right)^{\frac{r}{1-r}} \left[ 1 + o(1) \right] \right\}^{1-\frac{1}{r}} \frac{1}{n_s} \left[ 1 + o(1) \right]
+ \lambda \left\{ n_s(p^+_s)^{-1}(1 - p^+_s) \left[ 1 + o(1) \right] + o(1) \right\}
= (1 - \lambda) \left( 1 - \frac{1}{n_s} \right) + \lambda \left( 1 - \frac{1}{n_s} \right) + o(1) = 1 - \frac{1}{n_s} + o(1).
$$

For $p^*_w > \frac{v_w}{n_s}$, we have that

$$
\mathcal{F}(p^*_w) = (1 - \lambda) v_w (p^*_w)^{1-\frac{1}{r}}(1 - p^*_w) \left[ n_s(p^*_s)^{\frac{1}{r}} + n_w(p^*_w)^{\frac{1}{r}} \right]
+ \lambda \left\{ n_s(p^*_s)^{1-\frac{1}{r}}(1 - p^*_s) \left[ (p^*_s)^{\frac{1}{r}} - (p^*_w)^{\frac{1}{r}} \right] + n v_w p^*_w (1 - p^*_w) \right\}
\leq (1 - \lambda) v_w (p^*_w)^{1-\frac{1}{r}}(n_s p^*_s + n_w p^*_w) + \lambda \left[ n_s(p^*_s)^{1-\frac{1}{r}}(p^*_s)^{\frac{1}{r}}(1 - p^*_s) + n v_w p^*_w \right]
$$

A7
\[(1 - \lambda) v_w (p_w^*)^{1 - \frac{1}{r}} + \lambda \left[ n_s p_w^* (1 - p_w^*) + nv_w p_w^* \right] \]
\[\leq (1 - \lambda) \frac{1}{n_s} v_w + \lambda \left( 1 - \frac{1}{n_s} + n v_w \frac{2 + \epsilon}{n_s} \right) \]
\[= \lambda \left( 1 - \frac{1}{n_s} \right) + o(1) < F(p_w^*),\]

where the last inequality follows from \( \lambda < 1 \).

For \( p_w^* \leq v_w^{\frac{2 + \epsilon}{2r}} \), it follows from (A4) that
\[
F'(p_w^*) = (1 - \lambda) \times \left\{ \left( 1 - \frac{1}{r} \right) v_w (p_w^*)^{-\frac{1}{r}} n_s^{-\frac{1}{r}} \left[ 1 + o(1) \right] \right\} + \lambda \times O(1)
\]
\[\leq (1 - \lambda) \left( 1 - \frac{1}{r} \right) n_s^{-\frac{1}{r}} v_w^{\frac{2 + \epsilon}{2r}} \left[ 1 + o(1) \right] < 0.\]

To summarize, \( F(p_w^*) \) is strictly decreasing in \( p_w^* \) for \( p_w^* \in [p_w^{'\dagger}, v_w^{\frac{2 + \epsilon}{2r}}] \) and \( F(p_w^*) < F(p_w^{'\dagger}) \) for all \( p_w^* \in (v_w^{\frac{2 + \epsilon}{2r}}, 1/n] \) if \( v_w \) is sufficiently small, which in turn implies that there exists a threshold \( \bar{v}(\lambda, r) > 0 \) such that having no cap is optimal for all \( v_w / v_s < \bar{v}(\lambda, r) \).

Next, we consider the case where \( v_w \) is sufficiently large. In this case, we have that
\( p_w^{'\dagger} = 1/n + o(1) \) and \( p_s^{'\dagger} = 1/n + o(1) \). Therefore, for all \( p_w^* \in [p_w^{'\dagger}, 1/n] \), we have that
\[
F'(p_w^*) = (1 - \lambda) \times n \times \left\{ \left( 1 - \frac{1}{r} \right) \left( 1 - \frac{1}{n} \right) - n \times \frac{1}{n} + o(1) \right\}
\[+ \lambda \times \left\{ -n \times \frac{1}{r} \left( 1 - \frac{1}{n} \right) + n \times \left( 1 - \frac{2}{n} \right) + o(1) \right\} < 0,
\]
and thus \( F(p_w^*) \) is strictly decreasing in \( p_w^* \), which implies the optimality of imposing no cap on the contest. Therefore, there exists \( \bar{v}(\lambda, r) \) such that having no cap is optimal for all \( v_w / v_s > \bar{v}(\lambda, r) \). This concludes the proof. ■