Winner's Effort in Multi-Battle Team Contests*

Stefano Barbieri†, Marco Serena‡

April 23, 2024

Abstract

In multi-battle team contests with pairwise battles, how battles are organized—sequentially or (partially) simultaneously—may affect the expected winner's total effort (WE), a natural objective in R&D races, elections, and sports. We focus on noise (modeled via the contest success function's discriminatory power) and across-team heterogeneity, abstracting from player-specific heterogeneity. With sufficient noise, we show that: (1) If teams are symmetric, all temporal structures yield the same WE; and (2) If teams are asymmetric, WE is maximized by a fully simultaneous contest and minimized by a fully sequential one. With no noise, we show that: (3) If teams are symmetric, WE is maximized by a fully sequential contest and minimized by a fully simultaneous one; and (4) If teams are asymmetric, neither the fully sequential nor the fully simultaneous temporal structures maximize or minimize WE. Our results use a novel technique that simplifies temporal structure comparisons: extractions and mergers.

JEL classification codes: C72, D72, D74, D82

Keywords: team contest, winner's effort, temporal structures, team-asymmetry effect, stochastic-effort effect.

*We are pleased to acknowledge useful comments by four referees of this journal, the invaluable help of two Associate Editors, and Hervé Moulin as Editor-in-Chief. We also thank participants at various conferences and the Max Planck Institute for Tax Law and Public Finance for support and hospitality. All errors are our own.

†Tulane University. *Email*: sbarbier@tulane.edu

‡CUNEF Universidad, Madrid. *Email*: marco.serena@cunef.edu.

1 Introduction

Groups, coalitions, and joint ventures often compete over multiple confrontations in a multi-battle *team* contest. We focus on contests with pairwise battles, in which team members fight in separate battles and a team wins the overall competition if a sufficient number of battles is won by its members. Real-life examples of such multi-battle team contests are ubiquitous. In R&D races, the patent is granted to the first joint venture that achieves a breakthrough by winning a sufficient number of innovation stages against its rival as the patent race unfolds (Harris and Vickers, 1987). In elections around the world, coalitions of political parties compete and a coalition wins the control of the government if its members win in a sufficient number of local elections (Snyder, 1989). In sports, the victory of a team member in a match contributes to her team's overall success, which all team members benefit from (Szymanski, 2003). In multi-battle team contests, diverse "temporal structures" arise by nature or design: the battles between teams could be played out sequentially or (partially) simultaneously. We investigate how temporal structures affect the total expected equilibrium effort of the winning team players—henceforth WE.

WE is a key variable of interest in team contests (as well as in individualistic ones; see, e.g., Drugov and Ryvkin, 2017, and Serena, 2017). Among the three examples above, consider first R&D races. Efforts exerted on the winning project might be more relevant than those exerted on the losing project, as the winning project receives more funding, market share, visibility, or, in extreme cases, it is the *only* project ever implemented. Similarly, in elections, all rent-seeking activities may be considered wasteful, but especially problematic if campaign financing involves a quid pro quo with lobbyists that may lead to distortions in the economy. In this case, the expenditures of the winning party may be more relevant than those of the losing party because the winners' policies are ultimately implemented, whereas the losers' are not. As for sports, spectators may look unfavorably at a competition in which the difference between winners' and losers' efforts is too large, as in a runaway match.

We focus on the role of noise—modeled via the discriminatory power of the contest success function—and that of team heterogeneity—modeled via the across-team differences in the marginal costs of efforts of players. Across-team heterogeneity may or may not be present, battle-specific heterogeneity is allowed, but player-specific heterogeneity is ruled out. For instance, in R&D races all members of one joint venture may be favored because it is composed of domestic firms, in politics one party may have a better get-out-the-vote apparatus, and in sports a team may have better coaches. As it often happens in contests, sufficient noise yields a pure-strategy equilibrium so efforts are deterministic, while no noise (i.e., a perfectly discriminating contest) yields a mixed-strategy equilibrium so efforts are stochastic. The interplay of noise and team heterogeneity, regardless of battle-specific heterogeneity, gives rise to four main results.

- Result 1: when the contest is sufficiently noisy and teams are symmetric, WE is unaffected by the temporal structure.
- Result 2: when the contest is sufficiently noisy and teams are asymmetric, WE is maximized by a fully simultaneous contest and minimized by a fully sequential one.
- Result 3: when the contest is not noisy and teams are symmetric, WE is maximized by a fully sequential contest and minimized by a fully simultaneous one.
- Result 4: when the contest is not noisy and teams are asymmetric, neither the fully sequential nor the fully simultaneous temporal structures maximize or minimize WE.

Our work complements and builds on Fu, Lu, and Pan (2015)—henceforth FLP—who elegantly capture the key features of the competition in multi-battle team contests with pairwise battles, i.e., where each player of one team is matched in a battle against a player of the rival team, and each player fights and bears the effort cost of her battle only.[1] Our model in Section 2 is a version of FLP's.

FLP characterize how the temporal structure affects the expected total effort (henceforth TE), rather than WE. While they find that, for a wide range of multi-battle team contests and regardless of noise and heterogeneity, TE is not affected by the temporal structure, we find that WE may depend on the contest temporal structure. The intuition is as follows. In a team contest à la FLP, a member of the laggard team anticipates that the cost of future efforts will be borne by teammates and thus the *discouragement effect* that permeates individualistic contests does not apply.[2] The temporal-structure independence for TE derived by FLP obtains because a change in the temporal structure shifts efforts from some battles to

[1] The structure of FLP with pairwise battles is not the only possible one for team contests. Among others, in Baik et al. (2001) all efforts in a group are added up, in Barbieri et al. (2014) and Chowdhury et al. (2016) the group effort is identified as, respectively, the best shot or the weakest link, and in Häfner (2017) the overall contest is modeled as a tug-of-war. For a literature review on group contests, and more generally on contests with multiple battles, see Kovenock and Roberson (2012) and Fu and Wu (2019). For experimental evidence on the comparison of individualistic and team contests see Fu, Ke, and Tan (2015).

[2] The discouragement effect in multi-battle *individualistic* contests, where the same two players fight against each other in all battles, arises as the laggard anticipates having to provide more future effort than the leader in order to win, and competitiveness and efforts suffer. An analysis of the discouragement effect appears in Harris and Vickers (1987) and is further discussed, for instance, by Klumpp and Polborn (2006), Konrad and Kovenock (2009), Malueg and Yates (2010), and Feng and Lu (2018). For a review of the theoretical literature, see Konrad (2009), and for a review of the experimental and empirical literature testing and measuring the discouragement effect, see Dechenaux et al. (2015), Decamps et al. (2022), and references therein.

others, but without discouragement effect these shifts in efforts perfectly balance out when considering TE, which weighs equally all players' efforts. In contrast, WE does not weigh equally winners' and losers' efforts, and hence, WE *may be* affected by the above-described shifts in efforts, and we characterize how.

Results 1-2. Consider asymmetric teams. We show that the common knowledge of an early victory in a battle by the stronger team depresses both teams' efforts in late battles and thus reduces WE, compared to when battles are played out simultaneously. In contrast, the common knowledge of an early victory in a battle by the weaker team boosts both teams' efforts in late battles and thus increases WE. On average, the reduction dominates the increase because members of the stronger team are ex-ante more likely to win. We name this the "*team-asymmetry effect*." This WE-reducing effect arises when some battles are played out before others—e.g., sequentially—so that common knowledge of early victories is possible. The team-asymmetry effect drives Result 2 in the setup with pure-strategy equilibrium (due to the sufficient noise in the contest). If instead teams are symmetric, the team-asymmetry effect is absent and all temporal structures yield the same WE when the equilibrium is in pure strategies; this is Result 1. Results 1 and 2 are in Section 3.

Results 3-4. An additional force arises when the contest has no noise so equilibrium efforts are stochastic. In this case, the expected equilibrium effort of a player in a battle, conditional on winning that battle, is strictly larger than conditional on losing that battle. WE only considers the efforts of the winning team, regardless of whether its members won or lost their individual battle. Thus, WE "suffers" if the winning team contains players who lost their respective battles; we name this the "*stochastic-effort effect*." More sequential temporal structures mitigate this WE-reducing effect; e.g., in a best-of-three sequential contest, the third battle matters to players if and only if it is a tie-break, and in this case it is *always* won by the team that wins the overall contest, but the third battle may be won by the losing team if the best-of-three contest is simultaneous. The stochastic-effort effect drives Result 3 in the symmetric setup: i.e., with no team-asymmetry effect. Finally, the interplay of the team-asymmetry and stochastic-effort effects yields Result 4. Results 3 and 4 are in Section 4.

Our results are important from a contest design perspective and relevant for policymaking in practical contexts. For instance, our ranking of WE among different temporal structures with observable battle outcomes can also be interpreted as a ranking of WE among different feedback levels on battle outcomes when battles are played out sequentially (e.g., a simultaneous contest can be thought of as a sequential contest with full concealment on battle outcomes). In R&D contests, "the regulatory regimes in most of EU countries require mandatory disclosure of firms' R&D activities" (see Fu and Lu, 2020; p. 1830). Our results suggest that this mandatory disclosure policy may be beneficial if one interprets WE as the expected quality of the winning patent when the competition is not noisy and competitors are symmetric, but may be counterproductive when the competition is sufficiently noisy or asymmetries are present.[3] Also, our results can be applied to the design of elections; if one is interested in how much the winner of the election is beholden to special interests, then our results suggest that a sequential election minimizes the effort of (and contributions collected by) the winning team when the election outcome depends on campaign expenditures in a sufficiently noisy way. For the importance of sequential voting see, e.g., Battaglini (2005).

Our results are derived with a novel technique that simplifies temporal structure comparisons: extractions and mergers. The former makes the contest temporal structure "more sequential" and the latter "more simultaneous." Extractions and mergers can also be used to derive comparisons in the case of semi-mixed equilibrium (in Section 5) and in three extensions (in Section 6). For semi-mixed equilibrium, we obtain a result that straddles those with pure and mixed-strategy equilibrium. For the three extensions, first, in the simple setup with deterministic equilibrium efforts, we show that WE in a fully sequential contest is minimized when the *order of battles* is from hardest to easiest—i.e., from highest to lowest marginal effort cost.[4] Hence, when both the temporal structure and the order of battles are endogenous, WE is minimized by a fully sequential contest with battles ordered from highest to lowest marginal effort cost and maximized by a fully simultaneous contest. Second, in *private information* setups, stochastic efforts arise naturally due to the stochastic realization of, say, marginal effort costs. With symmetric teams, we show that WE is maximized by a fully sequential contest and minimized by a fully simultaneous one (as in Result 3) under mild extra regularity assumptions. Third, in the setup with pure-strategy equilibrium, we show that some form of *player-specific heterogeneity* can be introduced without affecting Result 2.

2 The model and WE

Two rival teams, A and B, consist of $2n+1$ players each, with n a finite, strictly positive integer. Players of rival teams are matched in $2n+1$ head-to-head battles. Battles are indexed

[3] However, while in individualistic contests temporal structures closely parallel information feedback policies (see, e.g., Hinnosaar, 2024), in team contests one needs to be more careful as teammates in sequential contests may still communicate with each other and get to know the outcome of previous battles even if the contest designer concealed such outcomes. Hence, the policy implications of our results in terms of information feedback go through in applications where one can safely abstract away from within-team communication concerns. We thank an Associate Editor and an anonymous reviewer for suggesting this point.

[4] We study the effect of the order of battles on WE; Fu and Lu (2020) study the winning-odds-maximizing teams' choices of the order of players. See also Hamilton and Romano (1998) and Konishi et al. (2022).

by m, with $m \in \{1, ..., 2n+1\}$. In any battle, the two matched players simultaneously choose non-negative efforts. Let the effort level actually chosen by the player of team $i \in \{A, B\}$ in battle m be denoted by $\xi_i^m \geq 0$ and the constant marginal effort cost by $c_i^m > 0$. Players of the team that wins at least $n+1$ battles obtain a prize of 1 each; hence, if team $i \in \{A, B\}$ wins at least $n+1$ battles, the player of team i fighting in battle m obtains a payoff of $1 - c_i^m \xi_i^m$ and the player of the rival team $j \neq i$ obtains a payoff of $-c_j^m \xi_j^m$.

Temporal structures and battle outcomes. Battles can be carried out sequentially or (partially) simultaneously. We define the temporal structure as a finite collection of strictly positive natural numbers $T = \{n_1, n_2, ...\}$, with $n_1 + n_2 + ... = 2n + 1$. We define a *battle outcome* as either "A" or "B," according to who wins that specific battle. The first n_1 pairs of players simultaneously choose their efforts in round 1, then battle outcomes realize; i.e., the winner of each of the n_1 battles is determined and publicly revealed. Then, the following n_2 pairs of players simultaneously choose their efforts in round 2 and battle outcomes realize, and so on. More formally, the $2n+1$ pairs of players are partitioned into sets, which number between 1 and $2n+1$. The size of the j^{th} set is denoted by n_j. Pairs of players partitioned into the j^{th} set simultaneously choose efforts in round j upon observing the outcomes of previous battles; i.e., who won all battles played in the k^{th} set, for all $k < j$. If pairs of players are partitioned into one unique set, then $n_1 = 2n+1$ (a fully simultaneous contest). If pairs of players are partitioned into $2n+1$ sets, then $n_1 = n_2 = ... = 1$ (a fully sequential contest). As in FLP, we assume that the pairs of players fighting a particular battle remain unchanged when we compare temporal structures.

From a temporal structure T, we define the set of all possible paths of battle outcomes Ω^T of the game; the paths that compose Ω^T record the player who wins each battle fought before the identity of the winning team becomes common knowledge (given T). Any specific path $\omega \in \Omega^T$ belongs to one of the two disjoint sets Ω_A^T and Ω_B^T that contain all the possible terminal paths of battle outcomes: that is, Ω_A^T and Ω_B^T contain all the paths of battle outcomes leading to the overall victory of team A or B, respectively. For instance, $\Omega_A^{\{1,1,1\}} = \{A, A; A, B, A; B, A, A\}$ and $\Omega_A^{\{3\}} = \{AAA; AAB; ABA; BAA\}$, where commas separate rounds and semicolons separate paths. Note that a team may win the overall contest although some of its members lose their battle and that the elements in Ω^T *exclude* battles occurring when the identity of the winning team is common knowledge. Thus, in any $\omega \in \Omega_A^T$, at least $n+1$ battle outcomes are "A" and at most n "B", and similarly for Ω_B^T.

Battles. The player of team A wins battle m with probability $p_A(\xi_A^m, \xi_B^m)$, and that of team B with probability $p_B(\xi_A^m, \xi_B^m) = 1 - p_A(\xi_A^m, \xi_B^m)$. We consider two alternative setups;

a generalized Tullock success function, where

$$p_A\left(\xi_A^m, \xi_B^m\right) = \begin{cases} \frac{\left(\xi_A^m\right)^r}{\left(\xi_A^m\right)^r + \left(\xi_B^m\right)^r} \text{ with } r \in (0, 2] & \text{if } \xi_A^m + \xi_B^m > 0, \\ 1/2 & \text{if } \xi_A^m + \xi_B^m = 0, \end{cases} \quad (1)$$

and an all-pay auction, where

$$p_A\left(\xi_A^m, \xi_B^m\right) = \begin{cases} 1 & \text{if } \xi_A^m > \xi_B^m, \\ 1/2 & \text{if } \xi_A^m = \xi_B^m, \\ 0 & \text{if } \xi_A^m < \xi_B^m. \end{cases} \quad (2)$$

Jia (2008), for instance, provides a foundation for the common interpretation of r as noise parameter in (1); the smaller r, the larger the noise. As $r \to \infty$, the expression for p_A for given efforts in (1) converges to (2). We do not consider $r > 2$ and finite, as no full characterization of the equilibrium behavior exists even in a one-shot contest: Ewerhart (2015) partially characterizes equilibrium strategies and Ewerhart (2017b) fully characterizes equilibrium payoffs. Calculating WE requires the full characterization of equilibrium strategies because WE depends on the expected efforts *conditional on battle outcomes*.

In battle m, the marginal cost of effort c_i^m of a player of team $i \in \{A, B\}$ is the product of two non-negative components: c_i and γ_m. The first component, c_i, captures the heterogeneity between teams; it is team-specific and identical across teammates. This first component is useful to capture the knife-edge case of symmetric teams $c_A = c_B$. The second component, γ_m with $m \in \{1, ..., 2n + 1\}$, captures the heterogeneity across battles; it is battle-specific, but identical between the two players of that battle. This second component allows, for instance, marginal costs of effort in battle m' to be larger than those in battle m'' ($\gamma_{m'} > \gamma_{m''}$), in which case we say that battle m' is harder than battle m''. For instance, political advertising in some districts may be more expensive than in others. Furthermore, while this second component does not play a key role in our main results (Results 1-4), it makes it meaningful to analyze the order of battles, as we do in the Extensions.

As in FLP, each battle is played under common knowledge of the foregoing description of the game, including the temporal structure, all marginal effort costs, and, importantly, past battle outcomes. Therefore, for given costs, a given contest success function (CSF) which is either (1) or (2), and a given temporal structure T, a pure strategy of a player is a function from the history of observed past battles to non-negative numbers for efforts. Mixed strategies are defined as standard in game theory (see, e.g., Osborne, 2004; page 142). Also as in FLP, the *state* of the overall contest right before the n_j battles of round j is summarized by a pair, which we denote by (b_A, b_B), with $b_A + b_B = n_1 + ... + n_{j-1}$, where b_i is the number of battles won by team $i \in \{A, B\}$. Note that not all battles are "actually fought"; that is, if the winning team is already decided before round j (i.e., $\max\{b_A, b_B\} \geq n+1$), then all players will exert zero efforts in round j and afterward. Throughout the paper, whenever confusion does not arise, we omit to specify that the equilibrium quantities and properties we discuss hold for battles that are actually fought, which we call "non-trivial" battles.

Preliminary equilibrium observations. We consider subgame perfect equilibrium; henceforth, simply, equilibrium. We build on well-known existing results and on FLP for the equilibrium characterizations and use them to provide three preliminary observations that help us compare temporal structures and write a compact equilibrium expression for WE.

- Preliminary equilibrium observation 1. As the winner of the overall contest is determined solely by the team that wins the majority of battles, backward induction logic implies that *previous battles* affect the equilibrium strategies of current battles only through the current state (b_A, b_B), as in FLP. Formally, this observation follows from (i) in Lemma 1 of FLP: in a generic battle, the common prize players are fighting for is independent of past efforts and future temporal structures, and only depends on the current state. Intuitively, consider the contingencies when the non-trivial tie-breaking battle $2n+1$ is reached as the unique battle of the last round. Efforts exerted in all the $2n$ previous battles neither affect what players are fighting for in battle $2n+1$ nor players' costs in battle $2n+1$. As there is a unique equilibrium in battle $2n+1$, past history cannot serve as a coordination device. The only thing that matters for battle $2n+1$'s incentives is that it is played out at state (n, n). A standard induction argument could be used to extend the same reasoning to all battles, making the state always a sufficient statistics for equilibrium play.

- Preliminary equilibrium observation 2. In our setup, a team's equilibrium probability of winning any non-trivial battle is *identical across battles and unaffected by the temporal structure*. Formally, this observation follows from our assumptions on costs and Theorem 1 of FLP, which states that, in any non-trivial battle, the probability of winning of a player is determined solely by the matched players' effort costs. Indeed, the prize of every battle is the resulting increase in the probability that a player's own team wins. Such a battle prize is in *common* between the two matched players; "because the contest is to be won by one and only one team, one team's advance is the other's backslide" (see FLP, p. 2124). Moreover, the absence of the discouragement effect allows us to analyze each battle as a one-shot contest with the above-mentioned common battle prize. As is well known in the literature, when the prize is in common in a one-shot, complete information, two-player contest under (1) or (2) with constant marginal effort costs, the equilibrium win probability of each player depends only on the ratio of marginal costs—see, e.g., Malueg and Yates (2005) and Wang (2010) under (1) and Baye, Kovenock, and De Vries (1996) under (2). Thus, the individual equilibrium probability of winning any specific battle m depends only on c_A^m/c_B^m and on the specific CSF, but not on the common battle prize. Our assumptions on marginal effort costs further imply that $c_A^m/c_B^m = c_A/c_B$ for any m because γ_m cancels out. Hence, the preliminary equilibrium observation 2 follows.

- Preliminary equilibrium observation 3. In any non-trivial battle, the common battle prize is not affected by the temporal structure of future battles, but only by the current state of battles won by each team: (b_A, b_B). Formally, this observation follows from (ii) of Theorem 2 of FLP, which demonstrates that the ex-ante likelihood of victory, calculated in state $(0, 0)$, is independent of the battle sequence. This result actually extends to any starting state (b_A, b_B), as FLP mention in Section III.B on head starts. Intuitively, recall that the common battle prize in state (b_A, b_B) is the difference in likelihood of victory of the overall contest for each team in state $(b_A + 1, b_B)$ as opposed to that in state $(b_A, b_B + 1)$: these likelihoods do not depend on the future arrangements of battles and so their difference does not either. (For more details, see our discussion after (7).)

Building on preliminary equilibrium observation 2, for simplicity and henceforth, we denote by p_A (p_B) the equilibrium value of the per-battle win probability of team A (B) in a non-trivial battle. Whenever confusion does not arise, we omit to specify that the battles we consider (and the corresponding equilibrium quantities and properties) are non-trivial. Building on preliminary equilibrium observation 3, we denote by $\nu(b_A, b_B)$ the common battle prize the two matched players in a battle are fighting for, in state (b_A, b_B). The specific CSF, the value of n, and the ratio c_A/c_B affect p_A, p_B, and $\nu(b_A, b_B)$. We omit to specify these functional dependencies as they do not change with the temporal structure, which is our focus. Building on the three preliminary equilibrium observations, for given costs, CSF, and temporal structure T, we write equilibrium efforts in battle m as $x_A^m(b_A, b_B)$ and $x_B^m(b_A, b_B)$; these equilibrium efforts are deterministic under (1) for sufficiently low r and stochastic under (2).

Formal definition of WE. The main question of our paper is: in equilibrium, for given costs and CSF, how do changes in the temporal structure T affect the total expected equilibrium effort of the winning team? We denote by WE^T the total expected equilibrium effort of the winning team in temporal structure T, where we only leave explicit the dependence of WE on T (and not on costs and CSF, for instance) as the effect of T is our main concern. The ex-ante equilibrium probability of path $\omega \in \Omega_i^T$ is $\Pr\{\omega\}$. Hence,

$$WE^T = \sum_{i\in\{A,B\}} \sum_{\omega\in\Omega_i^T} \Pr\{\omega\} \cdot \sum_{m=1}^{2n+1} E\left[x_i^m\left(b_A^m(\omega), b_B^m(\omega)\right)|\omega\right], \tag{3}$$

where $\left(b_A^m(\omega), b_B^m(\omega)\right)$ is the state of the contest at battle m along path ω given the temporal structure T. (Recall that expected efforts in trivial battles are zero, so that summing over all $2n+1$ battles makes sense.) More precisely, say battle m belongs to the n_k battles of round k; then $b_i^m(\omega)$ is the number of i's among the first $n_1 + .. + n_{k-1}$ battle outcomes of ω, with $i \in \{A, B\}$. As in FLP, the state determines the (unique) equilibrium efforts in battle m for given CSF, T, and costs. This is the first way in which ω affects $E\left[x_A^m\left(b_A^m(\omega), b_B^m(\omega)\right)|\omega\right]$. The second way is that the m^{th} element of ω describes whether A won or lost battle m, which in turn affects the conditioning in the expectation on player A's equilibrium effort itself if stochastic, as we intuitively explain in what follows.

If efforts are stochastic, considering for instance $T = \{3\}$, the fact that the A-player lost the first battle along path $\hat{\omega} = BAA$ or won the first battle along path $\tilde{\omega} = ABA$ has implications for WE even if in both $\hat{\omega}$ and $\tilde{\omega}$ team A wins the overall contest. Suppose CSF (2) and let costs be such that both players' first-battle equilibrium efforts are uniformly distributed on $[0, 1]$. Then, the first-battle component of WE that regards the effort of A along path $\hat{\omega} = BAA$ is $1/3$ (i.e., the expectation of the minimum of two uniform distributions on $[0, 1]$ because A loses the first battle but belongs to the team that wins the contest), while that along path $\tilde{\omega} = ABA$ is $2/3$ (i.e., the expectation of the maximum of two uniform distributions on $[0, 1]$ because A wins the first battle and belongs to the team that wins the contest). This conditioning depends on both the CSF being considered (with CSF (2), the maximum effort wins and the minimum loses) and on the equilibrium effort of both players. We omit the dependence of $E\left[x_A^m\left(b_A^m(\omega), b_B^m(\omega)\right)|\omega\right]$ on the effort of player B because ω, through the state $\left(b_A^m(\omega), b_B^m(\omega)\right)$, uniquely determines both players' equilibrium efforts.

If efforts are deterministic, instead, such conditioning on winning or losing the first battle cannot affect the expected efforts and, therefore, the expression for WE in (3) can be simplified; $E\left[x_i^m\left(b_A^m(\omega), b_B^m(\omega)\right)|\omega\right] = x_i^m\left(b_A^m(\omega), b_B^m(\omega)\right)$. In particular, two important remarks follow by well-known results applied to our setup.

Remark 1. *Under* (1) *with r sufficiently low, equilibrium efforts are* deterministic *in any battle and the expected effort of player $i \in \{A, B\}$ conditional on her winning the battle is* identical to *her expected effort conditional on losing.*

Remark 2. *Under* (2), *equilibrium efforts are* stochastic *in any battle and the expected effort of player $i \in \{A, B\}$ conditional on her winning the battle is* higher than *her expected effort conditional on losing.*

To illustrate WE^T in (3), consider $T = \{1, 1, 1\}$, so that the third battle might be trivial—e.g., $b_A^3(A, A) = 2$ as team A won the first two battles (when $\omega = A, A$) and thus $E[x_A^3(b_A^3(\omega), b_B^3(\omega)) | \omega = A, A] = 0$. Recalling from preliminary equilibrium observation 2 that p_i is the constant equilibrium win probability of team $i \in \{A, B\}$ in a non-trivial battle, then $\Pr\{\omega\} = p_A^j p_B^k$ in (3), where j is the number of A's in ω and k is the number of B's in ω; e.g., $\Pr\{B, A, A\} = p_A^2 p_B$. Hence, (3), specialized to $T = \{1, 1, 1\}$, reduces to

$$\begin{aligned} WE^{\{1,1,1\}} = & \, p_A^2 \cdot E\left[x_A^1(0,0) + x_A^2(1,0) | \omega = A, A\right] \\ & + p_A^2 p_B \cdot E\left[x_A^1(0,0) + x_A^2(1,0) + x_A^3(1,1) | \omega = A, B, A\right] \\ & + p_A^2 p_B \cdot E\left[x_A^1(0,0) + x_A^2(0,1) + x_A^3(1,1) | \omega = B, A, A\right] \\ & + \, p_B^2 \cdot E\left[x_B^1(0,0) + x_B^2(0,1) | \omega = B, B\right] \\ & + p_A p_B^2 \cdot E\left[x_B^1(0,0) + x_B^2(0,1) + x_B^3(1,1) | \omega = B, A, B\right] \\ & + p_A p_B^2 \cdot E\left[x_B^1(0,0) + x_B^2(1,0) + x_B^3(1,1) | \omega = A, B, B\right]. \end{aligned}$$

The first three lines of $WE^{\{1,1,1\}}$ account for $\Omega_A^{\{1,1,1\}}$, the remaining ones for $\Omega_B^{\{1,1,1\}}$. For instance, the first line considers $\omega = A, A$; the term p_A^2 accounts for the probability of $\omega = A, A$ and the remaining term accounts for the expected efforts exerted by players of team A only, as team A wins if $\omega = A, A$. Such an expectation is taken conditional on team A winning the first two battles. If equilibrium efforts are deterministic, the conditioning does not matter (Remark 1). If equilibrium efforts are stochastic, the conditioning matters and is affected by both teams' mixed strategies and the CSF; therefore, the details are equilibrium-dependent and elaborated on below (see, e.g., Lemma 5).

Finally, note that our Results 1–4 on WE can be immediately extended to a variety of objectives that are linear combinations of TE, WE, and the losers' effort LE. To see this more formally, let the maximand be $\alpha_W \cdot WE + \alpha_L \cdot LE + \alpha_T \cdot TE$, where α_W, α_L, and α_T are constants. Since $LE = TE - WE$ and TE is unaffected by the temporal structure (see FLP), the maximand simplifies to $(\alpha_W - \alpha_L) \cdot WE$ and our results apply straightforwardly. Nevertheless, for the sake of simplicity, we focus on WE alone throughout the remainder of the paper.

3 Pure-strategy equilibrium

In this section, we analyze the case of the CSF (1) with r "sufficiently low": as the prize is common in every battle m and $c_A^m/c_B^m = c_A/c_B$ for any m, then for any pair (c_A, c_B) there is an $\bar{r}(c_A, c_B) \in [1, 2]$ such that, for $r \leq \bar{r}(c_A, c_B)$ the equilibrium is in pure strategies in every battle. Wang (2010) shows $\bar{r}(c_A, c_B)$ is the unique $\bar{r}$ solving

$$\frac{1}{(\bar{r}-1)} = \left(\frac{\max\{c_A, c_B\}}{\min\{c_A, c_B\}}\right)^{\bar{r}}. \tag{4}$$

Throughout this section, we assume $r \leq \bar{r}(c_A, c_B)$ and leave the case $\bar{r}(c_A, c_B) < r \leq 2$ to Section 5. The assumption $r \leq \bar{r}(c_A, c_B)$ simplifies the analysis by Remark 1. If $c_A = c_B$, $\bar{r}(c_A, c_B) = 2$, so the equilibrium is in pure strategies for any $r \in (0, 2]$.

Main results. Using extractions and mergers, we prove that, if $c_A \neq c_B$, WE is minimized by a fully sequential contest and maximized by a fully simultaneous contest. An *extraction* is a process by which, from the first non-unitary n_j (the first round with multiple battles) of a starting temporal structure, one battle is extracted and fought on its own right before the remaining $n_j - 1$ battles; for instance, in $\{1, 4, 1, 3\}$ the first non-unitary n_j is $n_2 = 4$, and thus an extraction yields $\{1, 1, 3, 1, 3\}$. Proposition 1 proves that an extraction does not affect WE only if $c_A = c_B$ or we are comparing temporal structures $\{1, 1, ..., 1, 2\}$ and $\{1, 1, ..., 1\}$. In all other cases, an extraction strictly decreases WE. Starting from any arbitrary temporal structure, a sufficient number of extractions yields the fully sequential contest. Hence, WE is minimized, among all possible temporal structures, by a fully sequential contest.

A *merger* is a process by which the first two rounds, with n_1 and n_2 battles respectively, are merged into one round with $n_1 + n_2$ simultaneous battles; e.g., starting from $\{1, 4, 1, 3\}$ a merger yields $\{5, 1, 3\}$. Proposition 2 proves that a merger does not affect WE only if $c_A = c_B$. In all other cases, a merger strictly increases WE. Starting from any temporal structure, a sufficient number of mergers yields the fully simultaneous contest. Hence, WE is maximized, among all possible temporal structures, by a fully simultaneous contest.

3.1 Illustrative example; extractions and mergers

This section considers a simple example that illustrates extractions, mergers, and the intuition behind the differences between TE's temporal-structure independence and WE's temporal-structure dependence. In particular, we consider a best-of-three contest with Tullock CSF and identical players within teams (i.e., $n = 1$, $\gamma_1 = \gamma_2 = \gamma_3 = 1$, and (1) with $r = 1$). The only difference between teams is that, without loss of generality, team A is

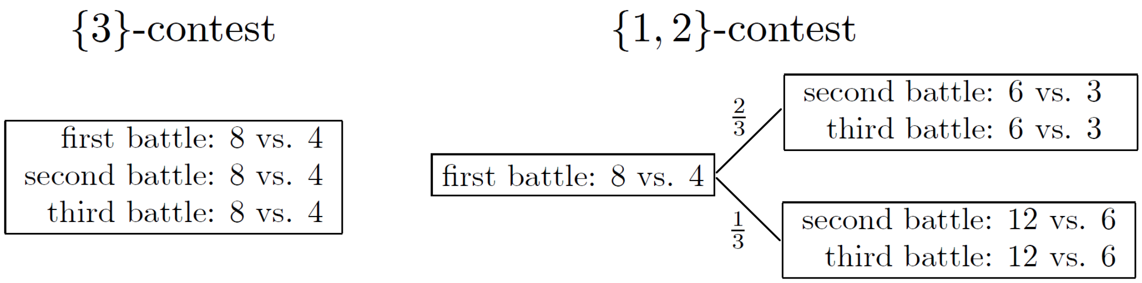

Figure 1: Equilibrium efforts in the $\{3\}$- and $\{1,2\}$-contests with $n = 1$, $c_A = 1/81$, $c_B = 2/81$, $\gamma_1 = \gamma_2 = \gamma_3 = 1$. The calculations follow by standard techniques. WE is 564/27 for the $\{3\}$-contest and 548/27 for the $\{1,2\}$-contest. One can also calculate that WE is 556/27 in a $\{2,1\}$-contest, and 548/27 in a $\{1,1,1\}$-contest.

stronger than team B: specifically, we assume that $c_A = 1/81$, $c_B = 2/81$, and therefore obtain $p_A = 2/3$ in every battle (recall the preliminary equilibrium observation 2). Figure 1 depicts equilibrium efforts when comparing temporal structures $\{3\}$ and $\{1,2\}$, highlighting both an extraction from $\{3\}$ to $\{1,2\}$ and a merger from $\{1,2\}$ to $\{3\}$.

We next explain why TE is identical in $\{3\}$ and $\{1,2\}$ whereas WE is larger in $\{3\}$ than in $\{1,2\}$. Note that $\Omega_A^{\{3\}} = \{AAA; AAB; ABA; BAA\}$, $\Omega_A^{\{1,2\}} = \{A, AA; A, AB; A, BA; B, AA\}$, $\Omega_B^{\{3\}} = \{BBB; BBA; BAB; ABB\}$, and $\Omega_B^{\{1,2\}} = \{B, BB; B, BA; B, AB; A, BB\}$. By Remark 1, the conditioning does not affect expected efforts in (3) and we can write

$$WE^{\{3\}} = \left(p_A^3 + 3p_A^2 p_B\right) \cdot (8+8+8) + \left(p_B^3 + 3p_A p_B^2\right) \cdot (4+4+4), \tag{5}$$

where $p_A^3 + 3p_A^2 p_B = \sum_{\omega \in \Omega_A^{\{3\}}} \Pr\{\omega\}$ and $p_B^3 + 3p_A p_B^2 = \sum_{\omega \in \Omega_B^{\{3\}}} \Pr\{\omega\}$. As the contest is simultaneous, efforts are identical across battles. Also, we can write

$$\begin{aligned} WE^{\{1,2\}} = \quad & (p_A^3 + 2p_A^2 p_B) \cdot (8+6+6) + p_A^2 p_B \cdot (8+12+12) \\ & + (p_B^3 + 2p_A p_B^2) \cdot (4+6+6) + p_A p_B^2 \cdot (4+3+3), \end{aligned} \tag{6}$$

where we have that $p_A^3 + 2p_A^2 p_B = \sum_{\omega \in \Omega_A^{\{1,2\}} \setminus B,AA} \Pr\{\omega\}$ and $p_A^2 p_B = \Pr\{B, AA\}$ for $\Omega_A^{\{1,2\}}$ in the first line of (6), and $p_B^3 + 2p_A p_B^2 = \sum_{\omega \in \Omega_B^{\{1,2\}} \setminus A,BB} \Pr\{\omega\}$ and $p_A p_B^2 = \Pr\{A, BB\}$ for $\Omega_B^{\{1,2\}}$ in the second line of (6). Note that the second and third battle efforts depend on the first battle outcome and so in (6) we distinguish whether A or B wins the first battle.

A first key feature that is clear from Figure 1 is that the efforts of the first battle are identical in the two temporal structures—i.e., 8 for the A-player and 4 for the B-player. To see this, recall the preliminary equilibrium observation 2; in particular, since $c_B = 2c_A$, under (1) with $r = 1$, $p_A = 2/3$ and $p_B = 1/3$ *in every battle of both temporal structures*. Consider now the first battle of the $\{1,2\}$-contest; each player fights for a prize equal to the difference between the winning probability of one's own team if the battle is won and that if the battle is lost. This difference is $4/9$ both for the A-player $(8/9 - 4/9)$ and the B-player $(5/9 - 1/9)$.[5] In the first battle of the $\{3\}$-contest, each player fights for a prize equal to the probability that this battle is pivotal, which also equals $4/9$ for both players.[6] Thus, in the first battle of the $\{3\}$- and $\{1,2\}$-contests, players are fighting for the same prize, and hence they exert the same effort.

A second key feature that is clear from Figure 1 is that, within each temporal structure, the efforts of the second battle equal those of the third battle. These two key features allow us to consider only the third battle when comparing TE or WE in the $\{3\}$- and $\{1,2\}$-contests. The simplification arising from these two key features can also be seen comparing common terms in (5) and (6): the first-battle efforts (8 and 4) are identical and have identical probability weights across temporal structures and can thus be ignored, and the second- and third-battle efforts are always repeated twice within each temporal structure and thus considering only one of them suffices.

To conclude the example, we now show that the comparison of third-battle efforts yields: *1)* identical TE in the $\{3\}$- and $\{1,2\}$-contests, and *2)* $WE^{\{3\}} > WE^{\{1,2\}}$. The upcoming reasoning can be followed with the help of Figure 1 and/or expressions (5) and (6).

1) Consider TE. In the $\{1,2\}$-contest, the low effort (which equals 6) of the A-player is due to the third battle's low pivotal probability (which equals 1/3), while the high effort (which equals 12) of the A-player is due to the third battle's high pivotal probability (which equals 2/3). Moreover, moving from the $\{3\}$- to the $\{1,2\}$-contest, the effort of the third battle of the A-player is lower (from 8 to 6) with probability 2/3, and higher (from 8 to 12) with probability 1/3. These two opposite forces perfectly cancel out in TE by the law of iterated expectations: $(2/3) \cdot 6 + (1/3) \cdot 12 = 8$. The same applies to the effort of the third battle of the B-player. Thus, TE is identical in the $\{3\}$- and $\{1,2\}$-contests. This is the temporal-structure independence of TE that FLP show under great generality.

2) Consider WE. Now, each player's effort is weighed by the probability that her team wins the overall contest (while in TE efforts are weighed equally across teams). First, focus on the A-player fighting the third battle in the $\{1,2\}$-contest; her team wins the overall contest with a relatively high probability when her team won the first battle ($\sum_{\omega \in \Omega_A^{\{1,2\}} \setminus B,AA} \Pr\{\omega\} =$

[5] If the A-player wins [loses] the first battle of the $\{1,2\}$-contest, team A wins the contest unless [only if] both the second and third battles are won by B [A], which occurs with probability $1 - (1/3)^2 = 8/9$ $[(2/3)^2 = 4/9]$. Similar calculations apply to B.

[6] The first battle of the $\{3\}$-contest is pivotal only if the second and third battles are won by different teams, which happens with probability $2\,(1/3)\,(2/3) = 4/9$.

$\frac{16}{27}$), and in this case her effort is lower than in the $\{3\}$-contest ($6 < 8$). Similarly, her team wins the overall contest with a relatively low probability when her team lost the first battle ($\Pr\{B, AA\} = \frac{4}{27}$), and in this case her effort is higher than in the $\{3\}$-contest ($12 > 8$). Hence, focusing only on the A-player fighting the third battle, high weight to the low effort and low weight to the high effort make WE *decrease* when moving from $\{3\}$ to $\{1, 2\}$. Second, focus now on the B-player fighting the third battle in the $\{1, 2\}$-contest; her team wins the overall contest with a relatively high probability when her team won the first battle ($\sum_{\omega \in \Omega_B^{\{1,2\}} \setminus A,BB} \Pr\{\omega\} = \frac{5}{27}$), and in this case her effort is higher than in the $\{3\}$-contest ($6 > 4$). Similarly, her team wins the overall contest with a relatively low probability when her team lost the first battle ($\Pr\{A, BB\} = \frac{2}{27}$), and in this case her effort is lower than in the $\{3\}$-contest ($3 < 4$). Hence, focusing only on the B-player fighting the third battle, high weight to the high effort and low weight to the low effort make WE *increase* when moving from $\{3\}$ to $\{1, 2\}$. Which effect dominates between the negative effect on WE of the A-player and the positive effect on WE of the B-player? The answer is the former and the reason is that the A-player's effort and probability of victory are larger than those of the B-player. Therefore, WE decreases when moving from $\{3\}$ to $\{1, 2\}$; an extraction (merger) decreases (increases) WE. We call this force the *team-asymmetry effect*; as soon as teams' marginal costs differ ($c_A \neq c_B$), the battles' pivotal probability and the probability of overall victory of the stronger team (whose members exert larger efforts) tend to move in opposite directions when early battle outcomes are revealed. The more sequential a contest, the larger the negative impact of the team-asymmetry effect on WE, because it applies "more often" (i.e., after every round). Conversely, the more simultaneous a contest, the milder the negative impact of the team-asymmetry effect on WE. The stylized reasoning of the above illustrative example can be extended to show that, regardless of which team is stronger, WE is maximized by a fully simultaneous contest and minimized by a fully sequential contest, despite TE being invariant to the temporal structure.

3.2 Equilibrium efforts

Sections 3.2-3.4 show that all the qualitative features of the example in Section 3.1 hold for any pair (c_A, c_B), any temporal structure T, any vector of γ_m's, any extraction or merger, and any CSF (1) with $r \leq \bar{r}(c_A, c_B)$ (i.e., with a pure-strategy equilibrium in each battle).

We begin this section by introducing the notation to describe the overall probability of victory. Recall that the state (b_A, b_B) identifies the number of battles already won by each team, with $b_A, b_B \leq n$ for non-trivial battles. Assume first that all battles from state (b_A, b_B) onwards are simultaneous and non-trivial. Since the equilibrium probability that team A

(B) wins each non-trivial battle is p_A (p_B), the equilibrium probability of victory of team A, given b_A victories for team A and b_B victories for team B, is

$$P_A(b_A, b_B) = \sum_{i=n+1-b_A}^{2n+1-(b_A+b_B)} \binom{2n+1-(b_A+b_B)}{i} p_A^i p_B^{2n+1-(b_A+b_B)-i}. \tag{7}$$

Note now that the probability resulting from formula (7) actually applies to all future possible temporal structures, including the (partially) sequential ones that may have trivial battles (by result (ii) of Theorem 2 in FLP and their discussion on team's head starts in Section III.B); this is the preliminary equilibrium observation 3. For instance, in temporal structure $\{1, 2\}$, battle 3 is never trivial and hence $P_A(1, 0) = p_A(p_A + p_B) + p_B p_A = p_A(1 + p_B)$, while in temporal structure $\{1, 1, 1\}$ battle 3 may become trivial, but $P_A(1, 0)$ still equals $p_A + p_B p_A = p_A(1 + p_B)$. For team B, $P_B(b_A, b_B) = 1 - P_A(b_A, b_B)$.

The equilibrium pivotal probability of non-trivial battle $b_A + b_B + 1$, which coincides with battle prize $\nu(b_A, b_B)$, is then

$$\begin{aligned} P_A(b_A+1, b_B) - P_A(b_A, b_B+1) &= P_B(b_A, b_B+1) - P_B(b_A+1, b_B) \\ &= \binom{2n-(b_A+b_B)}{n-b_A} p_A^{n-b_A} p_B^{n-b_B}. \end{aligned} \tag{8}$$

Such a pivotal probability will determine the equilibrium efforts of each battle, together with a multiplicative term that depends on the specific CSF we consider. In a contest with (1) and $r \leq \bar{r}(c_A, c_B)$, such multiplicative terms are

$$\mu_A^m \equiv r \frac{c_A^{r-1} c_B^r}{(c_A^r + c_B^r)^2 \gamma_m} \text{ and } \mu_B^m \equiv r \frac{c_A^r c_B^{r-1}}{(c_A^r + c_B^r)^2 \gamma_m}. \tag{9}$$

Indeed, standard equilibrium characterization results (see, e.g., Observation 2 and Lemma 1 of FLP) show that efforts in the n_j battles of round j are described by the following lemma.

Lemma 1. *Let (b_A, b_B) be the state at the beginning of round j with $b_A, b_B \leq n$ so that battle $m \in \{b_A + b_B + 1, ..., b_A + b_B + n_j\}$ is non-trivial. The equilibrium individual efforts in battle m are*

$$x_i^m(b_A, b_B) = \mu_i^m \cdot \binom{2n-(b_A+b_B)}{n-b_A} p_A^{n-b_A} p_B^{n-b_B} \text{ with } i \in \{A, B\}, \tag{10}$$

where the values of μ_i^m are in (9) and

$$p_A = \frac{c_B^r}{c_A^r + c_B^r} \text{ and } p_B = \frac{c_A^r}{c_A^r + c_B^r}. \tag{11}$$

Note that (9)-(11) show that each player of the stronger team—i.e., that with a lower marginal cost of effort—exerts larger effort and has a larger probability of winning the non-trivial battle than the matched player of the weaker team.

The approach of using Lemma 1 to derive and compare WE in (3) for every possible T proves not tractable and hence we next take a different tack: extractions and mergers.

3.3 Extractions decrease WE

We first provide the formal definition of an extraction.

Definition 1. *Consider a temporal structure $T = \{n_1, n_2, ...\}$ different from the fully sequential. Let $l \equiv \min\{j : n_j > 1, n_j \in T\}$; that is, l is the earliest round with multiple battles. The **extraction** from $T = \{n_1, n_2, ..., n_{l-1}, n_l, n_{l+1}, ...\}$ yields $T^{ext} \equiv \{n_1, ..., n_{l-1}, 1, n_l - 1, n_{l+1}, ...\}$.*

Using Definition 1, if $T = \{1, 3, 5\}$, then $T^{ext} = \{1, 1, 2, 5\}$, $n_l = 3$, and $l = 2$. We show extractions decrease WE through several intermediate steps.

First, by the preliminary equilibrium observations, extractions do not affect efforts of the first $l - 1$ battles (i.e., battles with index $m \leq l - 1$) and the last $2n + 1 - (l - 1) - n_l$ battles (i.e., with index $m \geq (l - 1) + n_l$). Consequently, these battles can be disregarded when comparing WE in T and T^{ext}. If $T = \{1, 3, 5\}$, the first battle can be disregarded as it occurs with certainty and has the same efforts in T and T^{ext}, and the last five battles can be disregarded too, as the possible realizations of the first four battles have the same probabilities in T and T^{ext}.

Second, at the one battle immediately after n_{l-1} (i.e., battle l), Lemma 1 implies identical efforts across players in T and T^{ext}, since the first $l-1$ battles are not affected by the extraction, and it does not matter whether battles from $l + 1$ to $2n + 1$ are fought simultaneously with or after battle l. Thus, battle l can be disregarded when comparing WE in T and T^{ext}. If $T = \{1, 3, 5\}$, the efforts of the second battle are identical across players in T and T^{ext}.

We store the two results above in the following corollary of Lemma 1.

Corollary 1. *When comparing WE in temporal structures T and T^{ext}, the first l and the last $2n + 1 - (l - 1) - n_l$ battles can be disregarded.*

Recall that, in the example of Section 3.1, we ignored the first battle since efforts and team probabilities of winning are identical in the $\{3\}$- and $\{1, 2\}$-contests; this analytical shortcut is generalized by Corollary 1. In the example of Section 3.1, as $n = l = 1$ and $n_l = 3$, Corollary 1 does not allow us to disregard any battle other than the first one (as $2n + 1 - (l - 1) - n_l = 0$).

Note that any extraction can only switch battles from being non-trivial to trivial. Furthermore, such a switch can only happen in the $n_l - 1$ battles of round l of T that are not extracted and end up being played in round $l+1$ of T^{ext}: these are the battles not mentioned in Corollary 1. Thus, in the proofs of the results of this section, we account for extractions that may change whether these battles are trivial.

The rest of this section generalizes the comparison of (second and) third battle efforts in the example of Section 3.1. Lemma 2, with proof in the Appendix, shows that, using the simplifications of Corollary 1, the effect of an extraction on the winner's effort is determined by the terms μ_A^m and μ_B^m in (9) for $m \in \{l+1, ..., l+n_l-1\}$ and by the behavior of the function

$$\psi(z) = \sum_{b_A=\max\{z-n,0\}}^{\min\{n,z\}} \binom{z}{b_A} \binom{2n-z}{n-b_A} \cdot P_A(b_A, z-b_A), \tag{12}$$

where $P_A(\cdot, \cdot)$ is defined in (7) with p_A and p_B in (11).

Lemma 2. *WE in T is larger (equal) [smaller] than in T^{ext} if*

$$\psi(l-1) \sum_{m=l+1}^{l+n_l-1} (\mu_A^m - \mu_B^m) > (=) [<] \psi(l) \sum_{m=l+1}^{l+n_l-1} (\mu_A^m - \mu_B^m).$$

Finally, Lemma 3, with proof in the Appendix, shows when $\psi(\cdot)$ is increasing or decreasing.

Lemma 3. *If $l = 2n$, then $\psi(l-1) = \psi(l)$. Otherwise, if $p_A > (=) [<] p_B$, then we obtain $\psi(l-1) > (=) [<] \psi(l)$.*

Lemmas 2 and 3 jointly characterize the effects of an extraction on WE. In particular,

1. if $l = 2n$ (i.e., $T = \{1, 1, ..., 1, 2\}$ and $T^{ext} = \{1, 1, ..., 1\}$), then $\psi(l-1) = \psi(l)$ by Lemma 3, and hence Lemma 2 shows that WE is constant after the extraction.

2. if $c_A = c_B$, which by (9) yields $\mu_A^m = \mu_B^m$ $\forall m$, then Lemma 2 implies that WE is constant after the extraction.

3. if $c_A < c_B$, which by (9) yields $\mu_A^m > \mu_B^m$ $\forall m$, then Lemma 2 shows that WE decreases after the extraction if and only if $\psi(l-1) > \psi(l)$. Lemma 3 shows that this is the case (unless $l = 2n$), since $c_A < c_B$ implies that $p_A > p_B$ by (11). The case of $c_A > c_B$ is analogous.

The above discussion proves the main result of this section: extractions decrease WE. This decrease is strict unless $l = 2n$ or there is no team-asymmetry effect ($c_A = c_B$). Starting from any arbitrary temporal structure, a sufficient number of extractions yields the fully sequential contest, and thus WE is minimized by a fully sequential contest.

Proposition 1. *Consider a team contest à la FLP described in Section 2 with CSF (1) and* $r \leq \bar{r}(c_A, c_B)$. *If* $c_A = c_B$, *extractions do not affect* WE. *If* $c_A \neq c_B$, *the extraction from* $\{1, 1, ..., 1, 2\}$ *to* $\{1, 1, ..., 1\}$ *(i.e.,* $l = 2n$*) does not affect* WE*; any other extraction strictly decreases* WE. *Therefore, if* $c_A \neq c_B$, *among all possible temporal structures there are two* WE*-minimizing temporal structures: the fully sequential contest and* $\{1, 1, ..., 1, 2\}$.

Our results in Proposition 1 are also in line with the illustrative example discussed in Section 3.1. The WE-reducing team-asymmetry effect applies at the end of every round. An extraction adds one round to the contest in an analytically tractable way. Thus extractions exacerbate the team-asymmetry effect and reduce WE.

3.4 Mergers increase WE

We first provide the formal definition of a merger.

Definition 2. *Consider a temporal structure* $T \equiv \{n_1, n_2, n_3, ...\}$, *different from the fully simultaneous. The* ***merger*** *from* T *yields* $T^{mer} \equiv \{n_1 + n_2, n_3, ...\}$.

Using Definition 2, if $T = \{3, 1, 5\}$, then $T^{mer} = \{4, 5\}$, $n_1 = 3$, and $n_2 = 1$. We show mergers increase WE through several intermediate steps.

First, by the preliminary equilibrium observations, mergers do not affect efforts in battles after the first $n_1 + n_2$ (battles with $m > n_1 + n_2$), or the probability that each later battle is played. Consequently, these battles can be disregarded when comparing WE in T and T^{mer}. If $T = \{3, 1, 5\}$, the battles from the fifth to the ninth can be disregarded as the possible realizations of the first four battles have the same probabilities in T and T^{mer}.

Second, for the first n_1 battles, Lemma 1 implies identical efforts across players in T and T^{mer}. If $T = \{3, 1, 5\}$, the first three battles can be disregarded as, in both T and T^{mer}, they occur with certainty and have the same efforts.

We store the two results above in the following corollary of Lemma 1.

Corollary 2. *When comparing* WE *in temporal structures* T *and* T^{mer}, *the first* n_1 *battles and all battles after the first* $n_1 + n_2$ *can be disregarded.*

If $T = \{3, 1, 5\}$, we can ignore the first three battles and the last five battles; and thus we can focus on the fourth battle only. The fourth battle is the only battle that can have different efforts in T and $T^{mer} = \{4,5\}$, as it is fought simultaneously with the first three battles in T^{mer}, but after the first three battles in T.

Note that any merger can only switch battles from being trivial to non-trivial; furthermore, such a switch can only happen in the n_2 battles of the second round of T that are merged in the first round of T^{mer}; these are the battles not mentioned in Corollary 2. For instance, considering a merger from $\{3,2\}$ to $\{5\}$, the fourth and fifth battles become certainly non-trivial because of the merger. Thus, in the proofs of the results of this section, we account for mergers that may change whether these battles are trivial.

Paralleling Section 3.3, using the simplifications of Corollary 2, the effect of a merger on WE is determined by the behavior of ψ defined in (12). The following lemma, with proof in the Appendix, mirrors Lemma 2.

Lemma 4. *WE in T^{mer} is larger (equal) [smaller] than in T if and only if*

$$\psi(0) \sum_{m=n_1+1}^{n_1+n_2} \left(\mu_A^m - \mu_B^m\right) > (=) [<] \, \psi(n_1) \sum_{m=n_1+1}^{n_1+n_2} \left(\mu_A^m - \mu_B^m\right).$$

Lemmas 3 and 4 jointly characterize the effects of a merger on WE. In particular,

1. if $c_A = c_B$, which by (9) yields $\mu_A^m = \mu_B^m \; \forall m$, then Lemma 4 implies that WE is constant after the merger.

2. if $c_A < c_B$, which by (9) yields $\mu_A^m > \mu_B^m \; \forall m$, then Lemma 4 shows that WE increases after the merger if and only if $\psi(0) > \psi(n_1)$. Applying $n_1 - 1$ times Lemma 3 shows that $\psi(0) > \psi(n_1)$, since $c_A < c_B$ implies that $p_A > p_B$ by (11). The case of $c_A > c_B$ is analogous.

The above discussion proves the main result of this section: mergers increase WE. This increase is strict unless there is no team-asymmetry effect ($c_A = c_B$). Starting from any arbitrary temporal structure, a sufficient number of mergers yields the fully simultaneous contest, and thus WE is maximized by a fully simultaneous contest.

Proposition 2. *Consider a team contest à la FLP described in Section 2 with CSF (1) and $r \leq \bar{r}(c_A, c_B)$. If $c_A = c_B$, mergers do not affect WE. If $c_A \neq c_B$, a merger strictly increases WE. Therefore, if $c_A \neq c_B$, among all possible temporal structures there is one WE-maximizing temporal structure: the fully simultaneous contest.*

Our results in Proposition 2 are also in line with the illustrative example discussed in Section 3.1. The WE-reducing team-asymmetry effect applies at the end of every round. A merger eliminates one round of the contest in an analytically tractable way. Thus mergers mitigate the team-asymmetry effect and increase WE.

Extractions and mergers need not be each other's inverse, and hence neither Proposition 2 implies Proposition 1, nor vice versa. This can be seen by considering $T = \{2, 1\}$; a merger followed by an extraction yields $\{1, 2\}$ rather than $\{2, 1\}$.

4 Mixed-strategy equilibrium

In this section, we analyze the all-pay auction CSF (2), so that the equilibrium is in mixed strategies due to the lack of noise.

4.1 Illustrative example

We now focus on $c_A = c_B$ to abstract from the team-asymmetry effect of Section 3. This allows us to highlight the new force, the stochastic-effort effect, that arises because equilibrium efforts are stochastic, rather than deterministic as in Section 3. For simplicity, we assume $\gamma_1 = \gamma_2 = \gamma_3 = c_A = c_B = 1$.

Consider temporal structures $\{3\}$ and $\{2, 1\}$. Standard calculations show that, in the first and second battle of both temporal structures and in the third battle of $\{3\}$, players mix uniformly on $[0, 1/2]$. The third battle of $\{2, 1\}$ is non-trivial only when the state is $(1, 1)$ and, if so, players mix uniformly on $[0, 1]$.

We now compare $WE^{\{3\}}$ and $WE^{\{2,1\}}$. The first two battles can be neglected in this comparison (as in Section 3.1) by the same reasoning behind Corollary 2. The third battle is crucially different between the two temporal structures: in $\{2, 1\}$ the third battle is always a tie-break if actually fought, and hence its winner wins also the overall contest, whereas in $\{3\}$ the third battle is won by a player belonging to the team that loses the overall contest when $\omega = AAB$ or $\omega = BBA$.

The contribution of the third battle to $WE^{\{3\}}$ is

$$\sum_{i \in \{A,B\}} \sum_{\omega \in \Omega_i^{\{3\}}} \Pr\{\omega\} E\left[x_i^3(0,0)\,|\omega\right] = 2 \cdot \frac{1}{8} \cdot \left(3 \cdot \frac{1}{3} + 1 \cdot \frac{1}{6}\right). \tag{13}$$

The two teams are identical, which explains the 2 at the beginning of the RHS of (13). Also, $\Pr\{\omega\} = 1/8 \ \forall \omega \in \Omega^{\{3\}}$ as, by $c_A = c_B$, $p_A = p_B = 1/2$ in every battle of any temporal structure. As $\Omega_A^{\{3\}} = \{AAA; AAB; ABA; BAA\}$, in three cases team A wins the overall contest *and* the third battle, while for $\omega = AAB$ team A wins the overall contest *but* loses the third battle. Finally, the expectation of the maximum (minimum) of two uniform distributions on $[0, 1/2]$ is 1/3 (1/6).

As for $\{2,1\}$, we have $\Omega_A^{\{2,1\}} = \{AA; AB, A; BA, A\}$ and $\Omega_B^{\{2,1\}} = \{BB; BA, B; AB, B\}$. As the third battle is actually fought only when the state is $(1,1)$ (in case of tie), the contribution of the third battle to $WE^{\{2,1\}}$ is

$$\sum_{\omega\in\{AB,A;BA,A\}} \Pr\{\omega\}E[x_A^3(1,1)|\omega] + \sum_{\omega\in\{BA,B;AB,B\}} \Pr\{\omega\}E[x_B^3(1,1)|\omega] = 2\cdot\frac{1}{8}\cdot\left(2\cdot\frac{2}{3}\right). \quad (14)$$

The key difference in (14) with respect to (13) is that now the third battle is always won by the team that wins the overall contest. Formally, 2/3 is the expectation of the maximum of two uniform distributions on $[0, 1]$ and its minimum now plays no role as the effort of the third-battle loser is never part of the efforts of the team that wins the overall contest. This is in contrast to the 1/6 in (13). Clearly, (14) is greater than (13): hence, $WE^{\{3\}} < WE^{\{2,1\}}$.

The main message of the above comparison is that, when equilibrium efforts are stochastic, there is an extra force which is absent when equilibrium efforts are deterministic: the effect on WE of having *battle-losers in the team that wins the overall contest*. This effect is negative by Remark 2. More sequential temporal structures mitigate such a negative effect; e.g., the third battle is always won by a member of the team that wins the overall contest in $\{2,1\}$, but not necessarily in $\{3\}$. This new force is the *stochastic-effort effect* and it only arises when equilibrium efforts are stochastic (as under (2)), but not when equilibrium efforts are deterministic (as under (1) with small r). In fact, in our illustrative example, we obtain that

$$WE^{\{3\}} = WE^{\{1,2\}} = \frac{7}{8} < \frac{11}{12} = WE^{\{2,1\}} = WE^{\{1,1,1\}}. \quad (15)$$

Note that in temporal structures $\{2,1\}$ and $\{1,1,1\}$ the third battle, if actually fought, cannot be lost by the winner of the overall contest; this is instead possible in temporal structures $\{3\}$ and $\{1,2\}$, which yield a lower WE.

Therefore, the stochastic-effort effect favors more sequential temporal structures. In contrast, the team-asymmetry effect highlighted in the deterministic setup of Section 3 favors less sequential temporal structures. As we explained in Section 3.1, when $c_A = c_B$ the team-asymmetry effect is muted. Instead, when $c_A \neq c_B$, the team-asymmetry effect kicks in and one can show that, keeping $\gamma_1 = \gamma_2 = \gamma_3 = 1$, the combination of team-asymmetry and stochastic-effort effects results in the following ranking of temporal structures:

$$WE^{\{1,2\}} < WE^{\{3\}} < WE^{\{1,1,1\}} < WE^{\{2,1\}}. \quad (16)$$

To understand (16), start from (15) which holds under $c_A = c_B$. As soon as $c_A \neq c_B$, the two equalities ($WE^{\{3\}} = WE^{\{1,2\}}$ and $WE^{\{2,1\}} = WE^{\{1,1,1\}}$) are broken in favor of the less sequential temporal structures, thus yielding (16). This can be seen in light of the intuition developed for the illustrative example of Section 3.1 (extractions reduce WE); in fact, $\{1,2\}$ is an extraction from $\{3\}$, and $\{1,1,1\}$ is an extraction from $\{2,1\}$.

4.2 Equilibrium efforts

Recall that, by the preliminary equilibrium observations, the equilibrium win probabilities (p_A, p_B) in a non-trivial battle are identical across battles and unaffected by the temporal structure. As functions of p_A, the expression for $P_A(b_A, b_B)$ (see (7)) and that of the battle prize $\nu(b_A, b_B)$ (see (8)) carry over. What changes from the setup with pure-strategy equilibrium to the present one with mixed-strategy equilibrium is that we now need four terms for conditional expected equilibrium efforts,

$$\mu_{A|A}^m \equiv \frac{1}{3c_B\gamma_m}\frac{3c_B - c_A}{2c_B - c_A},\ \mu_{A|B}^m \equiv \frac{1}{3c_B\gamma_m},\ \mu_{B|A}^m \equiv \frac{1}{3c_B\gamma_m}\frac{c_A}{2c_B - c_A},\ \mu_{B|B}^m \equiv \frac{2}{3c_B\gamma_m}, \quad (17)$$

where we assume, without loss of generality, that $c_A \leq c_B$. The reason why we now need twice as many terms as in the deterministic setup is clear from remarks 1 and 2.

The expected conditional efforts in the n_j battles of round j (with $n_j \in T \equiv \{n_1, n_2, ...\}$) are described by the following lemma, with proof in the Appendix.

Lemma 5. *Let (b_A, b_B) be the state at the beginning of round j with $b_A, b_B \leq n$ so that battle $m \in \{b_A + b_B + 1, ..., b_A + b_B + n_j\}$ is non-trivial. The* expected *equilibrium individual efforts in, conditional on the outcome of battle m, are*

$$E[x_i^m(b_A, b_B)|j \text{ wins battle } m] = \mu_{i|j}^m \binom{2n - (b_A + b_B)}{n - b_A} p_A^{n-b_A} p_B^{n-b_B} \text{ with } i, j \in \{A, B\}, \quad (18)$$

where the values of $\mu_{i|j}^m$ are in (17) *and those of p_A and p_B are*

$$p_A = \frac{2c_B - c_A}{2c_B} \text{ and } p_B = \frac{c_A}{2c_B}. \quad (19)$$

4.3 Symmetric teams: extractions increase WE and mergers decrease WE

In this section, we consider the case of symmetric teams ($c_A = c_B$) so as to single out the stochastic-effort effect and abstract from the team-asymmetry effect. We show that extractions (weakly) increase and mergers (weakly) decrease WE, in contrast with what we obtained in Section 3 when the CSF was (1) with r sufficiently small.

As for extractions, we obtain the following proposition, with proof in the Appendix.

Proposition 3. *Consider a team contest à la FLP described in Section 2 with CSF (2) and $c_A = c_B$. An extraction from $\{n_1, n_2, ..\}$ to $\{1, n_1 - 1, n_2, ..\}$ with $n_1 \geq 2$ (i.e., $l = 1$) does not affect WE. Any other extraction strictly increases WE. Therefore, among all possible temporal structures there are two WE-maximizing temporal structures: the fully sequential contest and $\{2, 1, ..., 1\}$.*

The ranking in (15) is consistent with the results of Proposition 3. Indeed, $\{1, 2\}$ is an extraction from $\{3\}$ with $n_1 = 3$ and hence $WE^{\{3\}} = WE^{\{1,2\}}$, and $\{1, 1, 1\}$ is an extraction from $\{2, 1\}$ with $n_1 = 2$ and hence $WE^{\{2,1\}} = WE^{\{1,1,1\}}$. However, $\{1, 1, 1\}$ is also an extraction from $\{1, 2\}$ with $n_1 = 1$ and hence the extraction strictly increases WE; that is, $WE^{\{1,2\}} < WE^{\{1,1,1\}}$.

As for mergers, we obtain the following proposition, with proof in the Appendix.

Proposition 4. *Consider a team contest à la FLP described in Section 2 with CSF (2) and $c_A = c_B$. A merger from $\{1, n_2, n_3, ..\}$ to $\{n_2 + 1, n_3, ..\}$ (i.e., $n_1 = 1$) does not affect WE. Any other merger strictly decreases WE. Therefore, among all possible temporal structures there are two WE-minimizing temporal structures: the fully simultaneous contest and $\{1, 2n\}$.*

The ranking in (15) is consistent with the results of Proposition 4. Indeed, $\{3\}$ is a merger from $\{1, 2\}$ with $n_1 = 1$ and hence $WE^{\{3\}} = WE^{\{1,2\}}$, and $\{2, 1\}$ is a merger from $\{1, 1, 1\}$ with $n_1 = 1$ and hence $WE^{\{2,1\}} = WE^{\{1,1,1\}}$. However, $\{3\}$ is also a merger from $\{2, 1\}$ with $n_1 = 2$ and hence the merger strictly decreases WE; that is, $WE^{\{3\}} < WE^{\{2,1\}}$.

4.4 Asymmetric teams: neither the fully sequential nor the fully simultaneous temporal structures maximize or minimize WE

As explained in Section 4.1, the team-asymmetry and stochastic-effort effect are countervailing. Recall that in Section 3 only the team-asymmetry effect was present and in Section 4.3 only the stochastic-effort effect was present. In both cases, we hence obtained clean results on the WE-maximizing and WE-minimizing temporal structures. When both effects are present, as it happens with CSF (2) and $c_A \neq c_B$, extractions and mergers need not have an equally unambiguous effect on WE. Indeed, we can show examples where the WE-maximizing and WE-minimizing temporal structures are different for different values of

(c_A, c_B).[7] However, we can still derive results of relevance for the two temporal structures of greatest interest: the fully simultaneous and the fully sequential ones. Importantly, these results are valid for any (c_A, c_B) and n. In particular, we now show that the fully simultaneous and the fully sequential temporal structures are "interior"; i.e., neither WE-maximizing nor WE-minimizing. The following proposition, with proof in the Appendix, assumes $c_A < c_B$ because we are interested in the case of $c_A \neq c_B$ and $c_A \leq c_B$ is without loss of generality.

Proposition 5. *Consider a team contest à la FLP described in Section 2 with CSF (2) and $c_A < c_B$. Among all possible temporal structures, the fully sequential and the fully simultaneous contests neither maximize nor minimize WE.*

The ranking in (16) is consistent with the results of Proposition 5. The method of proof is to show that $WE^{\{1,...,1,2\}} < WE^{\{1,....,1\}} < WE^{\{2,1,...,1\}}$ and $WE^{\{1,2n\}} < WE^{\{2n+1\}} < WE^{\{2n,1\}}$. Note that these comparisons are simplified by the use of extractions and mergers. Indeed, $\{1, ..., 1\}$ is an extraction starting from either $\{1, ..., 1, 2\}$ or $\{2, 1, ..., 1\}$, and $\{2n + 1\}$ is a merger starting from either $\{1, 2n\}$ or $\{2n, 1\}$. In particular, Proposition 3 showed that, when $c_A = c_B$, there are two WE-maximizing temporal structures: $\{1, 1, ..., 1\}$ and $\{2, 1, ..., 1\}$. As soon as $c_A < c_B$, the team-asymmetry effect kicks in, favoring the more simultaneous temporal structure. Hence, the equivalence $WE^{\{1,....,1\}} = WE^{\{2,1,...,1\}}$ is broken in favor of temporal structure $\{2, 1, ..., 1\}$, which is indeed more simultaneous than $\{1, 1, ..., 1\}$. All other arguments follow in a similar fashion.

5 Semi-mixed-strategy equilibrium

In this section, we consider CSF (1) with r such that $\bar{r}(c_A, c_B) < r \leq 2$. (This requires $c_A \neq c_B$, because when $c_A = c_B$, $\bar{r}(c_A, c_B) = 2$.) Assuming $c_A < c_B$, the equilibrium in any non-trivial battle is in semi-mixed strategies: the A-player uses a pure strategy and the B-player mixes between 0 effort and a strictly positive effort (see Wang, 2010; Feng and Lu, 2017; Ewerhart, 2017a). The (ex-ante) equilibrium probabilities of victory in a battle are

$$p_A = 1 - \frac{(r-1)^{\frac{r-1}{r}} c_A}{r c_B}, \text{ and } p_B = \frac{(r-1)^{\frac{r-1}{r}} c_A}{r c_B}. \tag{20}$$

[7] Consider a best-of-5 contest with CSF (2). When $(c_A, c_B) = (4, 5)$, the WE-minimizing temporal structure is $\{1, 4\}$ and the WE-maximizing temporal structure is $\{2, 1, 1, 1\}$. When $(c_A, c_B) = (2, 5)$, the WE-minimizing temporal structure is $\{1, 1, 3\}$ and the WE-maximizing temporal structure is $\{3, 1, 1\}$.

We next define

$$\hat{\mu}_A^m \equiv \frac{(r-1)^{\frac{r-1}{r}}}{rc_B\gamma_m},\ \hat{\mu}_{B|A}^m \equiv \frac{1}{rc_B\gamma_m}\frac{p_B}{p_A},\ \hat{\mu}_{B|B}^m \equiv \frac{r-1}{rc_B\gamma_m}. \tag{21}$$

Then, the analogue of Lemma 1 and Lemma 5 is:

Lemma 6. *Let (b_A, b_B) be the state at the beginning of round j with $b_A, b_B \leq n$ so that battle $m \in \{b_A + b_B + 1, ..., b_A + b_B + n_j\}$ is non-trivial. The equilibrium individual effort of an A-player in battle m is*

$$x_A^m(b_A, b_B) = \hat{\mu}_A^m \cdot \binom{2n-(b_A+b_B)}{n-b_A} p_A^{n-b_A} p_B^{n-b_B}, \tag{22}$$

and the expected *equilibrium individual effort of a B-player, conditional on the outcome of battle m, is*

$$E\left[x_B^m(b_A, b_B)\,|i \text{ wins battle } m\right] = \hat{\mu}_{B|i}^m \cdot \binom{2n-(b_A+b_B)}{n-b_A} p_A^{n-b_A} p_B^{n-b_B} \text{ with } i \in \{A, B\}. \tag{23}$$

The values of $\hat{\mu}_A^m$ and $\hat{\mu}_{B|i}^m$ are in (21) and of p_A and p_B are in (20).

Proof. The proof is similar to that of Lemma 5 and thus omitted. □

As the equilibrium is in semi-mixed strategies, the team-asymmetry effect applies to both teams, whereas the stochastic-effort effect only to the weaker team. Thus, as intuition suggests, we obtain a result, with proof in the Online Appendix, that straddles those for pure strategies in Section 3 and those for mixed strategies in Section 4.

Proposition 6. *Consider a team contest à la FLP described in Section 2 with CSF (1), $r > \bar{r}(c_A, c_B)$, and $c_A < c_B$. Among all possible temporal structures, the fully sequential contest neither maximizes nor minimizes WE and the fully simultaneous one does not minimize WE.*

The Proof of Proposition 6 uses (21) and the same technique in the Proof of Proposition 5 to show that the first three parts of the Proof of Proposition 5 carry over, while the fourth does not necessarily go through; thus, we obtain $WE^{\{1,\dots,1,2\}} < WE^{\{1,\dots,1\}} < WE^{\{2,1,\dots,1\}}$ and $WE^{\{1,2n\}} < WE^{\{2n+1\}}$, but the inequality $WE^{\{2n+1\}} < WE^{\{2n,1\}}$ may or may not hold. The following example shows that fully simultaneous temporal structure may or may not maximize WE, depending on parameter values.

Consider a best-of-three setup with $c_A = 1$ and $c_B = 2^{2/3}$, so that (4) gives $\bar{r}\left(1, 2^{2/3}\right) = 3/2$. For $r \in (3/2, 2]$, the Proof of Proposition 6 shows that $WE^{\{1,2\}} < WE^{\{1,1,1\}} < WE^{\{2,1\}}$ and $WE^{\{1,2\}} < WE^{\{3\}}$. The precise ranking of $WE^{\{3\}}$ depends on the value of r. When r

is not too large ($r \in (3/2, 1.67315)$) so that we "approach" the pure-strategy setup where the stochastic-effort effect is absent, we obtain $WE^{\{2,1\}} < WE^{\{3\}}$ as indeed we obtained in the pure-strategy setup. When r is large enough ($r \in (1.86198, 2]$) so that we are closer to the mixed-strategy setup where the stochastic-effort effect is important, we obtain $WE^{\{3\}} < WE^{\{1,1,1\}}$, as indeed we obtained in the mixed-strategy setup (see (16)). For the remaining intermediate values of r, we have $WE^{\{1,2\}} < WE^{\{1,1,1\}} < WE^{\{3\}} < WE^{\{2,1\}}$.

6 Extensions

We next endogenize the order of battles, relax the complete information assumption, and allow for player-specific heterogeneity. For each extension we use extractions and mergers.

Endogenous order of battles. We now assume that the vector of $\gamma_m's$ can be re-arranged across battles. We focus on the deterministic setup of Section 3 with CSF (1) and $r \leq \bar{r}(c_A, c_B)$, as abstracting from the stochastic-effort effect simplifies this analysis. First, the contest that maximizes WE when both the temporal structure and the order of battles are endogenous is a fully simultaneous contest: the order of battles does not matter (see Proposition 2). Second, as WE is minimized by a fully sequential contest (Proposition 1) for any fixed order of battles, we investigate how to order the battles of a fully sequential contest to further minimize WE. The following proposition, with proof in the Online Appendix, shows that the contest that minimizes WE when both the temporal structure and the order of battles are endogenous is a fully sequential contest with battles ordered from highest to lowest marginal cost of effort (i.e., from the hardest to the easiest battle).

Proposition 7. *Consider a team contest à la FLP described in Section 2 with CSF (1) and $r \leq \bar{r}(c_A, c_B)$.* In the fully sequential contest *with endogenous order of battles, WE is minimized by sorting battles in decreasing order of γ_m. WE is independent of the order of γ_m only if $c_A = c_B$ or all γ_m are the same.*

Intuitively, to minimize WE, the team-asymmetry effect should apply to battles with high efforts (i.e., low γ_m), which are then optimally placed in the final rounds. Our results in Proposition 7 speak, for instance, to the design of team sports competitions with pairwise matches. Teams can often choose the order of play of their players independently (e.g., the Ryder's cup in golf, the Laver cup in tennis, and the Swaythling cup in table tennis). Thus, our results call for the possibility of intervention by a contest organizer to impose a certain order of matches.

Private information. We modify our model by multiplying each player's costs by a random variable c (a player's marginal cost becomes $c \cdot c_i \cdot \gamma_m$); we assume c is private information of any player and independently and separately drawn for each player from a common distribution. Assume that primitives (CSF, distribution of c, etc.) are such that equilibrium efforts in each non-trivial battle are strictly decreasing functions from realizations of c to effort levels, and hence the equilibrium efforts are stochastic. As in Section 4.3, we assume ex-ante symmetric teams ($c_A = c_B$) so as to single out the stochastic-effort effect. We obtain the following proposition, with proof in the Online Appendix.

Proposition 8. *Consider a team contest à la FLP described in Section 2 with private information as described in the above paragraph and $c_A = c_B$, so that $p_A = p_B = 1/2$. Assume primitives are such that, whenever battle $m \in \{1, .., 2n+1\}$ is actually fought, equilibrium efforts are stochastic and*

$$E[x_i^m(b_A, b_B)|j \text{ wins battle } m] = \frac{\rho_{i|j}}{\gamma_m}\binom{2n-(b_A+b_B)}{n-b_A}p_A^{n-b_A}p_B^{n-b_B} \text{ with } i, j \in \{A, B\}, \quad (24)$$

where $\rho_{i|j} \in \mathbb{R}_+ \ \forall i, j \in \{A, B\}$. The results of Propositions 3 and 4 hold if

$$\rho_{i|i} - \rho_{i|j} = \rho_{j|j} - \rho_{j|i} \quad \forall i, j \in \{A, B\}. \quad (25)$$

Proposition 8 applies to two well-known setups. The first is when battles are as in Amann and Leininger (1996)'s setup with CSF (2) and c distributed with a strictly positive density f on $[1, \infty)$. In a one-shot contest à la Amann and Leininger (1996), when the marginal effort costs are $c \cdot c_i \cdot \gamma_m$ with $i \in \{A, B\}$ under $c_A = c_B$ and the prize $\check{\nu}$ is identical between players, equilibrium efforts are identical across players and characterized by

$$\check{x}(c) = \int_c^\infty \left(\frac{\check{\nu}}{c_A\gamma_m}\frac{f(y)}{y}\right)dy, \quad (26)$$

assuming that f is such that the integral in (26) is well-defined and $\lim_{c\to+\infty}\check{x}(c) = 0$. Note that equilibrium efforts are linear in the prize $\check{\nu}$ and in $1/\gamma_m$ and strictly decreasing in c.

In each battle of our setup, equilibrium efforts (and thus their expectations) are linear in $1/\gamma_m$ and the common, state-dependent battle prize $\binom{2n-(b_A+b_B)}{n-b_A}p_A^{n-b_A}p_B^{n-b_B}$. And these are the only two terms that depend on the battle. No other term in (26) changes in m as c_A and f are assumed to be fixed. Thus, conditional expected equilibrium efforts satisfy the functional form in (24): a term $\rho_{i|j}$ that does not depend on the state or m, multiplied by the battle-dependent term $1/\gamma_m$ and the state-dependent battle prize. To calculate the term $\rho_{A|A}$ ($\rho_{A|B}$) that enters the expected equilibrium effort of the A player that wins (loses) a generic battle m, we take $\int_c^\infty \left(\frac{1}{c_A}\frac{f(y)}{y}\right)dy$ from (26) and further integrate it with the probability density function of the minimum (maximum) of two i.i.d. draws for c. Similar considerations apply to $\rho_{B|B}$ and $\rho_{B|A}$. By symmetry across teams, and as the probability density function of the maximum first-order stochastically dominates that of the minimum, we obtain $\rho_{A|A} = \rho_{B|B} > \rho_{A|B} = \rho_{B|A}$, thus verifying (25).

The second setup where Proposition 8 applies is when battles are as in Malueg and Yates (2004)'s setup with CSF (1) and r sufficiently low. Using their Proposition 1 with $\sigma = 1/2$ so that draws are i.i.d. and converting prize values into reciprocals of costs to match their results and our setup, one can verify that the conditions in Proposition 8 hold.

Player-specific heterogeneity. So far, our setup allowed only for team- and battle-specific, but not player-specific, heterogeneity. We can add player-specific heterogeneity to our setup by letting c_i^m vary freely across battles. If such a player-specific heterogeneity is sufficiently small (i.e., the ratio c_A^m/c_B^m changes negligibly with m), then our results go through by continuity except for knife-edge results (e.g., temporal structures yielding identical WE). But, if player-specific heterogeneity is large, then our results may fail to carry over. We next derive more precise sufficient conditions such that our results go through. We focus on adding player-specific heterogeneity to the deterministic setup of Section 3 with CSF (1) and $r \leq 1$ to ensure a pure strategy equilibrium in every battle regardless of the specific values of marginal costs, as abstracting from the stochastic-effort effect simplifies the analysis.

The first main result of Section 3 is that, without player-specific heterogeneity so that p_A is constant across battles, extractions decrease WE. Here, we find that extractions decrease WE if player-specific heterogeneity is such that $c_A^m < c_B^m$ and the ratio c_A^m/c_B^m increases in m. These assumptions imply that p_A^m—defined as the value of the equilibrium probability of victory of the player of team A in battle m (and $p_B^m = 1-p_A^m$)—is strictly larger than 1/2 and decreasing in m. Note that, from FLP, it is still true that p_A^m is independent of the temporal structure and of the state (b_A, b_B) as long as the contest is not yet decided; it only depends on the marginal costs in battle m. Hence, while Section 3 required $p_A^1 = ... = p_A^{2n+1} \geq 1/2$, which is a non-full-dimensional subset in $[0,1]^{2n+1}$, we now show that our results go through for a full-dimensional subset, where $p_A^1 \geq ... \geq p_A^{2n+1} > 1/2$. We prove the following proposition in the Online Appendix.

Proposition 9. *Consider a team contest à la FLP described in Section 2 with CSF (1) and $r \leq 1$. Assume $c_A^m < c_B^m$ and the ratio c_A^m/c_B^m increases in m. If p_A^m is constant in m, an extraction from $\{1, 1, ..., 1, 2\}$ to $\{1, 1, ..., 1\}$ (i.e., $l = 2n$) does not affect WE; any other extraction strictly decreases WE. If p_A^m is not constant in m, any extraction strictly decreases WE so that, among all possible temporal structures, there is one WE-minimizing temporal structure: the fully sequential contest.*

Note that the equivalence of WE between $\{1, 1, ..., 1, 2\}$ and $\{1, 1, ..., 1\}$, established in

Proposition 1, ceases to hold as soon as p_A^m is not constant in m.

The second main result of Section 3 is that, without player-specific heterogeneity so that p_A is constant across battles, mergers increase WE. The following proposition, with proof in the Online Appendix, shows that a merger increases WE when $c_A^m < c_B^m$ and the ratio c_A^m/c_B^m increases in m.

Proposition 10. *Consider a team contest à la FLP described in Section 2 with CSF (1) and $r \leq 1$. Assume $c_A^m < c_B^m$ and the ratio c_A^m/c_B^m increases in m. Any merger increases WE. Thus, among all possible temporal structures, there is one WE-maximizing temporal structure: the fully simultaneous contest.*

For more general player-specific heterogeneity, the ranking of sequential and simultaneous structures may reverse: WE may end up being larger in a sequential structure than in a simultaneous one. For instance, when $n = 1$, $(c_A^1, c_A^2, c_A^3) = (1, 2, 3)$, and $(c_B^1, c_B^2, c_B^3) = (3, 2, 1)$, we obtain $WE^{\{1,1,1\}} = 7/32$ and $WE^{\{3\}} = 13/64$. The intuition we developed in Section 3.1 can be used to understand why $WE^{\{1,1,1\}} > WE^{\{3\}}$, as we now explain.

If player-specific heterogeneity is large enough, it becomes possible for team A to be advantaged in the first battle, and for team B to be advantaged afterward, even if the first battle is won by team A. It is clear that, if the identity of the advantaged team switches across battles, then the probability of winning for the team exerting the largest efforts and the pivotal probabilities of future battles may move in the same, rather than the opposite, direction (recall the intuition in Section 3.1). Therefore, WE may end up being larger in a sequential structure than in a simultaneous one. Importantly, if the identity of the advantaged player switches across battles and player-specific heterogeneity is large enough, also the conventional wisdom for TE in *individualistic* contests (i.e., TE is larger in a simultaneous than in a sequential contest) may reverse (see, e.g., Barbieri and Serena, 2020).

7 Conclusions

We study pairwise team contests, where each battle is fought by a different player who bears only the effort cost of her battle. Besides applying widely, this setup has several tractable and elegant features, and the consequent "neutrality results, which break the dynamic linkage among battles, sharply contrast with the conventional wisdom in the literature" (see FLP, p. 2134): sequential and (partially) simultaneous contests yield the same TE for a wide range of pairwise team contests. In contrast, we show that WE—which is of interest in applications such as R&D races, elections, and sports—may be affected by the temporal structure of the contest. A more sequential overall contest (e.g., a contest obtained through an extraction) generates shifts in future-battle efforts that keep the average effort constant. TE, which depends only on average efforts, is unaffected by such shifts. In contrast, WE does not depend only on the average effort and hence *is* in general affected by the above-described shifts in efforts. Another important difference between TE and WE is that, in a setup where equilibrium efforts are stochastic, a battle's efforts conditional on victory or loss typically differ, and while TE weighs them equally, WE does not. Abstracting from player-specific heterogeneity, we obtain four main results. With sufficient noise, we show that: (1) If teams are symmetric, all temporal structures yield the same WE; and (2) If teams are asymmetric, WE is maximized by a fully simultaneous contest and minimized by a fully sequential one. With no noise, we show that: (3) If teams are symmetric, WE is maximized by a fully sequential contest and minimized by a fully simultaneous one; and (4) If teams are asymmetric, neither the fully sequential nor the fully simultaneous temporal structures maximize or minimize WE.

Our results on the temporal structures that maximize or minimize WE are robust along two other dimensions. First, our results carry over to a setup where the player that wins each battle not only increases the probability that her team wins the overall contest but also obtains a per-battle prize $k > 0$. The effect of k is to add to efforts a term, linear in k, that does not change the win probability in a non-trivial battle m. Furthermore, the linear terms are independent of the pivotal probability of battle m, so the team-asymmetry effect does not apply to the linear terms. Moreover, when $k > 0$ all battles are actually fought, and therefore the stochastic-effort effect on linear term in k applies to all temporal structures equally. All in all, WE is increased by an additive term independent of the temporal structure and hence our results on temporal structure comparisons are not affected by k. Second, teams could have head-starts: team $i \in \{A, B\}$ wins when winning M_i battles, with $M_A \neq M_B$. Think of an R&D competition where one of the two competing joint ventures already developed some components of the final project prior to the contest, for instance, because of its past R&D efforts. The team-asymmetry effect carries over if the team—say team A—advantaged in every battle (its win probability is strictly greater than 1/2) is also the team having the head-start (that is, if $M_A \leq M_B$). In this case, our results on temporal structures are unchanged. If the other team has the head-start instead, our results may be reversed because the identity of the advantaged team switches over battles (as discussed at the end of Section 6).

In deriving our results, as the nature of WE and TE sharply differs, we introduce a novel technique, the analysis of extractions and mergers, which overcomes the loss of tractability that arises when moving from TE to WE. Our results underscore the importance of an application-driven specification of the objective and also reopen design questions, such as the effect of interim evaluations, that the focus on TE and FLP's independence results had closed. As we showed, these questions can be explored using FLP's elegant setup, extractions, and mergers.

Appendix

Throughout this appendix, we use combinatorial arguments a number of times. We borrow the convention in Knuth (1997; pages 53-56): with a non-negative integer n, the binomial coefficient $\binom{n}{k}$ equals 0 if $k < 0$ or $k > n$, and it equals 1 if $k = 0$. Finally, Pascal's rule is stated with k being any (even negative) integers (Knuth, 1997; p. 56, Equation (9)).

Proof of Lemma 2. By Corollary 1, when we compare WE in T and T^{ext}, we can focus only on battles from $l+1$ to $l+n_l-1$ in both temporal structures. These battles occur in round l in T and are not extracted, thus they occur in round $l+1$ in T^{ext}. Fix one such battle m. In (3), the effort of the A-player (B-player) in battle m enters in WE^T multiplied by $\sum_{\omega\in\Omega_A^T}\Pr\{\omega\}$ ($\sum_{\omega\in\Omega_B^T}\Pr\{\omega\}$). We decompose $\sum_{\omega\in\Omega_A^T}\Pr\{\omega\}$ into the probability of reaching all states for which battle m is played with positive efforts, and then the effort of the A-player is multiplied by the probability that the path continues with the victory of team A. The first l battles result in b_A victories for team A and $l-b_A$ victories for team B, and each state $(b_A, l-b_A)$ occurs with probability $\binom{l}{b_A}p_A^{b_A}p_B^{l-b_A}$.

Using the simplifications above, when comparing WE in T and T^{ext}, the contribution to WE^T of the battles from $l+1$ to $l+n_l-1$ is denoted by sWE_E^T, where

$$sWE_E^T = \sum_{b_A=\max\{l-1-n,0\}}^{\min\{n,l-1\}} \binom{l-1}{b_A} p_A^{b_A} p_B^{l-1-b_A} \sum_{m=l+1}^{l+n_l-1} \phi^m\left(b_A, l-1-b_A\right), \tag{A.1}$$

with

$$\phi^m\left(b_A, b_B\right) \equiv P_A\left(b_A, b_B\right) x_A^m\left(b_A, b_B\right) + P_B\left(b_A, b_B\right) x_B^m\left(b_A, b_B\right), \tag{A.2}$$

recalling that the value of $x_i^m(b_A, b_B)$ is in (10), and that of $P_A\left(b_A, b_B\right)$ is in (7), with $P_B\left(b_A, b_B\right) = 1 - P_A\left(b_A, b_B\right)$, all evaluated using (11). Similarly, the contribution to $WE^{T^{ext}}$ of the battles from $l+1$ to $l+n_l-1$ is denoted by $sWE_E^{T^{ext}}$, where

$$sWE_E^{T^{ext}} = \sum_{b_A=\max\{l-n,0\}}^{\min\{n,l\}} \binom{l}{b_A} p_A^{b_A} p_B^{l-b_A} \sum_{m=l+1}^{l+n_l-1} \phi^m\left(b_A, l-b_A\right). \tag{A.3}$$

Note that the min and max operators in the first summation of (A.1) and (A.3) ensure that we are taking into account only non-trivial battles (i.e., no team has won more than n battles). And the preliminary equilibrium observation 3 allows us to "pretend" that all future battles are non-trivial, see the discussion after equation (7). Therefore, WE in T is larger (equal) [smaller] than in T^{ext} if and only if $sWE_E^T > (=)[<] sWE_E^{T^{ext}}$.

Using (10) and (A.2) into (A.1), we obtain that sWE_E^T equals

$$
\begin{aligned}
& \sum_{b_A=\max\{l-1-n,0\}}^{\min\{n,l-1\}} \binom{l-1}{b_A} p_A^{b_A} p_B^{l-1-b_A} \sum_{m=l+1}^{l+n_l-1} \phi^m \left(b_A, l-1-b_A\right) \\
= & \sum_{b_A=\max\{l-1-n,0\}}^{\min\{n,l-1\}} \binom{l-1}{b_A} p_A^{b_A} p_B^{l-1-b_A} \\
& \cdot \sum_{m=l+1}^{l+n_l-1} \left[P_A(b_A, l-1-b_A) x_A^m(b_A, l-1-b_A) + P_B(b_A, l-1-b_A) x_B^m(b_A, l-1-b_A)\right] \\
= & \sum_{b_A=\max\{l-1-n,0\}}^{\min\{n,l-1\}} \binom{l-1}{b_A} p_A^{b_A} p_B^{l-1-b_A} \binom{2n-(l-1)}{n-b_A} p_A^{n-b_A} p_B^{n-(l-1-b_A)} \\
& \cdot \sum_{m=l+1}^{l+n_l-1} \left[P_A\left(b_A, l-1-b_A\right) \mu_A^m + P_B\left(b_A, l-1-b_A\right) \mu_B^m\right] \\
= & \; p_A^n p_B^n \sum_{b_A=\max\{l-1-n,0\}}^{\min\{n,l-1\}} \binom{l-1}{b_A} \binom{2n-(l-1)}{n-b_A} \\
& \cdot \sum_{m=l+1}^{l+n_l-1} \left[P_A\left(b_A, l-1-b_A\right) \mu_A^m + P_B\left(b_A, l-1-b_A\right) \mu_B^m\right] \\
= & \; p_A^n p_B^n \sum_{b_A=\max\{l-1-n,0\}}^{\min\{n,l-1\}} \binom{l-1}{b_A} \binom{2n-(l-1)}{n-b_A} \sum_{m=l+1}^{l+n_l-1} \left[P_A\left(b_A, l-1-b_A\right)\left(\mu_A^m - \mu_B^m\right) + \mu_B^m\right] \\
= & \; p_A^n p_B^n \psi\left(l-1\right) \sum_{m=l+1}^{l+n_l-1} \left[\mu_A^m - \mu_B^m\right] + p_A^n p_B^n \xi\left(l-1\right) \sum_{m=l+1}^{l+n_l-1} \mu_B^m,
\end{aligned}
$$

where ψ is defined in (12) and $\xi\left(z\right)$ as follows:

$$
\xi\left(z\right) \equiv \sum_{b_A=\max\{z-n,0\}}^{\min\{n,z\}} \binom{z}{b_A} \binom{2n-z}{n-b_A}. \tag{A.4}
$$

We now apply Vandermonde's convolution to show $\xi\left(z\right) = \binom{2n}{n}$:

$$
\begin{aligned}
\text{if } z & \leq n, \text{ then } \xi\left(z\right) = \sum_{b_A=0}^{z} \binom{z}{b_A} \binom{2n-z}{n-b_A} = \binom{2n}{n}, \text{ and} \\
\text{if } z & > n, \text{ then } \xi\left(z\right) = \sum_{b_A=z-n}^{n} \binom{z}{b_A} \binom{2n-z}{n-b_A} = \sum_{j=0}^{2n-z} \binom{z}{n-j} \binom{2n-z}{j} = \binom{2n}{n},
\end{aligned}
$$

where $j = n - b_A$. Thus, we obtain

$$\frac{sWE_E^T}{p_A^n p_B^n} = \binom{2n}{n} \sum_{m=l+1}^{l+n_l-1} \mu_B^m + \psi(l-1) \sum_{m=l+1}^{l+n_l-1} [\mu_A^m - \mu_B^m]. \tag{A.5}$$

Applying similar steps to (A.3), we have

$$\frac{sWE_E^{T^{ext}}}{p_A^n p_B^n} = \binom{2n}{n} \sum_{m=l+1}^{l+n_l-1} \mu_B^m + \psi(l) \sum_{m=l+1}^{l+n_l-1} [\mu_A^m - \mu_B^m]. \tag{A.6}$$

Comparing (A.5) and (A.6) establishes the lemma. □

Proof of Lemma 3. Throughout the proof we let $\Delta(z) \equiv \psi(z-1) - \psi(z)$. We show that $p_A \gtreqless p_B \Rightarrow \Delta(z) \gtreqless 0$ for $z \leq n$ and $z \neq 2$ in **Step 1**, for $z = 2$ in **Step 2**, and for $z > n$ and $z \neq 2n$ in **Step 3**. Finally, we prove that $\Delta(z) = 0$ for $z = 2n$ in **Step 4**.

Step 1. Consider $z \leq n$. Pascal's rule applied to (12) gives

$$\begin{aligned}
\psi(z-1) &= \sum_{b_A=0}^{z-1} P_A(b_A, z-1-b_A) \binom{z-1}{b_A} \binom{2n-z}{n-b_A} \\
&\quad + \sum_{b_A=0}^{z-1} P_A(b_A, z-1-b_A) \binom{z-1}{b_A} \binom{2n-z}{n-b_A-1} \\
&= \sum_{b_A=0}^{z-1} P_A(b_A, z-1-b_A) \binom{z-1}{b_A} \binom{2n-z}{n-b_A} \\
&\quad + \sum_{\tilde{b}_A=1}^{z} P_A\left(\tilde{b}_A - 1, z-1-\left(\tilde{b}_A-1\right)\right) \binom{z-1}{\tilde{b}_A-1} \binom{2n-z}{n-\left(\tilde{b}_A-1\right)-1} \\
&= \sum_{b_A=0}^{z-1} P_A(b_A, z-1-b_A) \binom{z-1}{b_A} \binom{2n-z}{n-b_A} \\
&\quad + \sum_{\tilde{b}_A=1}^{z} P_A\left(\tilde{b}_A - 1, z-\tilde{b}_A\right) \binom{z-1}{\tilde{b}_A-1} \binom{2n-z}{n-\tilde{b}_A},
\end{aligned}$$

where we used the change of variable $\tilde{b}_A - 1 = b_A$. Similarly,

$$\begin{aligned}
\psi(z) &= P_A(z,0)\binom{2n-z}{n-z} + \sum_{b_A=1}^{z-1} P_A(b_A, z-b_A)\binom{z}{b_A}\binom{2n-z}{n-b_A} + P_A(0,z)\binom{2n-z}{n} \\
&= P_A(z,0)\binom{2n-z}{n-z}
\end{aligned}$$

$$
\begin{aligned}
& + \sum_{b_A=1}^{z-1} P_A\left(b_A, z-b_A\right)\left(\binom{z-1}{b_A}+\binom{z-1}{b_A-1}\right)\binom{2n-z}{n-b_A} + P_A\left(0,z\right)\binom{2n-z}{n} \\
= \; & P_A\left(z,0\right)\binom{2n-z}{n-z} + \sum_{b_A=1}^{z-1} P_A\left(b_A, z-b_A\right)\binom{z-1}{b_A-1}\binom{2n-z}{n-b_A} \\
& + \sum_{b_A=1}^{z-1} P_A\left(b_A, z-b_A\right)\binom{z-1}{b_A}\binom{2n-z}{n-b_A} + P_A\left(0,z\right)\binom{2n-z}{n} \\
= \; & \sum_{b_A=1}^{z} P_A\left(b_A, z-b_A\right)\binom{z-1}{b_A-1}\binom{2n-z}{n-b_A} + \sum_{b_A=0}^{z-1} P_A\left(b_A, z-b_A\right)\binom{z-1}{b_A}\binom{2n-z}{n-b_A}.
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
\Delta\left(z\right) = \; & -\sum_{b_A=1}^{z}\left(P_A\left(b_A, z-b_A\right) - P_A\left(b_A-1, z-b_A\right)\right)\binom{z-1}{b_A-1}\binom{2n-z}{n-b_A} \\
& + \sum_{b_A=0}^{z-1}\left(P_A\left(b_A, z-1-b_A\right) - P_A\left(b_A, z-b_A\right)\right)\binom{z-1}{b_A}\binom{2n-z}{n-b_A}.
\end{aligned} \tag{A.7}
$$

Note now that $P_A\left(b_A-1, z-b_A\right) = p_A P_A\left(b_A, z-b_A\right) + p_B P_A\left(b_A-1, z-b_A+1\right)$, so that

$$
\begin{aligned}
P_A\left(b_A, z-b_A\right) - P_A\left(b_A-1, z-b_A\right) & = P_A\left(b_A, z-b_A\right) - p_A P_A\left(b_A, z-b_A\right) \\
& \quad - p_B P_A\left(b_A-1, z-b_A+1\right) \\
& = p_B\left[P_A\left(b_A, z-b_A\right) - P_A\left(b_A-1, z-b_A+1\right)\right] \\
& = p_B\binom{2n+1-z}{n+1-b_A} p_A^{n+1-b_A} p_B^{n-z+b_A} \\
& = \binom{2n+1-z}{n+1-b_A} p_A^{n+1-b_A} p_B^{n+1-z+b_A}.
\end{aligned} \tag{A.8}
$$

Similarly,

$$
\begin{aligned}
P_A\left(b_A, z-1-b_A\right) - P_A\left(b_A, z-b_A\right) & = p_A P_A\left(b_A+1, z-1-b_A\right) + p_B P_A\left(b_A, z-b_A\right) \\
& \quad - P_A\left(b_A, z-b_A\right) \\
& = p_A\left(P_A\left(b_A+1, z-1-b_A\right) - P_A\left(b_A, z-b_A\right)\right) \\
& = p_A\binom{2n+1-z}{n+1-\left(b_A+1\right)} p_A^{n+1-\left(b_A+1\right)} p_B^{n-z+b_A+1} \\
& = \binom{2n+1-z}{n-b_A} p_A^{n+1-b_A} p_B^{n+1-z+b_A}.
\end{aligned} \tag{A.9}
$$

Thus, using (A.8) and (A.9) into (A.7), we obtain

$$
\begin{aligned}
&\Delta (z) = -\psi (z) + \psi (z-1) \\
&\quad = -\sum_{b_A=1}^{z} \binom{2n+1-z}{n+1-b_A} p_A^{n+1-b_A} p_B^{n+1-z+b_A} \binom{z-1}{b_A-1} \binom{2n-z}{n-b_A} \\
&\quad + \sum_{b_A=0}^{z-1} \binom{2n+1-z}{n-b_A} p_A^{n+1-b_A} p_B^{n+1-z+b_A} \binom{z-1}{b_A} \binom{2n-z}{n-b_A} \\
&\quad = \binom{2n+1-z}{n} \binom{2n-z}{n} p_A^{n+1} p_B^{n+1-z} - \binom{2n+1-z}{n+1-z} \binom{2n-z}{n-z} p_A^{n+1-z} p_B^{n+1} \\
&\quad + \sum_{b_A=1}^{z-1} \binom{2n-z}{n-b_A} \left[\binom{2n+1-z}{n-b_A} \binom{z-1}{b_A} - \binom{2n+1-z}{n+1-b_A} \binom{z-1}{b_A-1} \right] p_A^{n+1-b_A} p_B^{n+1-z+b_A} \\
&\quad = \binom{2n+1-z}{n} \binom{2n-z}{n} (p_A p_B)^{n+1-z} (p_A^z - p_B^z) \\
&\quad + p_A^{n+1} p_B^{n+1-z} \sum_{b_A=1}^{z-1} \binom{2n-z}{n-b_A} \left(\frac{p_B}{p_A} \right)^{b_A} \left[\binom{2n+1-z}{n-b_A} \binom{z-1}{b_A} - \binom{2n+1-z}{n+1-b_A} \binom{z-1}{b_A-1} \right].
\end{aligned}
\tag{A.10}
$$

Recall that we want to show that $p_A \gtreqless p_B \Rightarrow \Delta (z) \gtreqless 0$ for $z \leq n$ and $z \neq 2$. Given (A.10) and $p_B = 1 - p_A$, it suffices to prove that

$$
\begin{aligned}
p_A \; &\gtreqless \; p_B \Longrightarrow \sum_{b_A=1}^{z-1} d(b_A) \left(\frac{1-p_A}{p_A} \right)^{b_A} \gtreqless 0 \\
\text{where } d(b_A) \; &\equiv \; \binom{2n-z}{n-b_A} \left[\binom{2n+1-z}{n-b_A} \binom{z-1}{b_A} - \binom{2n+1-z}{n+1-b_A} \binom{z-1}{b_A-1} \right].
\end{aligned}
\tag{A.11}
$$

The sign of $d(b_A)$ in (A.11) equals

$$
\begin{aligned}
& sgn \left[\binom{2n+1-z}{n-b_A} \binom{z-1}{b_A} \left(1 - \frac{(n+1-z+b_A) b_A}{(n+1-b_A)(z-b_A)} \right) \right] \\
&= sgn \left[(n+1-b_A)(z-b_A) - (n+1-z+b_A) b_A \right] \\
&= sgn \left[(z-2b_A)(n+1) \right] \\
&= sgn (z-2b_A).
\end{aligned}
\tag{A.12}
$$

Therefore, $d(b_A) > 0$ when $b_A < z/2$, $d(b_A) < 0$ when $b_A > z/2$, and $d(b_A) = 0$ when z is even and $b_A = z/2$. In words, if we view the expression in (A.11) as a polynomial in $\frac{1-p_A}{p_A}$, the first $\lfloor \frac{z-1}{2} \rfloor$ coefficients are positive and the last $\lfloor \frac{z-1}{2} \rfloor$ coefficients are negative since in the summation b_A runs from 1 to $z-1$. Next, we show that $d(1) = -d(z-1)$, $d(2) = -d(z-2)$, and so on until all coefficients are covered. With compact notation, we show that $d(1+k) = -d(z-1-k)$ for any $k \in \{0,..,z-1\}$. Indeed, by the symmetry rule of binomial coefficients, we have

$$\begin{aligned}
-d(z&-1-k) \\
&= \binom{2n-z}{n-(z-1-k)}\left[-\binom{2n+1-z}{n-z+1+k}\binom{z-1}{z-1-k}+\binom{2n+1-z}{n+2-z+k}\binom{z-1}{z-2-k}\right] \\
&= \binom{2n-z}{n-(1+k)}\left[-\binom{2n+1-z}{n-k}\binom{z-1}{k}+\binom{2n+1-z}{n-k-1}\binom{z-1}{k+1}\right] \\
&= d(1+k).
\end{aligned}$$

To summarize, we proved that the coefficients in the polynomial (A.11) satisfy $d(1) = -d(z-1) > 0$, $d(2) = -d(z-2) > 0$, and so on. To visualize this structure of coefficients and its immediate consequence on proving (A.11), consider $n = 5$. When also $z = 5$, the coefficients are $\{270, 600, -600, -270\}$. When $z = 4$, the coefficients are $\{1260, 0, -1260\}$. Thus, it is clear from this structure of coefficients that, since each of these coefficients is multiplied respectively by $\left\{\left(\frac{1-p_A}{p_A}\right)^1, \left(\frac{1-p_A}{p_A}\right)^2, ..\right\}$, whenever $\frac{1-p_A}{p_A} \in (0,1)$ (or equivalently, $p_A > p_B$ or $p_A \in (1/2, 1)$) the weight given to the first and positive coefficients is greater than the weight given to the last and negative coefficients, resulting in an overall positive summation in (A.11). On the contrary, when $\frac{1-p_A}{p_A} \in (1,\infty)$, or equivalently, $p_A \in (0, 1/2)$, the opposite holds, and the overall summation is negative, thus demonstrating (A.11). This concludes the proof that $p_A \gtreqless p_B \Rightarrow \Delta(z) \gtreqless 0$ for $z \leq n$ and $z \neq 2$. The case $z = 2$ is special in that there is only one coefficient in the polynomial (A.11) and it equals 0. Hence, we analyze it next in **Step 2**.

Step 2. Consider $z = 2$. The summation in (A.10) has only one term ($b_A = 1$). However, since in (A.12) we proved that $sgn[d(b_A)] = sgn(z - 2b_A)$, when $z = 2$ and $b_A = 1$ we have that $d(1) = 0$, and thus the summation in (A.10) equals 0. Therefore,

$$\Delta(2) = \binom{2n-1}{n}\binom{2n-2}{n}(p_A p_B)^{n-1}\left(p_A^2 - p_B^2\right),$$

which concludes the proof that $p_A \gtreqless p_B \Rightarrow \Delta(z) \gtreqless 0$ for $z = 2$.

Step 3. Turning now to the case $z > n$, we have

$$\psi(z-1) = \binom{z-1}{z-1-n} P_A(z-1-n, n) + \binom{z-1}{n} P_A(n, z-1-n)$$

$$
\begin{aligned}
&+ \sum_{b_A=z-n}^{n-1} \binom{z-1}{b_A}\binom{2n-(z-1)}{n-b_A} P_A\left(b_A, z-1-b_A\right) \\
= \; & \binom{z-1}{n} P_A\left(n, z-1-n\right) + \sum_{b_A=z-n}^{n-1} \binom{z-1}{b_A}\binom{2n-z}{n-b_A} P_A\left(b_A, z-1-b_A\right) \\
&+ \sum_{b_A=z-n}^{n-1} \binom{z-1}{b_A}\binom{2n-z}{n-b_A-1} P_A(b_A, z-1-b_A) + \binom{z-1}{z-1-n} P_A(z-1-n, n) \\
= \; & \binom{z-1}{n} P_A\left(n, z-1-n\right) + \sum_{b_A=z-n}^{n-1} \binom{z-1}{b_A}\binom{2n-z}{n-b_A} P_A\left(b_A, z-1-b_A\right) \\
&+ \sum_{j=z-n+1}^{n} \binom{z-1}{j-1}\binom{2n-z}{n-j} P_A\left(j-1, z-j\right) + \binom{z-1}{z-1-n} P_A\left(z-1-n, n\right) \\
= \; & \sum_{b_A=z-n}^{n} \binom{z-1}{b_A}\binom{2n-z}{n-b_A} P_A\left(b_A, z-1-b_A\right) \\
&+ \sum_{j=z-n}^{n} \binom{z-1}{j-1}\binom{2n-z}{n-j} P_A\left(j-1, z-j\right).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\psi\left(z\right) &= \sum_{b_A=z-n}^{n} \left(\binom{z-1}{b_A} + \binom{z-1}{b_A-1}\right)\binom{2n-z}{n-b_A} P_A\left(b_A, z-b_A\right) \\
&= \sum_{b_A=z-n}^{n} \binom{z-1}{b_A-1}\binom{2n-z}{n-b_A} P_A\left(b_A, z-b_A\right) + \sum_{b_A=z-n}^{n} \binom{z-1}{b_A}\binom{2n-z}{n-b_A} P_A\left(b_A, z-b_A\right).
\end{aligned}
$$

So,

$$
\begin{aligned}
\Delta\left(z\right) = & - \sum_{b_A=z-n}^{n} \binom{z-1}{b_A-1}\binom{2n-z}{n-b_A} \left(P_A\left(b_A, z-b_A\right) - P_A\left(b_A-1, z-b_A\right)\right) \\
&+ \sum_{j=z-n}^{n} \binom{z-1}{b_A}\binom{2n-z}{n-b_A} \left(P_A\left(b_A, z-1-b_A\right) - P_A\left(b_A, z-b_A\right)\right) \\
&= - \sum_{b_A=z-n}^{n} \binom{z-1}{b_A-1}\binom{2n-z}{n-b_A}\binom{2n+1-z}{n+1-b_A} p_A^{n+1-b_A} p_B^{n+1-z+b_A} \\
&+ \sum_{b_A=z-n}^{n} \binom{z-1}{b_A}\binom{2n-z}{n-b_A}\binom{2n+1-z}{n-b_A} p_A^{n-b_A+1} p_B^{n-z+b_A+1}
\end{aligned}
$$

$$= p_A^{n+1} p_B^{n+1-z} \sum_{b_A=z-n}^{n} \binom{2n-z}{n-b_A} \left(\frac{p_B}{p_A}\right)^{b_A} \left[\binom{2n+1-z}{n-b_A}\binom{z-1}{b_A} - \binom{2n+1-z}{n+1-b_A}\binom{z-1}{b_A-1}\right], \tag{A.13}$$

where we used (A.8) and (A.9).

Therefore, the claim that $p_A \gtreqqless p_B \Rightarrow \Delta(z) \gtreqqless 0$ for $z > n$ and $z \neq 2n$ reads

$$p_A \gtreqqless p_B \Rightarrow \sum_{b_A=z-n}^{n} d(b_A) \left(\frac{1-p_A}{p_A}\right)^{b_A} \gtreqqless 0, \tag{A.14}$$

which is identical to (A.11), except for the range of b_A, which now goes from $z-n$ to n. Thus, the only structural step that changes in the proof of (A.14) with respect to that of (A.11) is that instead of having to prove that $d(1+k) = -d(z-1-k)$, we now have to prove that $d(z-n+k) = -d(n-k)$. Indeed, we have

$$\begin{aligned}
d(n-k) &= \binom{2n-z}{n-(n-k)}\left[\binom{2n+1-z}{k}\binom{z-1}{n-k} - \binom{2n+1-z}{k+1}\binom{z-1}{n-k-1}\right] \\
&= \binom{2n-z}{2n-z-k}\left[\binom{2n+1-z}{2n+1-z-k}\binom{z-1}{z-n+k-1} - \binom{2n+1-z}{2n-z-k}\binom{z-1}{z-n+k}\right] \\
&= \binom{2n-z}{n-(z-n+k)}\left[\binom{2n+1-z}{n-(z-n+k-1)}\binom{z-1}{z-n+k-1}\right] \\
&\quad -\binom{2n-z}{n-(z-n+k)}\left[\binom{2n+1-z}{n-(z-n+k)}\binom{z-1}{z-n+k}\right] \\
&= -d(z-n+k),
\end{aligned}$$

and the proof of (A.14) then follows exactly as that of (A.11). This concludes the proof that $p_A \gtreqqless p_B \Rightarrow \Delta(z) \gtreqqless 0$ for $z > n$ and $z \neq 2n$.

Step 4. Consider the case $z = 2n$. In (A.13), b_A only equals n and hence the square brackets of (A.13) equal 0. So, $\Delta(2n) = 0$ and the proof of the lemma is complete. □

Proof of Lemma 4. Following the same steps at the beginning of the Proof of Lemma 2, using Corollary 2, we can focus only on battles from $n_1 + 1$ to $n_1 + n_2$ in T and T^{mer} and decompose $\sum_{\omega \in \Omega_i^T} \Pr\{\omega\}$ with $i \in \{A, B\}$ similarly. The resulting simplified terms are denoted by sWE_M^T and $sWE_M^{T^{mer}}$, where

$$sWE_M^T = \sum_{b_A=\max\{n_1-n,0\}}^{\min\{n_1,n\}} \binom{n_1}{b_A} p_A^{b_A} p_B^{n_1-b_A} \sum_{m=n_1+1}^{n_1+n_2} \phi^m(b_A, n_1-b_A), \tag{A.15}$$

and

$$sWE_M^{T^{mer}} = \sum_{m=n_1+1}^{n_1+n_2} \phi^m(0,0), \tag{A.16}$$

where the definition of $\phi^m(\cdot,\cdot)$ is in $(A.2)$. Note that the min and max operators in the first summation of (A.15) ensure that we are taking into account only non-trivial battles (i.e., no team has won more than n battles), while in (A.16) these battles are never trivial. Therefore, WE in T^{mer} is larger (equal) [smaller] than in T if and only if $sWE_M^{T^{mer}} > (=)[<] sWE_M^T$.

Using (10) and (A.2) into (A.15), we obtain that sWE_M^T equals

$$\begin{aligned}
& \sum_{b_A=\max\{n_1-n,0\}}^{\min\{n_1,n\}} \binom{n_1}{b_A} p_A^{b_A} p_B^{n_1-b_A} \sum_{m=n_1+1}^{n_1+n_2} \phi^m(b_A, n_1-b_A) \\
= & \sum_{b_A=\max\{n_1-n,0\}}^{\min\{n_1,n\}} \binom{n_1}{b_A} p_A^{b_A} p_B^{n_1-b_A} \sum_{m=n_1+1}^{n_1+n_2} \left[\begin{array}{c} P_A(b_A, n_1-b_A)\, x_A^m(b_A, n_1-b_A) \\ +P_B(b_A, n_1-b_A)\, x_B^m(b_A, n_1-b_A) \end{array} \right] \\
= & \sum_{b_A=\max\{n_1-n,0\}}^{\min\{n_1,n\}} \binom{n_1}{b_A} p_A^{b_A} p_B^{n_1-b_A} \binom{2n-n_1}{n-b_A} p_A^{n-b_A} p_B^{n-(n_1-b_A)} \sum_{m=n_1+1}^{n_1+n_2} \left[\begin{array}{c} P_A(b_A, n_1-b_A)\, \mu_A^m \\ +P_B(b_A, n_1-b_A)\, \mu_B^m \end{array} \right] \\
= & \; p_A^n p_B^n \sum_{b_A=\max\{n_1-n,0\}}^{\min\{n_1,n\}} \binom{n_1}{b_A} \binom{2n-n_1}{n-b_A} \sum_{m=n_1+1}^{n_1+n_2} \left[P_A(b_A, n_1-b_A)\, \mu_A^m + P_B(b_A, n_1-b_A)\, \mu_B^m \right] \\
= & \; p_A^n p_B^n \sum_{b_A=\max\{n_1-n,0\}}^{\min\{n_1,n\}} \binom{n_1}{b_A} \binom{2n-n_1}{n-b_A} \sum_{m=n_1+1}^{n_1+n_2} \left[P_A(b_A, n_1-b_A)\, (\mu_A^m - \mu_B^m) + \mu_B^m \right] \\
= & \; p_A^n p_B^n \psi(n_1) \sum_{m=n_1+1}^{n_1+n_2} [\mu_A^m - \mu_B^m] + \xi(n_1) \sum_{m=n_1+1}^{n_1+n_2} \mu_B^m,
\end{aligned}$$

where ψ is defined in (12) and ξ is defined in (A.4). As shown in the Proof of Lemma 2, ξ is constant and equal to $\binom{2n}{n}$. Thus, we obtain

$$\frac{sWE_M^T}{p_A^n p_B^n} = \psi(n_1) \sum_{m=n_1+1}^{n_1+n_2} [\mu_A^m - \mu_B^m] + \binom{2n}{n} \sum_{m=n_1+1}^{n_1+n_2} \mu_B^m. \tag{A.17}$$

Applying similar steps to (A.16), we have

$$\begin{aligned}
sWE_M^{T^{mer}} &= \sum_{m=n_1+1}^{n_1+n_2} \left[P_A(0,0)\, x_A^m(0,0) + P_B(0,0)\, x_B^m(0,0) \right] \\
&= \binom{2n}{n} p_A^n p_B^n \sum_{m=n_1+1}^{n_1+n_2} \left[P_A(0,0)\, \mu_A^m + P_B(0,0)\, \mu_B^m \right]
\end{aligned}$$

$$= \binom{2n}{n} p_A^n p_B^n \left[P_A(0,0) \sum_{m=n_1+1}^{n_1+n_2} (\mu_A^m - \mu_B^m) + \sum_{m=n_1+1}^{n_1+n_2} \mu_B^m \right],$$

so that

$$\frac{sWE_M^{T^{mer}}}{p_A^n p_B^n} = \psi(0) \sum_{m=n_1+1}^{n_1+n_2} [\mu_A^m - \mu_B^m] + \binom{2n}{n} \sum_{m=n_1+1}^{n_1+n_2} \mu_B^m. \tag{A.18}$$

Comparing (A.17) and (A.18) establishes the lemma. □

Proof of Lemma 5. In our setup, recalling that the discouragement effect does not apply, the results on one-shot contests can be used to derive battle-specific efforts, as soon as the appropriate value of the prize is specified. We build on the canonical equilibrium characterization of Baye, Kovenock, and de Vries (1996). In a one-shot two-player contest under (2), with marginal costs $c_A\gamma_m$ and $c_B\gamma_m$ ($c_A \leq c_B$) and common prize $\tilde{\nu}$, player A bids $x_A(0,0)$, which is uniformly distributed on $[0, \tilde{\nu}/(c_B\gamma_m)]$, and player B bids $x_B(0,0)$, which equals 0 with probability $1 - c_A/c_B$ and is uniformly distributed on $[0, \tilde{\nu}/(c_B\gamma_m)]$ with probability c_A/c_B. As in our setup battle prizes are in common, we can use the above characterization to derive that the ex-ante equilibrium probabilities of victory of A and B are as in (19). We next calculate conditional expected efforts. (To lighten the notation, we use x_A for $x_A(0,0)$ and x_B for $x_B(0,0)$.)

Using Bayes' rule, the effort of player A conditional on A winning the battle is:

$$\Pr\{x_A \leq y | A \text{ wins}\} = \frac{\Pr\{A \text{ wins} | x_A \leq y\} \Pr\{x_A \leq y\}}{\Pr\{A \text{ wins}\}}.$$

Note that x_A conditional on $x_A \leq y$ is a uniform distribution on $[0, y]$. Therefore,

$$\Pr\{A \text{ wins} | x_A \leq y\} = \int_0^y \frac{1}{y} \left(1 - \frac{c_A}{c_B} + \frac{c_A}{c_B} \frac{z}{\frac{\tilde{v}}{c_B\gamma_m}}\right) dz = 1 - \frac{c_A}{c_B} + \frac{\gamma_m c_A}{2\tilde{v}} y;$$

and since $\Pr\{x_A \leq y\} = \frac{c_B\gamma_m}{\tilde{v}} y$ and $\Pr\{A \text{ wins}\} = \frac{2c_B - c_A}{2c_B}$, we have

$$\Pr\{x_A \leq y | A \text{ wins}\} = \frac{\left(1 - \frac{c_A}{c_B} + \frac{\gamma_m c_A}{2\tilde{v}} y\right)\left(\frac{c_B\gamma_m}{\tilde{v}} y\right)}{\left(\frac{2c_B - c_A}{2c_B}\right)}.$$

Thus,

$$E[x_A | A \text{ wins}] = \int_0^{\frac{\tilde{v}}{c_B\gamma_m}} y d\left(\frac{\left(1 - \frac{c_A}{c_B} + \frac{\gamma_m c_A}{2\tilde{v}} y\right)\left(\frac{c_B\gamma_m}{\tilde{v}} y\right)}{\left(\frac{2c_B - c_A}{2c_B}\right)}\right) = \frac{1}{3\gamma_m c_B} \cdot \frac{3c_B - c_A}{2c_B - c_A} \tilde{v} = \mu_{A|A}^m \tilde{v},$$

where $\mu^m_{A|A}$ is in (17). Proceeding similarly to the above, the corresponding quantities can be computed for the A-player's effort conditional on her losing the battle—$\mu^m_{A|B}\tilde{\nu}$—and for the B-player's effort conditional on her winning or losing the battle—$\mu^m_{B|B}\tilde{\nu}$ and $\mu^m_{B|A}\tilde{\nu}$—leading to (17).

Note that, in battle m and state (b_A, b_B), the battle prize $\tilde{\nu}$ is equal to $\binom{2n-(b_A+b_B)}{n-b_A} p_A^{n-b_A} p_B^{n-b_B}$ (recall that the battle prize is the pivotal probability of battle m), thus proving (18). □

Proof of Proposition 3. Corollary 1 holds under mixed strategies too and allows us to neglect the first l battles and the last $2n+1-(l-1)-n_l$ battles in the comparisons of WE^T and $WE^{T^{ext}}$. Focusing on a generic battle m of the remaining $n_l - 1$ battles, we can decompose $\sum_{\omega \in \Omega_A^T} \Pr\{\omega\}$ as in the Proof of Lemma 2 with one crucial change: the formulas for WE now needs to take into account Remark 2 because the equilibrium is in mixed strategies. Thus, $\phi(b_A, b_B)$ is substituted by

$$
\begin{aligned}
\tilde{\phi}^m(b_A, b_B) \equiv \quad & p_A \cdot E\left[x_A^m(b_A, b_B)\,|A \text{ wins battle } m\right] P_A(b_A+1, b_B) \\
+ \quad & p_B \cdot E\left[x_A^m(b_A, b_B)\,|B \text{ wins battle } m\right] P_A(b_A, b_B+1) \\
+ \quad & p_B \cdot E\left[x_B^m(b_A, b_B)\,|B \text{ wins battle } m\right] P_B(b_A, b_B+1) \\
+ \quad & p_A \cdot E\left[x_B^m(b_A, b_B)\,|A \text{ wins battle } m\right] P_B(b_A+1, b_B),
\end{aligned}
\tag{A.19}
$$

where the value of $E[x_i^m(b_A, b_B)|j \text{ wins battle } m]$ is in (18), that of $P_A(b_A, b_B)$ is in (7), with $P_B(b_A, b_B) = 1 - P_A(b_A, b_B)$, and that of p_i is in (19). Note that $(A.19)$ differs from $(A.2)$ because conditioning on the identity of the winner of battle m affects the expectation of the effort of battle m itself, and such expectation enters WE.

Therefore, the appropriate formulas for WE in $(A.1)$ and $(A.3)$ in the setup with mixed strategies now have $\tilde{\phi}^m(b_A, b_B)$, rather than $\phi^m(b_A, b_B)$. That is, for extractions, $(A.1)$ and $(A.3)$ are replaced by

$$
\tilde{s}WE_E^{T^{ext}} \equiv \sum_{b_A=\max\{l-n,0\}}^{\min\{n,l\}} \binom{l}{b_A} p_A^{b_A} p_B^{l-b_A} \sum_{m=l+1}^{l+n_l-1} \tilde{\phi}^m(b_A, l-b_A), \tag{A.20}
$$

and

$$
\tilde{s}WE_E^T \equiv \sum_{b_A=\max\{l-1-n,0\}}^{\min\{n,l-1\}} \binom{l-1}{b_A} p_A^{b_A} p_B^{l-1-b_A} \sum_{m=l+1}^{l+n_l-1} \tilde{\phi}^m(b_A, l-1-b_A): \tag{A.21}
$$

WE in T is larger (equal) [smaller] than in T^{ext} if and only if $\tilde{s}WE_E^T > (=)[<]\tilde{s}WE_E^{T^{ext}}$.

We can rewrite $\tilde{s}WE_E^{T^{ext}}$ as

$$\tilde{s}WE_E^{T^{ext}} = p_A^n p_B^n \sum_{b_A=\max\{l-n,0\}}^{\min\{n,l\}} \binom{l}{b_A}\binom{2n-l}{n-b_A} \cdot \sum_{m=l+1}^{l+n_l-1} \begin{pmatrix} p_A P_A(b_A+1,l-b_A)\mu_{A|A}^m + p_B P_A(b_A,l-b_A+1)\mu_{A|B}^m \\ +p_A P_B(b_A+1,l-b_A)\mu_{B|A}^m + p_B P_B(b_A,l-b_A+1)\mu_{B|B}^m \end{pmatrix} \quad \text{(A.22)}$$

and $\tilde{s}WE_E^T$ as

$$\tilde{s}WE_E^T = p_A^n p_B^n \sum_{b_A=\max\{l-1-n,0\}}^{\min\{n,l-1\}} \binom{l-1}{b_A}\binom{2n-(l-1)}{n-b_A} \cdot \sum_{m=l+1}^{l+n_l-1} \begin{pmatrix} p_A P_A\left(b_A+1,l-1-b_A\right)\mu_{A|A}^m + p_B P_A\left(b_A,l-b_A\right)\mu_{A|B}^m \\ +p_A P_B(b_A+1,l-1-b_A)\mu_{B|A}^m + p_B P_B(b_A,l-b_A)\mu_{B|B}^m \end{pmatrix} \text{(A.23)}$$

By $c_A = c_B$, we obtain $p_A = p_B = \frac{1}{2}$ and, by (17),

$$\mu_{A|A}^m \equiv \frac{2}{3c_B\gamma_m}, \mu_{A|B}^m \equiv \frac{1}{3c_B\gamma_m}, \mu_{B|A}^m \equiv \frac{1}{3c_B\gamma_m}, \mu_{B|B}^m \equiv \frac{2}{3c_B\gamma_m}.$$

Thus, $\tilde{s}WE_E^{T^{ext}}$ and $\tilde{s}WE_E^T$ can be simplified. In particular, $\tilde{s}WE_E^{T^{ext}}$ becomes

$$\begin{aligned} & \frac{1}{3c_B}\left(\frac{1}{2}\right)^{2n+1} \sum_{b_A=\max\{l-n,0\}}^{\min\{n,l\}} \binom{l}{b_A}\binom{2n-l}{n-b_A} \\ & \cdot \sum_{m=l+1}^{l+n_l-1} \frac{2P_A(b_A+1,l-b_A)+P_A(b_A,l-b_A+1)+P_B(b_A+1,l-b_A)+2P_B(b_A,l-b_A+1)}{\gamma_m} \\ = & \frac{1}{3c_B}\left(\frac{1}{2}\right)^{2n+1} \cdot \left(\sum_{m=l+1}^{l+n_l-1} \frac{1}{\gamma_m}\right) \\ & \cdot \left(\sum_{b_A=\max\{l-n,0\}}^{\min\{n,l\}} \binom{l}{b_A}\binom{2n-l}{n-b_A} (P_A(b_A+1,l-b_A)-P_A(b_A,l-b_A+1)+3)\right), \end{aligned}$$

where we used $P_A(x,y)+P_B(x,y)=1$, and similarly $\tilde{s}WE_E^T$ becomes

$$\frac{1}{3c_B}\left(\frac{1}{2}\right)^{2n+1} \cdot \left(\sum_{m=l+1}^{l+n_l-1} \frac{1}{\gamma_m}\right)$$

$$\cdot\left(\sum_{b_A=\max\{l-1-n,0\}}^{\min\{n,l-1\}}\binom{l-1}{b_A}\binom{2n-(l-1)}{n-b_A}(P_A(b_A+1,l-1-b_A)-P_A(b_A,l-b_A)+3)\right).$$

Therefore, $\tilde{s}WE_E^{T^{ext}} \geq \tilde{s}WE_E^{T}$ if and only if

$$\begin{aligned}&\sum_{b_A=\max\{l-n,0\}}^{\min\{n,l\}}\tbinom{l}{b_A}\tbinom{2n-l}{n-b_A}\left(P_A\left(b_A+1,l-b_A\right)-P_A\left(b_A,l-b_A+1\right)+3\right)\\ \geq&\sum_{b_A=\max\{l-1-n,0\}}^{\min\{n,l-1\}}\tbinom{l-1}{b_A}\tbinom{2n-(l-1)}{n-b_A}\left(P_A\left(b_A+1,l-1-b_A\right)-P_A\left(b_A,l-b_A\right)+3\right).\end{aligned} \tag{A.24}$$

First, proceed under $l \leq n$, so $(A.24)$ becomes

$$\begin{aligned}&\sum_{b_A=0}^{l}\tbinom{l}{b_A}\tbinom{2n-l}{n-b_A}\left(P_A\left(b_A+1,l-b_A\right)-P_A\left(b_A,l-b_A+1\right)+3\right)\\ \geq&\sum_{b_A=0}^{l-1}\tbinom{l-1}{b_A}\tbinom{2n-(l-1)}{n-b_A}\left(P_A\left(b_A+1,l-1-b_A\right)-P_A\left(b_A,l-b_A\right)+3\right).\end{aligned}$$

Since

$$\sum_{b_A=0}^{l}\binom{l}{b_A}\binom{2n-l}{n-b_A}=\binom{2n}{n}=\sum_{b_A=0}^{l-1}\binom{l-1}{b_A}\binom{2n-(l-1)}{n-b_A} \tag{A.25}$$

(by Vandermonde's identity), the above (and hence $(A.24)$) is equivalent to

$$\begin{aligned}&\sum_{b_A=0}^{l}\tbinom{l}{b_A}\tbinom{2n-l}{n-b_A}\left(P_A\left(b_A+1,l-b_A\right)-P_A\left(b_A,l-b_A+1\right)\right)\\ \geq&\sum_{b_A=0}^{l-1}\tbinom{l-1}{b_A}\tbinom{2n-(l-1)}{n-b_A}\left(P_A\left(b_A+1,l-1-b_A\right)-P_A\left(b_A,l-b_A\right)\right).\end{aligned} \tag{A.26}$$

Note that $P_A\left(b_A+1,l-b_A\right)-P_A\left(b_A,l-b_A+1\right)$ is the pivotal probability of a battle played in state $(b_A, l-b_A)$, so it equals

$$\binom{2n-l}{n-b_A}p_A^{n-b_A}p_B^{2n-l-(n-b_A)}=\binom{2n-l}{n-b_A}\left(\frac{1}{2}\right)^{2n-l},$$

and $P_A\left(b_A+1,l-1-b_A\right)-P_A\left(b_A,l-b_A\right)$ is the pivotal probability of a battle played in state $(b_A, l-1-b_A)$, so it equals

$$\binom{2n-l+1}{n-b_A}\left(\frac{1}{2}\right)^{2n-l+1}.$$

Thus, $(A.26)$ simplifies to

$$2\sum_{b_A=0}^{l}\binom{l}{b_A}\binom{2n-l}{n-b_A}^2 \geq \sum_{b_A=0}^{l-1}\binom{l-1}{b_A}\binom{2n-(l-1)}{n-b_A}^2. \tag{A.27}$$

Applying Pascal's rule, we obtain

$$2\sum_{b_A=0}^{l-1}\binom{l-1}{b_A}\binom{2n-l}{n-b_A}^2 + 2\sum_{b_A=1}^{l}\binom{l-1}{b_A-1}\binom{2n-l}{n-b_A}^2 \geq \sum_{b_A=0}^{l-1}\binom{l-1}{b_A}\left(\binom{2n-l}{n-b_A}+\binom{2n-l}{n-b_A-1}\right)^2,$$

where we erased the null terms of the LHS (i.e., the one with $b_A = l$ in the first summation and with $b_A = 0$ in the second). Hence, a change of variable in the second summation yields

$$2\sum_{b_A=0}^{l-1}\binom{l-1}{b_A}\binom{2n-l}{n-b_A}^2 + 2\sum_{b_A=0}^{l-1}\binom{l-1}{b_A}\binom{2n-l}{n-b_A-1}^2 \geq \sum_{b_A=0}^{l-1}\binom{l-1}{b_A}\left(\binom{2n-l}{n-b_A}+\binom{2n-l}{n-b_A-1}\right)^2$$

or

$$\sum_{b_A=0}^{l-1}\binom{l-1}{b_A}\left(\binom{2n-l}{n-b_A}-\binom{2n-l}{n-b_A-1}\right)^2 \geq 0, \tag{A.28}$$

which proves $(A.27)$. Note that $l = 1$ is the only case when the LHS of $(A.28)$ equals 0.

Second, we now proceed instead under $2n-1 \geq l-1 \geq n$. Similar steps to the case of $l \leq n$, and in particular, using

$$\sum_{b_A=l-n}^{n}\binom{l}{b_A}\binom{2n-l}{n-b_A} = \sum_{B=0}^{2n-l}\binom{2n-l}{B}\binom{l}{n-B} = \binom{2n}{n} = \sum_{b_A=l-n-1}^{n}\binom{l-1}{b_A}\binom{2n-(l-1)}{n-b_A}, \tag{A.29}$$

we obtain that $(A.24)$ becomes

$$2\sum_{b_A=l-n}^{n}\binom{l}{b_A}\binom{2n-l}{n-b_A}^2 \geq \sum_{b_A=l-n-1}^{n}\binom{l-1}{b_A}\binom{2n-(l-1)}{n-b_A}^2. \tag{A.30}$$

Applying Pascal's rule to the LHS of $(A.30)$, we obtain

$$2\sum_{b_A=l-n}^{n}\binom{l}{b_A}\binom{2n-l}{n-b_A}^2 = 2\sum_{b_A=l-n}^{n}\left(\binom{l-1}{b_A}+\binom{l-1}{b_A-1}\right)\binom{2n-l}{n-b_A}^2.$$

Recalling that $2n-1 \geq l-1 \geq n$, we have that $b_A - 1 > 0$ as $b_A - 1 \geq l - n \geq 1$ and that $b_A \leq l-1$ as $b_A \leq n \leq l-1$. Hence, the binomial coefficients on the RHS of the above-displayed equation are standard.

Using Pascal's rule to the RHS of $(A.30)$, we obtain that $\sum_{b_A=l-n-1}^{n} \binom{l-1}{b_A}\binom{2n-(l-1)}{n-b_A}^2$ equals

$$\begin{aligned}
& \sum_{b_A=l-n}^{n-1} \binom{l-1}{b_A}\binom{2n-(l-1)}{n-b_A}^2 + \binom{l-1}{l-n-1}\binom{2n-(l-1)}{2n-(l-1)}^2 + \binom{l-1}{n}\binom{2n-(l-1)}{0}^2 \\
= & \sum_{b_A=l-n}^{n-1} \binom{l-1}{b_A}\binom{2n-(l-1)}{n-b_A}^2 + \binom{l-1}{l-n-1} + \binom{l-1}{n} \\
= & \sum_{b_A=l-n}^{n-1} \binom{l-1}{b_A}\left(\binom{2n-l}{n-b_A} + \binom{2n-l}{n-b_A-1}\right)^2 + \binom{l-1}{l-n-1} + \binom{l-1}{n}. \qquad (A.31)
\end{aligned}$$

We next bring the last two addends above back into the summation by noting that

$$\begin{aligned}
& \binom{l-1}{l-n-1}\left(\binom{2n-l}{n-(l-n-1)} + \binom{2n-l}{n-(l-n-1)-1}\right)^2 \\
= & \binom{l-1}{l-n-1}\left(\binom{2n-l}{2n-l+1} + \binom{2n-l}{2n-l}\right)^2 \\
= & \binom{l-1}{l-n-1}(0+1)^2,
\end{aligned}$$

by $\binom{n}{k} = 0$ with $k > n$ (see Knuth, 1997; page 55) and that,

$$\begin{aligned}
\binom{l-1}{n}\left(\binom{2n-l}{n-n} + \binom{2n-l}{n-n-1}\right)^2 & = \binom{l-1}{n}\left(\binom{2n-l}{0} + \binom{2n-l}{-1}\right)^2 \\
& = \binom{l-1}{n}(1+0)^2,
\end{aligned}$$

by $\binom{n}{k} = 0$ with $k < 0$ (see, Knuth, 1997; page 55). So, the RHS of $(A.31)$ equals

$$\sum_{b_A=l-n-1}^{n} \binom{l-1}{b_A}\left(\binom{2n-l}{n-b_A} + \binom{2n-l}{n-b_A-1}\right)^2.$$

Hence, $(A.30)$ is equivalent to

$$\begin{aligned}
& 2\sum_{b_A=l-n}^{n} \binom{l-1}{b_A}\binom{2n-l}{n-b_A}^2 + 2\sum_{b_A=l-n}^{n} \binom{l-1}{b_A-1}\binom{2n-l}{n-b_A}^2 \\
\geq & \sum_{b_A=l-n-1}^{n} \binom{l-1}{b_A}\left(\binom{2n-l}{n-b_A} + \binom{2n-l}{n-b_A-1}\right)^2. \qquad (A.32)
\end{aligned}$$

Note that the element with $b_A = l-n$ on the second summation on the LHS is double that with $b_A = l - n - 1$ on the summation on the RHS. Also, we can add the null term with $b_A = n + 1$ to the second summation on the LHS. Thus, a sufficient condition for $(A.32)$ is

$$2 \sum_{b_A = l-n}^{n} \binom{l-1}{b_A} \binom{2n-l}{n-b_A}^2 + 2 \sum_{b_A = l-n+1}^{n+1} \binom{l-1}{b_A - 1} \binom{2n-l}{n-b_A}^2$$
$$\geq \sum_{b_A = l-n}^{n} \binom{l-1}{b_A} \left(\binom{2n-l}{n-b_A} + \binom{2n-l}{n-b_A-1} \right)^2 .$$

A change of variable from b_A to $b_A - 1$ in the second summation on the LHS yields

$$2 \sum_{b_A = l-n}^{n} \binom{l-1}{b_A} \binom{2n-l}{n-b_A}^2 + 2 \sum_{b_A = l-n}^{n} \binom{l-1}{b_A} \binom{2n-l}{n-b_A-1}^2$$
$$\geq \sum_{b_A = l-n}^{n} \binom{l-1}{b_A} \left(\binom{2n-l}{n-b_A} + \binom{2n-l}{n-b_A-1} \right)^2 ,$$

or

$$\sum_{b_A = l-n}^{n} \binom{l-1}{b_A} \left(\binom{2n-l}{n-b_A} - \binom{2n-l}{n-b_A-1} \right)^2 \geq 0,$$

which proves $(A.30)$. □

Proof of Proposition 4. Corollary 2 holds under mixed strategies too and allow us to neglect the first n_1 battles and all battles after the first $n_1 + n_2$ in the comparisons of WE^T and $WE^{T^{mer}}$. Similarly to the Proof of Proposition 3, we now use $\tilde{\phi}^m (b_A, b_B)$ in $(A.19)$ and replace the formulas for WE in $(A.15)$ and $(A.16)$ in the Proof of Lemma 4 by

$$\tilde{s}WE_M^T = \sum_{b_A = \max\{n_1 - n, 0\}}^{\min\{n_1, n\}} \binom{n_1}{b_A} p_A^{b_A} p_B^{n_1 - b_A} \sum_{m = n_1 + 1}^{n_1 + n_2} \tilde{\phi}^m (b_A, n_1 - b_A) , \tag{A.33}$$

and

$$\tilde{s}WE_M^{T^{mer}} = \sum_{m = n_1 + 1}^{n_1 + n_2} \tilde{\phi}^m (0, 0) ; \tag{A.34}$$

so WE in T^{mer} is larger (equal) [smaller] than in T if and only if $\tilde{s}WE_M^{T^{mer}} > (=)[<] \tilde{s}WE_M^T$. Hence, a merger decreases WE if and only if

$$\sum_{m = n_1 + 1}^{n_1 + n_2} \tilde{\phi}^m (0, 0) \leq \sum_{b_A = \max\{n_1 - n, 0\}}^{\min\{n_1, n\}} \binom{n_1}{b_A} p_A^{b_A} p_B^{n_1 - b_A} \sum_{m = n_1 + 1}^{n_1 + n_2} \tilde{\phi}^m (b_A, n_1 - b_A) . \tag{A.35}$$

Using $(A.19)$ and (18), $(A.35)$ is equivalent to

$$p_A^n p_B^n \binom{2n}{n} \sum_{m=n_1+1}^{n_1+n_2} \left(p_A \mu_{A|A}^m P_A(1,0) + p_B \mu_{A|B}^m P_A(0,1) + p_A \mu_{B|A}^m P_B(1,0) + p_B \mu_{B|B}^m P_B(0,1)\right)$$
$$\leq \sum_{b_A=\max\{n_1-n,0\}}^{\min\{n_1,n\}} \binom{n_1}{b_A} p_A^{b_A} p_B^{n_1-b_A}$$
$$\sum_{m=n_1+1}^{n_1+n_2} \binom{2n-n_1}{n-b_A} p_A^{n-b_A} p_B^{n-(n_1-b_A)} \left(\begin{array}{c} p_A \mu_{A|A}^m P_A(b_A+1, n_1-b_A) + p_B \mu_{A|B}^m P_A(b_A, n_1-b_A+1) \\ +p_A \mu_{B|A}^m P_B(b_A+1, n_1-b_A) + p_B \mu_{B|B}^m P_B(b_A, n_1-b_A+1) \end{array} \right).$$

Next, by $c_A = c_B$ so that $p_A = p_B = 1/2$, and by (17) so that $\mu_{A|A}^m = \mu_{B|B}^m = 2\mu_{A|B}^m = 2\mu_{B|A}^m$, $(A.35)$ is equivalent to

$$\binom{2n}{n} \left(2P_A(1,0) + P_A(0,1) + P_B(1,0) + 2P_B(0,1)\right) \sum_{m=n_1+1}^{n_1+n_2} \frac{1}{\gamma_m}$$
$$\leq \sum_{b_A=\max\{n_1-n,0\}}^{\min\{n_1,n\}} \binom{n_1}{b_A} \binom{2n-n_1}{n-b_A} \left(\begin{array}{c} 2P_A(b_A+1, n_1-b_A) + P_A(b_A, n_1-b_A+1) \\ +P_B(b_A+1, n_1-b_A) + 2P_B(b_A, n_1-b_A+1) \end{array} \right) \sum_{m=n_1+1}^{n_1+n_2} \frac{1}{\gamma_m}.$$

As $P_A(x,y) + P_B(x,y) = 1$, we further obtain that $(A.35)$ is equivalent to

$$\binom{2n}{n} \left(3 + P_A(1,0) - P_A(0,1)\right)$$
$$\leq \sum_{b_A=\max\{n_1-n,0\}}^{\min\{n_1,n\}} \binom{n_1}{b_A} \binom{2n-n_1}{n-b_A} \left(3 + P_A(b_A+1, n_1-b_A) - P_A(b_A, n_1-b_A+1)\right).$$

Note that, regardless of whether $n_1 \geq n$ or $n_1 < n$,

$$\sum_{b_A=\max\{n_1-n,0\}}^{\min\{n_1,n\}} \binom{n_1}{b_A} \binom{2n-n_1}{n-b_A} = \binom{2n}{n},$$

so that $(A.35)$ is equivalent to

$$\binom{2n}{n} \left(P_A(1,0) - P_A(0,1)\right)$$
$$\leq \sum_{b_A=\max\{n_1-n,0\}}^{\min\{n_1,n\}} \binom{n_1}{b_A} \binom{2n-n_1}{n-b_A} \left(P_A(b_A+1, n_1-b_A) - P_A(b_A, n_1-b_A+1)\right).$$

$P_A(b_A+1, n_1-b_A) - P_A(b_A, n_1-b_A+1)$ is the pivotal probability of a battle played in state $(b_A, n_1 - b_A)$, so it equals

$$\binom{2n-n_1}{n-b_A} p_A^{n-b_A} p_B^{2n-n_1-(n-b_A)} = \binom{2n-n_1}{n-b_A}\left(\frac{1}{2}\right)^{2n-n_1}.$$

So, we obtain that $(A.35)$ implies that a merger decreases WE if and only if

$$\binom{2n}{n}^2 \leq \sum_{b_A=\max\{n_1-n,0\}}^{\min\{n_1,n\}} \binom{n_1}{b_A}\binom{2n-n_1}{n-b_A}^2 2^{n_1}. \tag{A.36}$$

Consider now the case of $n_1 \leq n$, so that $(A.36)$ is

$$\binom{2n}{n}^2 \leq \sum_{b_A=0}^{n_1} \binom{n_1}{b_A}\binom{2n-n_1}{n-b_A}^2 2^{n_1},$$

which, by

$$\binom{2n}{n} = \sum_{b_A=0}^{n_1} \binom{n_1}{b_A}\binom{2n-n_1}{n-b_A} \text{ and } 2^{n_1} = \sum_{b_A=0}^{n_1}\binom{n_1}{b_A},$$

is equivalent to

$$\sum_{b_A=0}^{n_1} \frac{\binom{n_1}{b_A}}{\sum_{b_A=0}^{n_1}\binom{n_1}{b_A}}\binom{2n-n_1}{n-b_A}^2 - \left(\sum_{b_A=0}^{n_1} \frac{\binom{n_1}{b_A}}{\sum_{b_A=0}^{n_1}\binom{n_1}{b_A}}\binom{2n-n_1}{n-b_A}\right)^2 \geq 0. \tag{A.37}$$

The above-displayed inequality holds as its LHS is the variance of the random variable that takes values $\binom{2n-n_1}{n-b_A}$ for $b_A \in \{0,..,n_1\}$, with probability $\binom{n_1}{b_A}/\sum_{b_A=0}^{n_1}\binom{n_1}{b_A}$. Condition $(A.37)$ holds with equality only if $n_1 = 1$.

Consider now the case of $n_1 > n$, so that $(A.36)$ becomes

$$\binom{2n}{n}^2 \leq \sum_{b_A=n_1-n}^{n} \binom{n_1}{b_A}\binom{2n-n_1}{n-b_A}^2 2^{n_1}.$$

Using $\binom{2n}{n} = \sum_{b_A=n_1-n}^{n} \binom{n_1}{b_A}\binom{2n-n_1}{n-b_A}$, the above is equivalent to

$$\sum_{b_A=n_1-n}^{n} \binom{n_1}{b_A}\binom{2n-n_1}{n-b_A}^2 2^{n_1} - \left(\sum_{b_A=n_1-n}^{n} \binom{n_1}{b_A}\binom{2n-n_1}{n-b_A}\right)^2 \geq 0,$$

or

$$\sum_{b_A=n_1-n}^{n} \frac{\binom{n_1}{b_A}}{\sum\limits_{b_A=n_1-n}^{n} \binom{n_1}{b_A}} \binom{2n-n_1}{n-b_A}^2 \frac{2^{n_1}}{\sum\limits_{b_A=n_1-n}^{n} \binom{n_1}{b_A}} - \left(\sum_{b_A=n_1-n}^{n} \frac{\binom{n_1}{b_A}}{\sum\limits_{b_A=n_1-n}^{n} \binom{n_1}{b_A}} \binom{2n-n_1}{n-b_A} \right)^2 \geq 0. \tag{A.38}$$

As

$$2^{n_1} = \sum_{b_A=0}^{n_1} \binom{n_1}{b_A} = \sum_{b_A=0}^{n_1-n-1} \binom{n_1}{b_A} + \sum_{b_A=n_1-n}^{n} \binom{n_1}{b_A} + \sum_{b_A=n+1}^{n_1} \binom{n_1}{b_A} \geq \sum_{b_A=n_1-n}^{n} \binom{n_1}{b_A},$$

a sufficient condition for $(A.38)$ is

$$\sum_{b_A=n_1-n}^{n} \frac{\binom{n_1}{b_A}}{\sum\limits_{b_A=n_1-n}^{n} \binom{n_1}{b_A}} \binom{2n-n_1}{n-b_A}^2 - \left(\sum_{b_A=n_1-n}^{n} \frac{\binom{n_1}{b_A}}{\sum\limits_{b_A=n_1-n}^{n} \binom{n_1}{b_A}} \binom{2n-n_1}{n-b_A} \right)^2 \geq 0,$$

which holds true by the same variance argument we discussed right after equation $(A.37)$. □

Proof of Proposition 5. Note that, from (17), simple algebra yields

$$\mu^m_{B|B} \geq \mu^m_{A|A} > \mu^m_{A|B} \geq \mu^m_{B|A}, \tag{A.39}$$

with strict inequalities throughout as soon as $c_A < c_B$. We will use $(A.39)$ repeatedly below.

First, we show that $WE^{\{1,\dots,1\}} > WE^{\{1,\dots,1,2\}}$. By Corollary 1, we can neglect the first $2n$ battles for the comparison and focus only on expected efforts in the very last battle. Using (3), we then obtain that the contribution of the very last battle to WE is

$$\sum_{i\in\{A,B\}} \sum_{\omega\in\Omega_i^T} \Pr\{\omega\} \cdot E\left[x_i^{2n+1}\left(b_A^{2n+1}(\omega), b_B^{2n+1}(\omega)\right)|\omega\right]. \tag{A.40}$$

With a similar logic to that in the Proof of Lemma 2, we can decompose any path $\omega \in \Omega_A^T$ into the probability of reaching all possible states in which battle $2n+1$ is played with strictly positive efforts, multiplied by the probability of victory conditional on A winning the overall contest (and similarly for paths $\omega \in \Omega_B^T$ conditional on B winning the overall contest). Therefore, using $(A.19)$ in the Proof of Proposition 3 and the fact that the probability of reaching state (n,n) is $\binom{2n}{n}p_A^n p_B^n$, we obtain that $(A.40)$ in temporal structure $\{1,..,1\}$ is

$$\binom{2n}{n}p_A^n p_B^n \cdot \tilde{\phi}^{2n+1}(n,n)\cdot 1, \tag{A.41}$$

recalling that the prize of the very last battle is 1. Similarly, $(A.40)$ in temporal structure $\{1,...,1,2\}$ is

$$\binom{2n-1}{n-1}p_A^n p_B^{n-1}\cdot \tilde{\phi}^{2n+1}(n,n-1)\cdot p_B + \binom{2n-1}{n-1}p_A^{n-1}p_B^n \cdot \tilde{\phi}^{2n+1}(n-1,n)\cdot p_A, \tag{A.42}$$

recalling that the prize at state $(n,n-1)$ is p_B and at state $(n-1,n)$ is p_A.

Next, note that $(A.41)$ is equivalent to

$$\binom{2n-1}{n-1}p_A^{n-1}p_B^n \cdot \tilde{\phi}^{2n+1}(n,n)\, p_A + \binom{2n-1}{n-1}p_A^n p_B^{n-1}\cdot \tilde{\phi}^{2n+1}(n,n)\, p_B. \tag{A.43}$$

Subtracting $(A.42)$ from $(A.43)$, the sign $WE^{\{1,....,1\}} - WE^{\{1,...,1,2\}}$ is identical to the sign of

$$\left(\tilde{\phi}^{2n+1}(n,n) - \tilde{\phi}^{2n+1}(n-1,n)\right) + \left(\tilde{\phi}^{2n+1}(n,n) - \tilde{\phi}^{2n+1}(n,n-1)\right).$$

Finally, using (18) and $(A.19)$, the above-displayed term is identical to

$$\begin{aligned}
& p_A\mu_{A|A}^{2n+1} + p_B\mu_{B|B}^{2n+1} - p_A\mu_{A|A}^{2n+1}p_A - p_A\mu_{B|A}^{2n+1}p_B - p_B\mu_{B|B}^{2n+1} \\
& + p_A\mu_{A|A}^{2n+1} + p_B\mu_{B|B}^{2n+1} - p_A\mu_{A|A}^{2n+1} - p_B\mu_{A|B}^{2n+1}p_A - p_B\mu_{B|B}^{2n+1}p_B \\
= \; & p_A p_B\left(\mu_{A|A}^{2n+1} - \mu_{B|A}^{2n+1} + \mu_{B|B}^{2n+1} - \mu_{A|B}^{2n+1}\right),
\end{aligned} \tag{A.44}$$

where the last step uses $p_B = 1 - p_A$. This last term is positive from $(A.39)$, thus concluding the proof the $WE^{\{1,....,1\}} > WE^{\{1,...,1,2\}}$.

Second, we show that $WE^{\{2,1,....,1\}} > WE^{\{1,...,1\}}$. By Corollary 1, we can neglect all battles but the second one for the comparison. Proceeding as above, we obtain that the contribution of the second battle to WE in temporal structure $\{1,..,1\}$ is

$$p_A\tilde{\phi}^2(1,0)\binom{2n-1}{n-1}p_A^{n-1}p_B^n + p_B\tilde{\phi}^2(0,1)\binom{2n-1}{n-1}p_A^n p_B^{n-1},$$

recalling that the prize in state $(1,0)$ is $\binom{2n-1}{n-1}p_A^{n-1}p_B^n$ and in state $(0,1)$ is $\binom{2n-1}{n-1}p_A^n p_B^{n-1}$. Similarly, the contribution of the second battle to WE in temporal structure $\{2,1,...,1\}$ is

$$\tilde{\phi}^2(0,0)\binom{2n}{n}p_A^n p_B^n,$$

recalling that the prize at state $(0,0)$ is $\binom{2n}{n} p_A^n p_B^n$.

As $\binom{2n}{n} = 2\binom{2n-1}{n-1}$, the sign of $WE^{\{2,1,\dots,1\}} - WE^{\{1,\dots,1\}}$ is identical to that of

$$2\tilde{\phi}^2(0,0) - \tilde{\phi}^2(1,0) - \tilde{\phi}^2(0,1).$$

Finally, using (18) and $(A.19)$, the above-displayed term is identical to

$$\begin{aligned} & p_A \mu_{A|A}^2 (2P_A(1,0) - P_A(2,0) - P_A(1,1)) + p_A \mu_{B|A}^2 (2P_B(1,0) - P_B(2,0) - P_B(1,1)) \\ & + p_B \mu_{A|B}^2 (2P_A(0,1) - P_A(1,1) - P_A(0,2)) + p_B \mu_{B|B}^2 (2P_B(0,1) - P_B(1,1) - P_B(0,2)) \\ = \; & p_A \mu_{A|A}^2 (2P_A(1,0) - P_A(2,0) - P_A(1,1)) - p_A \mu_{B|A}^2 (2P_A(1,0) - P_A(2,0) - P_A(1,1)) \\ & + p_B \mu_{A|B}^2 (2P_A(0,1) - P_A(1,1) - P_A(0,2)) - p_B \mu_{B|B}^2 (2P_A(0,1) - P_A(1,1) - P_A(0,2)) \\ = \; & p_A \left(\mu_{A|A}^2 - \mu_{B|A}^2\right) (2P_A(1,0) - P_A(2,0) - P_A(1,1)) \\ & + p_B \left(\mu_{A|B}^2 - \mu_{B|B}^2\right) (2P_A(0,1) - P_A(1,1) - P_A(0,2)), \end{aligned}$$

where the first step used $P_B(b_A, b_B) = 1 - P_A(b_A, b_B)$. Next, using $P_A(1,0) = p_A P_A(2,0) + p_B P_A(1,1)$ and $P_A(0,1) = p_A P_A(1,1) + p_B P_A(0,2)$, the above is equivalent to

$$\begin{aligned} & p_A \left(\mu_{A|A}^2 - \mu_{B|A}^2\right) (2p_A P_A(2,0) + 2p_B P_A(1,1) - (p_A + p_B) P_A(2,0) - (p_A + p_B) P_A(1,1)) \\ & + p_B \left(\mu_{A|B}^2 - \mu_{B|B}^2\right) (2p_A P_A(1,1) + 2p_B P_A(0,2) - (p_A + p_B) P_A(1,1) - (p_A + p_B) P_A(0,2)), \end{aligned}$$

thus showing that the sign of $WE^{\{2,1,\dots,1\}} - WE^{\{1,\dots,1\}}$ is the same as that of

$$p_A \left(\mu_{A|A}^2 - \mu_{B|A}^2\right) (p_A - p_B) (P_A(2,0) - P_A(1,1)) + p_B \left(\mu_{A|B}^2 - \mu_{B|B}^2\right) (p_A - p_B) (P_A(1,1) - P_A(0,2)). \tag{A.45}$$

Finally, note that

$$P_A(2,0) - P_A(1,1) = \binom{2n-1}{n-1} p_A^{n-1} p_B^n \text{ and } P_A(1,1) - P_A(0,2) = \binom{2n-1}{n} p_A^n p_B^{n-1},$$

hence, as $p_A > p_B$, the sign of $(A.45)$ is the same as that of

$$\mu_{A|A}^2 - \mu_{B|A}^2 + \mu_{A|B}^2 - \mu_{B|B}^2, \tag{A.46}$$

which is positive by $(A.39)$. This concludes the proof the $WE^{\{2,1,\dots,1\}} > WE^{\{1,\dots,1\}}$.

Third, we show that $WE^{\{2n+1\}} > WE^{\{1,2n\}}$. By Corollary 2, we can neglect the very first battle for the comparison. Proceeding as above, we obtain that the contribution of the last $2n$ battles to WE in temporal structure $\{2n+1\}$ is

$$\begin{aligned}
& \sum_{m=2}^{2n+1}\binom{2n}{n}p_A^n p_B^n \tilde{\phi}^m(0,0) \\
= & \binom{2n}{n}p_A^n p_B^n\left(p_A\frac{3c_B-c_A}{2c_B-c_A}P_A(1,0)+p_B P_A(0,1)+p_A\frac{c_A}{2c_B-c_A}P_B(1,0)+2p_B P_B(0,1)\right) \\
& \cdot\sum_{m=2}^{2n+1}\frac{1}{3c_B\gamma_m},
\end{aligned}$$

where we used (17), (18), and $(A.19)$. Similarly, the contribution of the last $2n$ battles to WE in temporal structure $\{1,2n\}$ is

$$\begin{aligned}
& p_A\binom{2n-1}{n-1}p_A^{n-1}p_B^n\left(\begin{array}{c} p_A\frac{3c_B-c_A}{2c_B-c_A}P_A(2,0)+p_B P_A(1,1) \\ +p_A\frac{c_A}{2c_B-c_A}P_B(2,0)+2p_B P_B(1,1)\end{array}\right)\sum_{m=2}^{2n+1}\frac{1}{3c_B\gamma_m} \\
& +p_B\binom{2n-1}{n-1}p_A^n p_B^{n-1}\left(\begin{array}{c} p_A\frac{3c_B-c_A}{2c_B-c_A}P_A(1,1)+p_B P_A(0,2) \\ +p_A\frac{c_A}{2c_B-c_A}P_B(1,1)+2p_B P_B(0,2)\end{array}\right)\sum_{m=2}^{2n+1}\frac{1}{3c_B\gamma_m}.
\end{aligned}$$

As $\binom{2n}{n}=2\binom{2n-1}{n-1}$, we obtain that the sign of $WE^{\{2n+1\}}-WE^{\{1,2n\}}$ is identical to that of

$$\begin{aligned}
& 2\left(p_A\frac{3c_B-c_A}{2c_B-c_A}P_A(1,0)+p_B P_A(0,1)+p_A\frac{c_A}{2c_B-c_A}P_B(1,0)+2p_B P_B(0,1)\right) \\
& -\left(p_A\frac{3c_B-c_A}{2c_B-c_A}P_A(2,0)+p_B P_A(1,1)+p_A\frac{c_A}{2c_B-c_A}P_B(2,0)+2p_B P_B(1,1)\right) \\
& -\left(p_A\frac{3c_B-c_A}{2c_B-c_A}P_A(1,1)+p_B P_A(0,2)+p_A\frac{c_A}{2c_B-c_A}P_B(1,1)+2p_B P_B(0,2)\right),
\end{aligned}$$

or, using $P_B(x,y)=1-P_A(x,y)$ and collecting terms,

$$p_A\frac{3c_B-2c_A}{2c_B-c_A}\left(2P_A(1,0)-P_A(2,0)-P_A(1,1)\right)-p_B\left(2P_A(0,1)-P_A(1,1)-P_A(0,2)\right).$$

Finally, using $P_A(x,y)=p_B P_A(x,y+1)+p_A P_A(x+1,y)$, the above-displayed expression equals

$$\begin{aligned}
& p_A\frac{3c_B-2c_A}{2c_B-c_A}\left(2p_B P_A(1,1)+2p_A P_A(2,0)-P_A(2,0)-P_A(1,1)\right) \\
& -p_B\left(2p_B P_A(0,2)+2p_A P_A(1,1)-P_A(1,1)-P_A(0,2)\right) \\
= & (p_A-p_B)\left(p_A\frac{3c_B-2c_A}{2c_B-c_A}\left(P_A(2,0)-P_A(1,1)\right)-p_B\left(P_A(1,1)-P_A(0,2)\right)\right),
\end{aligned}$$

or, as $p_A > p_B$ and

$$P_A(2,0) - P_A(1,1) = \binom{2n-1}{n-1} p_A^{n-1} p_B^n \text{ and } P_A(1,1) - P_A(0,2) = \binom{2n-1}{n} p_A^n p_B^{n-1}, \quad \text{(A.47)}$$

we obtain that the sign of $WE^{\{2n+1\}} - WE^{\{1,2n\}}$ is the same as that of

$$\frac{3c_B - 2c_A}{2c_B - c_A} - 1 = \frac{c_B - c_A}{2c_B - c_A} > 0.$$

This concludes the proof the $WE^{\{2n+1\}} > WE^{\{1,2n\}}$.

Fourth, we show that $WE^{\{2n,1\}} > WE^{\{2n+1\}}$. By Corollary 2, we can neglect the first $2n$ battle for the comparison. Proceeding as above, we obtain that the contribution of the very last battle to WE in temporal structure $\{2n+1\}$ is

$$\binom{2n}{n} p_A^n p_B^n \left(p_A \mu_{A|A}^{2n+1} P_A(1,0) + p_B \mu_{A|B}^{2n+1} P_A(0,1) + p_A \mu_{B|A}^{2n+1} P_B(1,0) + p_B \mu_{B|B}^{2n+1} P_B(0,1) \right),$$

and in temporal structure $\{2n,1\}$ is

$$\binom{2n}{n} p_A^n p_B^n \left(p_A \mu_{A|A}^{2n+1} + p_B \mu_{B|B}^{2n+1} \right).$$

Therefore, $WE^{\{2n,1\}} - WE^{\{2n+1\}}$ takes the sign of

$$\begin{aligned}
(p_A \mu_{A|A}^{2n+1} + p_B \mu_{B|B}^{2n+1}) &- (p_A \mu_{A|A}^{2n+1} P_A(1,0) + p_B \mu_{A|B}^{2n+1} P_A(0,1) + p_A \mu_{B|A}^{2n+1} P_B(1,0) + p_B \mu_{B|B}^{2n+1} P_B(0,1)) \\
&= (p_A \mu_{A|A}^{2n+1} P_B(1,0) + p_B \mu_{B|B}^{2n+1} P_A(0,1)) - (p_B \mu_{A|B}^{2n+1} P_A(0,1) + p_A \mu_{B|A}^{2n+1} P_B(1,0)) \\
&= p_A (\mu_{A|A}^{2n+1} - \mu_{B|A}^{2n+1}) P_B(1,0) + p_B (\mu_{B|B}^{2n+1} - \mu_{A|B}^{2n+1}) P_A(0,1),
\end{aligned}$$

which is strictly positive by $(A.39)$. This concludes the proof that $WE^{\{2n,1\}} > WE^{\{2n+1\}}$. □

References

[1] **Amann, Erwin, and Wolfgang Leininger**. 1996. "Asymmetric all-pay auctions with incomplete information: the two-player case." *Games and Economic Behavior*, 14(1), 1-18.

[2] **Baik, Kyung Hwan, In-Gyu Kim, and Sunghyun Na**. 2001. "Bidding for a group-specific public-good prize." *Journal of Public Economics,* 82(3): 415-429.

[3] **Barbieri, Stefano, David A. Malueg, and Iryna Topolyan**. 2014. "The best-shot all-pay (group) auction with complete information." *Economic Theory,* 57(3): 603-640.

[4] **Barbieri, Stefano, and Marco Serena**. 2020. "Fair Representation in Primaries: Heterogeneity and the New Hampshire Effect." Working Paper of the Max Planck Institute for Tax Law and Public Finance No. 2020-07. Available at SSRN: https://ssrn.com/abstract=3635360.

[5] **Battaglini, Marco.** 2005. "Sequential voting with abstention." *Games and Economic Behavior*, 51(2): 445-463.

[6] **Baye, Michael R., Dan Kovenock, Casper G. De Vries**. 1996. "The all-pay auction with complete information." *Economic Theory*, 8(2): 291-305.

[7] **Chowdhury, Subhasish M., Dongryul Lee, and Iryna Topolyan**. 2016. "The Max-Min Group Contest: Weakest-link (Group) All-Pay Auction." *Southern Economic Journal*, 83(1): 105-125.

[8] **Decamps, Ambroise, Changxia Ke, and Lionel Page**. 2022. "How success breeds success." *Quantitative Economics*, 13(1), 355-385.

[9] **Dechenaux, Emmanuel, Dan Kovenock, and Roman M. Sheremeta**. 2015. "A survey of experimental research on contests, all-pay auctions and tournaments." *Experimental Economics,* 18(4): 609-669.

[10] **Drugov, Mikhail, and Dmitry Ryvkin**. 2017. "Biased contests for symmetric players." *Games and Economic Behavior,* 103: 116-144.

[11] **Ewerhart, Christian**. 2010. "Rent-Seeking Contests with Independent Private Values." Working Paper No. 490. University of Zurich.

[12] **Ewerhart, Christian**. 2015. "Mixed equilibria in Tullock contests." *Economic Theory*, 60(1), 59-71.

[13] **Ewerhart, Christian**. 2017a. "Revenue ranking of optimally biased contests: The case of two players." *Economics Letters*, 157, 167-170.

[14] **Ewerhart, Christian**. 2017b. "Contests with small noise and the robustness of the all-pay auction." *Games and Economic Behavior*, 105, 195-211.

[15] **Feng, Xin, and Jingfeng Lu.** 2017. "Uniqueness of equilibrium in two-player asymmetric Tullock contests with intermediate discriminatory power." *Economics Letters*, 159, 61-64.

[16] **Feng, Xin, and Lu, Jingfeng.** 2018. "How to Split the Pie: Optimal Rewards in Dynamic Multi-Battle Competitions." *Journal of Public Economics*, 160: 82-95.

[17] **Fey, Mark**. 2008. "Rent-seeking contests with incomplete information." *Public Choice,* 135: 225-236.

[18] **Fu, Qiang, Changxia Ke, and Fangfang Tan.** 2015. ""Success breeds success" or "Pride goes before a fall"?: Teams and individuals in multi-contest tournaments." *Games and Economic Behavior,* 94: 57-79.

[19] **Fu, Qiang, and Jingfeng Lu.** 2020. "On Equilibrium Player Ordering in Dynamic Team Contests." *Economic Inquiry*, doi.org/10.1111/ecin.12907.

[20] **Fu, Qiang, Jingfeng Lu, and Yue Pan**. 2015. "Team Contests with Multiple Pairwise Battles." *American Economic Review*, 105(7): 2120-40.

[21] **Fu, Qiang, and Zenan Wu**. 2019. "Contests: Theory and Topics." the *Oxford Research Encyclopedia of Economics.*

[22] **Häfner, Samuel**. 2017. "A tug-of-war team contest." *Games and Economic Behavior,* 104: 372-391.

[23] **Hamilton, J., and Romano, R. E.** 1998. "Equilibrium assignments in pairwise team contests: How to form political slates and tennis teams." *European Journal of Political Economy*, 14(1), 101-114.

[24] **Harris, Christopher, and John Vickers**. 1987. "Racing with uncertainty." *The Review of Economic Studies*, 54(1): 1-21.

[25] **Hinnosaar, Toomas** . 2024. "Optimal sequential contests." forthcoming, *Theoretical Economics.*

[26] **Hurley, Terrance and Jason Shogren.** 1998. "Asymmetric information contests." *European Journal of Political Economy*, 14: 645-665.

[27] **Jia, Hao.** 2008. "A stochastic derivation of the ratio form of contest success functions." *Public Choice*, 135(3), 125-130.

[28] **Klumpp, Tilman and Mattias K. Polborn**. 2006. "Primaries and the New Hampshire effect." *Journal of Public Economics*, 90(6-7): 1073-1114.

[29] **Knuth, Donald E**. 1997. The Art of Computer Programming: Fundamental Algorithms (volume 1). Addison-Wesley Professional.

[30] **Konishi, Hideo, Pan, Chen-Yu, and Simeonov, Dimitar.** 2022. "Equilibrium player choices in team contests with multiple pairwise battles." *Games and Economic Behavior*, 132, 274-287.

[31] **Konrad, Kai A**. 2009. *Strategy and Dynamics in Contests*. Oxford: Oxford University Press.

[32] **Konrad, Kai A., and Dan Kovenock**. 2009. "Multi-battle contests." *Games and Economic Behavior*, 66: 256-274.

[33] **Kovenock, Dan, and Brian Roberson.** 2012. "Conflicts with multiple battlefields." In M. R. Garfinkel & S. Skaperdas (Eds.), Oxford handbook of the economics of peace and conflict (pp. 503–531). New York: Oxford University Press.

[34] **Malueg, David A., and Andrew J. Yates**. 2004. "Rent seeking with private values." *Public Choice*, 119: 161-178.

[35] **Malueg, David A., and Andrew J. Yates**. 2010. "Testing contest theory: evidence from best-of-three tennis matches." *Review of Economics and Statistics*, 92(3): 689-692.

[36] **Osborne, Martin J.** 2004. *An introduction to game theory*. New York: Oxford University Press.

[37] **Samuels, Stephen M.** 1965. "On the number of successes in independent trials." *The Annals of Mathematical Statistics*, 1272-1278.

[38] **Serena, Marco**. 2017. "Quality Contests." *European Journal of Political Economy*, 46: 15-25.

[39] **Serena, Marco**. 2021. "Harnessing beliefs to optimally disclose contestants' types." *Economic Theory*, 1-30.

[40] **Snyder, James M**. 1989. "Election goals and the allocation of campaign resources." *Econometrica*, 57(3): 637-660.

[41] **Szymanski, Stefan**. 2003. "The economic design of sporting contests." *Journal of Economic Literature*, 41(4): 1137-1187.

[42] **Wang, Zhewei.** 2010. "The optimal accuracy level in asymmetric contests." *The BE Journal of Theoretical Economics*, 10(1), 13.

[43] **Wasser, Cédric**. 2013. "Incomplete information in rent-seeking contests." *Economic Theory*, 53(1): 239-268.