

Winners' Efforts in Team Contests*

Stefano Barbieri[†], Marco Serena[‡]

January 17, 2019

Abstract

The conventional wisdom for *individualistic* dynamic contests is that expected total effort is larger in a simultaneous than in a sequential contest, as only the latter is affected by the “discouragement” effect. In contrast, all temporal structures of *team* contests yield the same expected total effort, as shown by Fu, Lu, and Pan (2015)’s remarkable neutrality result. Rather than expected total effort, we analyze the consequences of different temporal structures of team contests on expected *winners’ efforts*, a natural objective in a number of applications such as R&D races, elections, and sports. We show that, among all possible temporal structures, expected winners’ efforts are maximized by a fully simultaneous and minimized by a fully sequential contest. This result thus parallels the conventional wisdom for individualistic contests. And the neutrality of expected total effort ceases to hold as soon as winners’ and losers’ efforts are not equally valuable.

JEL classification codes: C72, D72, D74, D82

Keywords: team contest, winners’ efforts, temporal structures.

*We are pleased to acknowledge useful comments by Qiang Fu, Tilman Klumpp, Kai Konrad, Jingfeng Lu, Dave Malueg, and participants to the 32nd Tax Day (Max Planck Institute for Tax Law and Public Finance). All errors are our own.

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

[‡]Max Planck Institute for Tax Law and Public Finance. *Email:* marco.serena85@gmail.com

1 Introduction

The analysis of dynamic contests is a staple of economic theory. The key insight therein is the “discouragement” effect: as the competition unfolds, an endogenous advantage arises for the winner of the early battles, and competitiveness and efforts suffer. Indeed, Klumpp and Polborn (2006) show that expected total effort in an individualistic, simultaneous contest is larger than that in a sequential one, as only the latter is affected by the discouragement effect.¹ This result has acquired the force of conventional wisdom in the literature.²

This conventional wisdom has been countered by Fu, Lu, and Pan (2015)—henceforth FLP—for team, rather than individualistic, contests; each player of a team is matched to a player of the rival team in a battle, and each player fights and bears the costs of her battle only.³ Their key result is that all possible temporal structures—i.e., sequential and (partially) simultaneous contests—yield the same expected total efforts, because the discouragement effect does not arise in such team contests.⁴

We focus on the expected winners’ efforts (henceforth WE), rather than on the conventional analysis of expected total effort (henceforth TE). We find that, unlike TE , WE does depend on the contest’s temporal structure. And the neutrality result, which holds for TE , ceases to hold as soon as winners’ and losers’ efforts are not equally valuable; this is natural in many instances, including the three main applications presented by FLP: R&D races, elections, and sports. Consider first R&D races. Efforts exerted on the winning project might be more relevant than that exerted on the losing project: 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 loser because the winners’ policies are ultimately implemented,

¹In individualistic contests the same two players fight against each other in all battles. For other examples of discouragement effect see, for instance, Konrad and Kovenock (2009) and Malueg and Yates (2010). For a review and applications, see Konrad (2009).

²For a review of the large body of the experimental and empirical literature testing and measuring the discouragement effect, see Dechenaux et al. (2015) and Decamps et al. (2018), and references therein.

³The structure of FLP with majoritarian pairwise battles is not the only 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, contests with multiple battles see Kovenock and Roberson (2012) and Fu and Wu (2018).

⁴To see this, note that in individualistic contests the discouragement effect arises because the laggard anticipates having to provide more future effort than the leader in order to win. In contrast, in FLP’s team contest a member of the laggard team is not discouraged because the cost of future efforts will be borne by teammates.

whereas the losers’ are not. As for sports, spectators may be especially thrilled if the winning team also establishes a world record. Or they may look unfavorably at a competition in which the difference between winners’ and losers’ efforts is too large, as in a runaway match.

The neutrality result for TE derived by FLP obtains because a change in the temporal structure shifts efforts from some battles to others. But as the discouragement effect does not apply, these shifts of efforts perfectly balance out when considering TE , which weights equally all players’ efforts. In contrast, if one maximizes WE alone, the efforts of the team more likely to win have a greater weight than those of its competitor. Thus, in contrast to TE , for WE shifts of efforts across battles due to a change in the temporal structure do not perfectly balance out. In particular we show that, among all possible temporal structures, WE is minimized by a fully sequential (Section 5) and maximized by a fully simultaneous (Section 6) contest—thus paralleling the conventional wisdom—unless all players are equally skilled.⁵ Our results hold in several commonly used contest models (Section 7), and they provide a ranking of simultaneous and sequential temporal structures for any linear combination of winners’, losers’, and total effort (Section 2).

2 Model and main results

Two rival teams, A and B , consist of $2n + 1$ players each. Players are matched in $2n + 1$ head-to-head battles. Players of the team that wins at least $n + 1$ battles obtain a prize of 1 each. In a battle where the player of team A (B) exerts effort x_A (x_B), the player of team A wins the battle with probability $p_A(x_A, x_B)$, which is homogeneous of degree zero. The player of team B wins with the complementary probability. (For brevity we write p_i with $i \in \{A, B\}$ for the equilibrium probability of winning.) Marginal costs of effort are c_A (c_B) for members of team A (B).

As shown by FLP,

Fact 1. p_i is: 1) constant across battles, 2) independent of the temporal structure, and 3) independent of the common prize of victory in a battle.

The valuation of winning a battle for a player is the resulting increase in her team’s winning probability; denoting this battle prize by v , the expected equilibrium contribution x_i is linear in v ,

$$x_i = \mu_i(c_A, c_B) \cdot v, \quad i \in \{A, B\}, \quad (1)$$

⁵If players are all equally skilled, efforts and thus probabilities of victory are identical across teams. Therefore, WE boils down to a rescaling of TE and FLP’s neutrality result carries over.

where the details of the function μ_i depend on the function $p_i(x_A, x_B)$. (For brevity we write μ_i with $i \in \{A, B\}$ for the equilibrium $\mu_i(c_A, c_B)$.) The linearity (1) applies widely in contest models (see Model 1-7, Section I.C of FLP) and suffices to generate FLP’s result of neutrality of TE to changes in the temporal structure.

However, WE is a more complex objective than TE ; in WE , unlike in TE , each player’s effort is weighted by its team’s probability of victory which in turn depends on the outcome of past matches.⁶ Due to this extra complexity, we require two additional properties, on top of linearity (1), for our result on WE :

$$p_A \gtrless p_B \Leftrightarrow c_A \gtrless c_B, \tag{2}$$

$$\mu_A \gtrless \mu_B \Leftrightarrow c_A \gtrless c_B. \tag{3}$$

In Section 7 we show that (2) and (3) hold in three commonly adopted setups; namely, generalized Tullock contests under complete information, and all-pay auctions under complete or incomplete information.⁷ Furthermore, properties (2) and (3) are intuitive; each player of the stronger team—i.e., with lower marginal cost of effort—exerts larger (expected) effort and wins her battle with larger probability.

As in FLP, we consider all possible temporal structures; i.e., battles can be carried out sequentially or (partially) simultaneously. We define the temporal structure of a contest as a sequence of strictly positive natural numbers $\{n_1, n_2, \dots\}$. The first n_1 players simultaneously choose their efforts, then battles’ outcomes realize and the following n_2 players simultaneously choose their efforts, and so on. Thus, $n_1 + n_2 + \dots = 2n + 1$. The fully sequential contest is $n_1 = n_2 = \dots = 1$. The fully simultaneous contest is $n_1 = 2n + 1$.

Main Result. Despite all contest’s temporal structures yielding the same TE , as FLP show, we prove that WE is minimized by a fully sequential contest (Proposition 1) and maximized by a fully simultaneous contest (Proposition 2). In order to derive such results, we cannot rely on shifts of efforts cancelling out, as for TE . Hence, we provide an original proof which relies on a novel technique—namely, “extractions” and “mergers.”

First, we define an extraction as a temporal structure where, from the first non-unitary n_j of the starting temporal structure, one battle is extracted and fought on its own right before; that is, for instance, in $\{1, 4, 1, 3\}$ the first non-unitary n_j is $n_2 = 4$, and thus the extraction yields $\{1, 1, 3, 1, 3\}$. In Proposition 1 we prove that such an extraction decreases WE , and

⁶See Serena (2017) for the main driving forces in two-player contests when maximizing WE or TE .

⁷While under complete information c_i with $i \in \{A, B\}$ is the marginal costs of effort, under incomplete information costs are stochastic and, if the realization of cost is c' , then the marginal cost is $c'c_i$ for members of team i . Hence, in both complete and incomplete information settings, all players are ex-ante identical except for a team-specific parameter c_i for members of team i that characterizes marginal costs of effort.

the decrease is strict unless $c_A = c_B$ or we are comparing $\{1, 1, \dots, 1, 2\}$ with $\{1, 1, \dots, 1\}$. Starting from any arbitrary contest temporal structure, a sufficient number of extractions yields the fully sequential contest. Hence, Proposition 1 can be used to show that WE is minimized, among all possible temporal structures, by a fully sequential contest.

Second, we define a merger as a temporal structure where the first two groups of simultaneous battles, n_1 and n_2 , are merged into $n_1 + n_2$ simultaneous battles; that is, for instance, starting from $\{1, 4, 1, 3\}$ the merger yields $\{5, 1, 3\}$. In Proposition 2 we prove that such a merger increases WE , and the increase is strict unless $c_A = c_B$. Starting from any arbitrary contest temporal structure, a sufficient number of mergers yields the fully simultaneous contest. Hence, Proposition 2 can be used to show that WE is maximized, among all possible temporal structures, by a fully simultaneous contest.

Remark. Propositions 1 and 2 can be immediately extended beyond WE , to a variety of objectives that are linear combinations of TE , WE , and the losers' effort $LE = TE - WE$. More formally, let the maximand be $\alpha \cdot WE + \beta \cdot LE + \gamma \cdot TE$, where α , β , and γ are constants. Since $LE = TE - WE$, the objective can be equivalently written as $(\alpha - \beta) \cdot WE + (\gamma + \beta) \cdot TE$, and TE is unaffected by the temporal structure of the contest, so the objective simplifies to $(\alpha - \beta) \cdot WE$. Then, the optimality of a fully simultaneous or a fully sequential temporal structure, among all possible temporal structures, only depends on whether $\alpha \geq \beta$. Thus, neutrality ceases to hold as soon as $\alpha \neq \beta$ and what holds for the maximization of WE alone—namely, a simultaneous temporal structure dominates any other—holds more generally when WE is weighted more than LE . Conversely, a sequential structure dominates any other when LE is weighted more than WE .⁸ For the sake of simplicity, we focus on the maximization of WE alone throughout the remaining of the paper.

3 Intuition; extractions and mergers

Extractions and mergers are useful for understanding the intuition behind FLP's result for maximization of TE and our result for maximization of WE . In this section, we simplify the setup by setting $n = 1$ (i.e., a best-of-three contest), $c_A = 1$ and $c_B = 2$, and considering the workhorse model in the literature (a Tullock contest⁹ under complete information). This simple setup helps us to illustrate extractions and mergers, and to spell out the main forces driving the difference between FLP's result for TE and our results for WE .

⁸Minimization of WE is also covered—e.g., $\alpha = -1$ and $\beta = \gamma = 0$.

⁹In a Tullock contest the player of team A wins with probability $p_i(x_A, x_B) = x_i / (x_A + x_B)$.

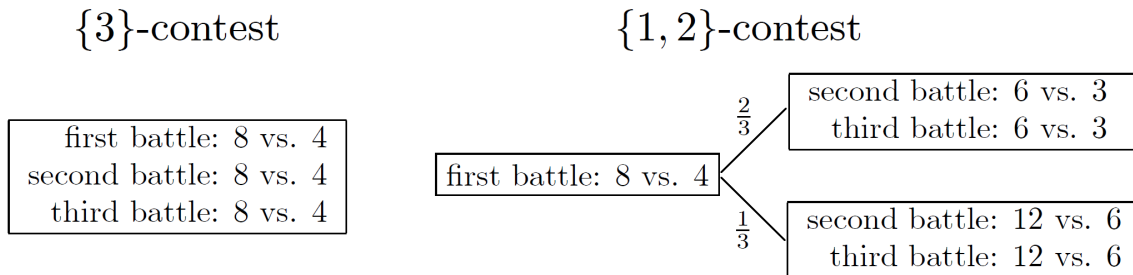


Figure 1: Efforts (multiplied by 81) in the $\{3\}$ - and $\{1, 2\}$ -contests with $n = 1$, $c_A = 1$, $c_B = 2$. The calculations follow by standard techniques. WE , multiplied by 2187, is 564 for the $\{3\}$ -contest and 548 for the $\{1, 2\}$ -contest. One can also calculate that WE is 556 in a $\{2, 1\}$ -contest, and 548 in a $\{1, 1, 1\}$ -contest.

Figure 1 depicts efforts when comparing temporal structures $\{3\}$ and $\{1, 2\}$ for a best-of-three contest. This comparison captures both an extraction from $\{3\}$ to $\{1, 2\}$ and a merger from $\{1, 2\}$ to $\{3\}$. (In general extractions and mergers need not be each other’s inverse.) In what follows, we explain why TE is identical in the two temporal structures while WE is larger in $\{3\}$, the fully simultaneous contest. First, this comparison shows that FLP’s neutrality does not extend to WE . Second, this comparison can be extended to a further extraction from $\{1, 2\}$ to $\{1, 1, 1\}$, hence intuitively suggesting that WE is larger in a simultaneous than a sequential contest.

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. In other words, future arrangements of battles do not distort the efforts of past battles. To see this, we start from FLP’s result that the outcomes of past battles do not distort the probability of victory for each team in future battles (effort levels change in each battle, but they do so proportionally across teams); that is, the discouragement effect does not apply. In particular, since $c_B = 2c_A$, the Tullock contest success function implies $p_A = 2/3$ 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 her team’s win probability if she wins that battle and that if she loses it. This difference is $4/9$ both for the A-player ($8/9 - 4/9$) and the B-player ($5/9 - 1/9$).¹⁰ In the first battle of the $\{3\}$ -contest, each player fights for a prize equal to the probability that her battle is pivotal, which also equals $4/9$ for both players.¹¹ Thus, in the first battle of the $\{3\}$ - and $\{1, 2\}$ -contests, players

¹⁰If 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.

¹¹The first battle of the $\{3\}$ -contest is pivotal only if the second and third battles are won by different

are fighting for the same prize, and hence they exert identical effort.

A second key feature that is clear from Figure 1 is that 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. In what follows, we show that such comparison of the efforts of the third battle yields: 1) identical TE in the $\{3\}$ - and $\{1, 2\}$ -contests, as FLP's neutrality result shows, and 2) larger WE in the $\{3\}$ -contest than in the $\{1, 2\}$ -contest, as our propositions show.

1) 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 $2/3$).¹² 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 , since $(2/3) \cdot 6 + (1/3) \cdot 12 = 8$. 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 neutrality result that FLP show under great generality; all contest's temporal structures yield the same TE .

2) Consider now WE . Focus first on the A-player. The low effort 6 is exerted *conditional* on A having won the first battle, and its weight in WE is relatively large because team A 's winning odds increase when team A is one battle ahead of team B : indeed, the weight is $8/9$ (see fn. 10). Symmetrically, the high effort 12 is exerted *conditional* on B having won the first battle, and its weight in WE is relatively low because team A 's winning odds decrease when team A is one battle behind team B : indeed, the weight is $4/9$ (see fn. 10). In a nutshell: greater weight to lower effort and lower weight to greater effort make WE *decrease*. Focus now on the B-player. The argument for WE reverses: greater (lower) weight to greater (lower) effort makes WE *increase*. However, the decrease in WE for the A-player overcomes the increase in WE for the B-player because A-players' efforts and probability of victory are larger than those of B-players. Therefore, WE in the $\{3\}$ -contest is larger than in the $\{1, 2\}$ -contest. This stylized reasoning can be extended to show that WE is maximized by a fully simultaneous contest and minimized by a fully sequential contest, despite TE being invariant to the temporal structure as FLP show.

teams, which happens with probability $2(1/3)(2/3) = 4/9$.

¹²The third battle's pivotal probability is the probability that the second battle is won by the team who lost the first.

4 Preliminaries

In this section we state and prove the intermediate building blocks for propositions 1 and 2; namely, extractions decrease WE , and mergers increase WE . In doing so, we show that all the qualitative features of the example in Section 3 hold for any triple (c_A, c_B, n) , any extractions or mergers, and any contest success function satisfying the extra properties (2) and (3). All proofs are in the Appendix.

For a given temporal structure $T \equiv \{n_1, n_2, \dots\}$, we characterize efforts in a general set of battles $n_j \geq 1, j \in \{1, 2, \dots\}$.

Lemma 1. *Let (b_A, b_B) identify the number of early battles (fought before the set of battles n_j) won by each team. If $(b_A, b_B) \in \{0, n\}^2$, then individual efforts in the set of battles n_j obey*

$$x_i(b_A, b_B) = \mu_i \cdot \binom{2n - (b_A + b_B)}{n - b_A} p_A^{n-b_A} p_B^{n-b_B}; \quad i \in \{A, B\} \quad (4)$$

otherwise, the contest is already decided in the early battles and efforts in the set of battles n_j and afterwards are zero.

Next, we introduce some notation on the probability of victory. Let (β_A, β_B) identify the number of battles already won by each team, with $(\beta_A, \beta_B) \in \{0, \dots, n\}^2$. Since the probability that team A (B) wins each battle is p_A (p_B), then the probability of victory of team A given β_A victories for team A and β_B victories for team B is

$$P_A(\beta_A, \beta_B) = \sum_{i=n+1-\beta_A}^{2n+1-(\beta_A+\beta_B)} \binom{2n+1-(\beta_A+\beta_B)}{i} p_A^i p_B^{2n+1-(\beta_A+\beta_B)-i}, \quad (5)$$

and $P_B(\beta_A, \beta_B)$ is the complementary probability. Define as $\phi(\beta_A, \beta_B)$ the efforts exerted at (β_A, β_B) by the player of the expected winning team, conditional on reaching (β_A, β_B) . We have

$$\phi(\beta_A, \beta_B) = P_A(\beta_A, \beta_B) x_A(\beta_A, \beta_B) + P_B(\beta_A, \beta_B) x_B(\beta_A, \beta_B). \quad (6)$$

One approach is to use Lemma 1 to derive an explicit formula for WE for every possible temporal structure T , and derive comparative statics in T of such an explicit formula. This would provide a complete ranking of temporal structures, including the fully sequential and fully simultaneous ones. However, this approach turns out to involve expressions that are not tractable for WE , as opposed to TE in which shifts of efforts cancel out when comparing any two temporal structures. Hence, we take a different tack, based on a novel technique we develop—namely, extractions and mergers. Section 5 shows that extractions decrease WE , and Section 6 shows that mergers increase WE .

5 Extractions decrease WE

We first provide the formal definition of an extraction, which we informally defined in Section 2.

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

In order to show that extractions decrease WE , we derive a number of intermediate steps. Throughout this section, we keep as a running example the case of $T = \{1, 3, 5\}$, thus $T^{ext} = \{1, 1, 2, 5\}$, $n_l = 3$, $n_{l-1} = 1$, and $l = 2$.

First, by Fact 1, extractions do not affect efforts of earlier battles $\{n_1, \dots, n_{l-1}\}$ and later battles $\{n_{l+1}, \dots\}$. Consequently, these battles can be disregarded when comparing WE in T and T^{ext} . In our running example, 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^{mer} by Fact 1.

Second, at the one battle immediately after n_{l-1} (i.e., at the $(n_{l-1} + 1)^{th}$ battle), Lemma 1 implies identical effort across players in T and T^{ext} , since early battles $\{n_1, \dots, n_{l-1}\}$ are not affected by the extraction. In our running example, the efforts of the second battle are identical across players in T and T^{ext} .

We store these two results as (i) and (ii) of the following corollary.

Corollary 1. (Efforts and WE after an extraction). When comparing WE in temporal structures T and T^{ext} , two simplifying facts arise:

- (i) Battles $\{n_1, \dots, n_{l-1}\}$ and $\{n_{l+1}, \dots\}$ can be disregarded.
- (ii) The $(n_{l-1} + 1)^{th}$ battle has the same per-player effort in T and T^{ext} . However, in T^{ext} this battle is separately fought, while in T it is fought simultaneously with other $n_l - 1$ battles; hence, in T we can consider only $n_l - 1$ battles, while in T^{ext} we disregard the $(n_{l-1} + 1)^{th}$ battle.

So, WE in T is larger (equal) [smaller] than in T^{ext} if and only if $WE^T > (=) [<] WE^{T^{ext}}$, where

$$WE^{T^{ext}} \equiv (n_l - 1) \sum_{b_A = \max\{l - n, 0\}}^{\min\{n, l\}} \binom{l}{b_A} p_A^{b_A} p_B^{l - b_A} \phi(b_A, l - b_A), \quad (7)$$

and

$$WE^T \equiv (n_l - 1) \sum_{b_A = \max\{l-1-n, 0\}}^{\min\{n, l-1\}} \binom{l-1}{b_A} p_A^{b_A} p_B^{l-1-b_A} \phi(b_A, l-1-b_A). \quad (8)$$

We explain formula (7). Exploiting (i) and (ii) in the corollary, in $WE^{T^{ext}}$ we ignore the first l battles $\{n_1, \dots, n_{l-1}, 1\}$, and battles $\{n_{l+1}, \dots\}$. Hence, $WE^{T^{ext}}$ accounts only for the $n_l - 1$ battles left. The first l ignored battles result in b_A victories for team A and $l - b_A$ victories for team B, and each $(b_A, l - b_A)$ occurs with probability $\binom{l}{b_A} p_A^{b_A} p_B^{l-b_A}$. Starting from $(b_A, l - b_A)$, $\phi(b_A, l - b_A)$ is the expected winning effort of the $n_l - 1$ battles left to be considered in $WE^{T^{ext}}$ after the simplifications arising from (i) and (ii).

Formula (8) is similar, but one less battle is fought sequentially. Note, however, that we keep the term $(n_l - 1)$ as in (7), rather than substituting it with n_l , because we count one less player as discussed in (ii) of the corollary.

In our running example with $T = \{1, 3, 5\}$ and $T^{ext} = \{1, 1, 2, 5\}$, from (i) and (ii) of the corollary we can focus on the third and fourth battles only, as it can be seen in formulae (7) and (8). These are the only two battles that can have different efforts in T and T^{ext} , as they are fought simultaneously with the second battle in T and after the second battle in T^{ext} .

Corollary 1 is a key step to understand the intuition why WE decreases with extractions. In Figure 1 we ignored the first battle both in TE and WE since efforts and team probabilities of winning are identical in the $\{3\}$ - and $\{1, 2\}$ -contests; this analytical shortcut is generalized by (ii) of Corollary 1. Additionally, (i) of Corollary 1 allows us to ignore battles earlier and later than n_l when comparing WE in T and T^{ext} . For instance, when comparing WE in $\{1, 5, 3\}$ - and $\{1, 1, 4, 3\}$ -contest, we can ignore the first battle by (i), and the seventh, eighth, and ninth battles by (ii).

Recall that moving from the analysis of TE to that of WE entails a tractability loss; each player's effort is weighted by its team's probability of victory and, as a result, shifts of efforts from some battles to others do not perfectly cancel out as in TE . But Corollary 1 makes up for such a tractability loss: using the simplifications of Corollary 1, the effect of an extraction on WE is determined by the behavior of the function

$$\psi(z) = \sum_{b_A = \max\{z-n, 0\}}^{\min\{n, z\}} \binom{z}{b_A} \binom{2n-z}{n-b_A} \cdot P_A(b_A, z-b_A), \quad (9)$$

which we characterize in the following two lemmas.

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

$$(\mu_A - \mu_B) \psi(l-1) > (=) [<] (\mu_A - \mu_B) \psi(l). \quad (10)$$

Lemma 3. *Suppose $p_A \in (0, 1)$. If $l = 2n$, then $\psi(l-1) = \psi(l)$. Otherwise, if $p_A \geq p_B$, then $\psi(l-1) \geq \psi(l)$.*

Lemmas 2 and 3, jointly, provide conditions under which an extraction increases or decreases WE . In particular,

1. if $l = 2n$ (i.e., $T = \{1, 1, \dots, 1, 2\}$ and $T^{ext} = \{1, 1, \dots, 1\}$), then by Lemma 3 $\psi(l-1) = \psi(l)$, and hence Lemma 2 shows that WE is constant after the extraction.
2. if $c_A = c_B$, which by (3) is equivalent to $\mu_A = \mu_B$, then Lemma 2 implies that WE is constant after the extraction.
3. if, say, $c_A < c_B$, which by (3) is equivalent to $\mu_A > \mu_B$, then Lemma 2 shows that WE increases 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 (2). The case of $c_A > c_B$ is symmetric.

From the above discussion, the main result of this section follows; extractions decrease WE . Starting from any arbitrary temporal structure, a sufficient number of extractions yields the fully sequential contest, and thus the fact that extractions decrease WE implies that WE is minimized by a fully sequential contest.

Proposition 1. *Consider a team contest à la FLP, in which the battles satisfy (2) and (3). An extraction decreases WE . The decrease is strict unless $c_A = c_B$ or $l = 2n$.*

Therefore, WE is minimized, among all possible temporal structures, by a fully sequential contest.

6 Mergers increase WE

We first provide the formal definition of a merger, which we informally defined in Section 2.

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

In order to show that extractions decrease WE , we derive a number of intermediate steps. Throughout this section, we keep as a running example the case of $T = \{3, 1, 5\}$, thus $T^{mer} = \{4, 5\}$, $n_1 = 3$, and $n_2 = 1$.

First, by Fact 1, mergers do not affect efforts of later battles $\{n_3, \dots\}$. Consequently, these battles can be disregarded when comparing WE in T and T^{mer} . In our running example, 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} by Fact 1.

Second, at the first n_1 battles, Lemma 1 implies identical effort across players in T and T^{mer} . In our running example, 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 these two results as (i) and (ii) of the following corollary.

Corollary 2. (*Efforts and WE after a merger*). When comparing WE in temporal structures T and T^{mer} , two simplifying facts arise:

(i) Battles after the first $n_1 + n_2$ can be disregarded, if any exist.

(ii) The first n_1 battles have the same per-player effort in T and T^{mer} , and hence they can be disregarded.

So, WE in T^{mer} is larger (equal) [smaller] than in T if and only if $WE^{T^{mer}} > (=) [<] WE^T$, where¹³

$$WE^T \equiv n_2 \sum_{b_A=\max\{n_1-n,0\}}^{\min\{n_1,n\}} \binom{n_1}{b_A} p_A^{b_A} p_B^{n_1-b_A} \phi(b_A, n_1 - b_A), \quad (11)$$

and

$$WE^{T^{mer}} \equiv n_2 \phi(0, 0). \quad (12)$$

We explain formula (11). Exploiting (i) and (ii) in the corollary, WE^T accounts only for the n_2 battles following the first n_1 battles; that is, from the $(n_1 + 1)^{th}$ to the $(n_1 + n_2)^{th}$

¹³The expected winners' efforts under the arbitrary temporal structure T is the same in Section 5 and Section 6. However, the formulae in (11) and (8) differ as the simplifications when comparing T with T^{ext} and T^{mer} differ.

battles. The first n_1 battles result in b_A victories for team A and $n_1 - b_A$ victories for team B, and each $(b_A, n_1 - b_A)$ occurs with probability $\binom{n_1}{b_A} p_A^{b_A} p_B^{n_1 - b_A}$. Starting from $(b_A, n_1 - b_A)$, $\phi(b_A, n_1 - b_A)$ is the expected winning effort of the n_2 battles left to be considered in WE^T after the simplifications arising from (i) and (ii). Formula (12) accounts only for n_2 battles out of the total $n_1 + n_2$ initial battles and it is simpler than (11) as, after a merger, the n_2 battles are fought simultaneously with the first n_1 battles, hence there are no possible past victories to consider.

In our running example with $T = \{3, 1, 5\}$ and $T^{mer} = \{4, 5\}$, from (i) and (ii) in the corollary we can focus on the fourth battle only, as it can be seen in (11). The fourth battle is the only battle that can have different efforts in T and T^{mer} , as it is fought simultaneously with the first three battles in T^{mer} and after the first three battles in T .

As in Section 5, using the simplifications of Corollary 2, the effect of a merger on WE is determined by the behavior of ψ defined in (9). The following lemma mirrors Lemma 2.

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

$$(\mu_A - \mu_B) \psi(0) > (=) [<] (\mu_A - \mu_B) \psi(n_1).$$

Lemmas 3 and 4, jointly, provide conditions under which a merger increases or decreases WE . In particular,

1. if $c_A = c_B$, which by (3) is equivalent to $\mu_A = \mu_B$, then Lemma 4 implies that WE is constant after the merger.
2. if, say, $c_A < c_B$, which by (3) is equivalent to $\mu_A > \mu_B$, 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 (2). The case of $c_A > c_B$ is symmetric.

From the above discussion, the main result of this section follows; mergers increase WE . Starting from any arbitrary temporal structure, a sufficient number of mergers yields the fully simultaneous contest, and thus the fact that mergers increase WE implies that WE is maximized by a fully simultaneous contest.

Proposition 2. *Consider a team contest à la FLP, in which the battles satisfy (2) and (3). A merger increases WE . The increase is strict unless $c_A = c_B$.*

Therefore, WE is maximized, among all possible temporal structures, by a fully simultaneous contest.

7 Contests fulfilling properties (1), (2) and (3)

In this section we show that properties (1), (2) and (3) hold for three commonly adopted setups.

Generalized Tullock contests under complete information. Consider the contest success function

$$p_i(x_A, x_B) \equiv \frac{h_i(x_i)}{h_A(x_A) + h_B(x_B)},$$

where h_1 and h_2 are strictly increasing, strictly positive functions, and $p_i(x_A, x_B)$ is homogeneous of degree zero. Malueg and Yates (2005) consider a contest model in which the contest success function is homogeneous of degree zero and satisfies some boundary conditions that guarantee existence of a pure strategy equilibrium with non-negative efforts (see, e.g., their Proposition 2). Linearity (1) is provided in their Proposition 1. In particular, they define the winning probability of A as $p(x_A, x_B)$, and that of B as $1 - p(x_A, x_B)$. Matching their notation to ours, $p_A = p(1, c_A/c_B)$, which satisfies our property (2). Furthermore,

$$\mu_A = \frac{1}{c_A} p' \left(1, \frac{c_A}{c_B} \right),$$

where p' is the derivative of the contest success function with respect to its first argument. A similar expression holds for B . Hence, their setup satisfies also our property (3). In a closely related and complementary work, Baik (2004) analyzes $p_i(x_A, x_B) \equiv f(x_A/x_B)$, thus yielding in equilibrium $p_A = f(c_B/c_A)$, $\mu_A = f'(c_B/c_A)/c_B$ and $\mu_B = f'(c_B/c_A)/c_A$. Hence, his setup also satisfies properties (1), (2) and (3). Therefore, our propositions 1 and 2 apply, for instance, to the commonly adopted ratio-form $p_i(x_A, x_B) \equiv x_i^r / (x_A^r + x_B^r)$.

All-pay auctions under complete information. From Hillman and Reily (1989) and Baye, Kovenock and de Vries (1996), if payoffs are $v/c_A > v/c_B$, then the A-player bids uniformly on $[0, v/c_B]$ and the B-player bids zero with probability $1 - c_A/c_B$, and bids uniformly on $[0, v/c_B]$ with probability c_A/c_B . So, the probabilities of winning are $p_A = 1 - c_A/(2c_B)$, $p_B = c_A/(2c_B)$, and the expected bids are $x_A = v/(2c_B)$, $x_B = vc_A/(2c_B^2)$. Thus, $\mu_A = 1/(2c_B)$ and $\mu_B = c_A/(2c_B^2)$. Hence, this setup also satisfies properties (1), (2) and (3). Therefore, our propositions 1 and 2 apply.

All-pay auctions under incomplete information. Recall that the common prize of battles is v . Each player privately observes an independent draw c from a common distribution F on $[1, +\infty)$, with bounded density f , and her resulting marginal cost of effort is $c \cdot c_i$, with $i \in \{A, B\}$. As before, c_A and c_B are common knowledge. Therefore, her utility

of contributing x_i is

$$v \cdot \Pr(\text{win}|x_i) - c \cdot c_i x_i.$$

Following Amann and Leininger (1996), and letting $k(c) = x_B^{-1}(x_A(c))$, the differential equation characterizing the equilibrium is

$$f(k(c)) k'(c) k(c) = \frac{c_A}{c_B} c \cdot f(c), \quad (13)$$

with initial condition $k(1) = 1$. The differential equations characterizing x_A and x_B are

$$x'_A(c) = -\frac{v \cdot f(c)}{c_B \cdot k(c)} \text{ and } x'_B(c) = -\frac{v \cdot f(c)}{c_A \cdot k^{-1}(c)}. \quad (14)$$

We next show that this setup satisfies properties (1), (2) and (3).

Making explicit the dependence of the solution on c_A/c_B , we write $k(c; c_A/c_B)$. We show that $k(c; c_A/c_B)$ is increasing pointwise in c_A/c_B .

Lemma 5. *Consider two values of the ratio $\left(\frac{c_A}{c_B}\right)_H > \left(\frac{c_A}{c_B}\right)_L$. For any $c > 1$,*

$$k\left(c; \left(\frac{c_A}{c_B}\right)_H\right) > k\left(c; \left(\frac{c_A}{c_B}\right)_L\right).$$

If $c_A = c_B$, then the differential equation for $k(c; 1)$ is solved by $k(c; 1) = c$. Therefore, by Lemma 5, if $c_A > c_B$, then $k(c) > c$. This, together with $x_B(k(c)) = x_A(c)$ and x_i decreasing, implies $x_A(c) < x_B(c)$. Symmetrically, $c_A < c_B \implies k(c) < c \implies x_A(c) > x_B(c)$. In the remainder of this section we show that this setup satisfies (1), (2) and (3).

First, we prove (1). As $c \rightarrow +\infty$, both $x_A(c)$ and $x_B(c)$ must converge to zero, otherwise payoffs would be negative. Then, (14) yields that the expected effort for the A-player is

$$\begin{aligned} E(x_A) &= \frac{v}{c_B} \int_1^\infty \left(\int_c^\infty \frac{f(z)}{k\left(z; \frac{c_A}{c_B}\right)} dz \right) f(c) dc \\ &= \frac{v}{c_B} \int_1^\infty \int_1^z \frac{f(z)}{k\left(z; \frac{c_A}{c_B}\right)} f(c) dc dz \\ &= \frac{v}{c_B} \int_1^\infty \frac{f(z) F(z)}{k\left(z; \frac{c_A}{c_B}\right)} dz \\ &= \frac{v}{2c_B} \int_1^\infty \frac{1}{k\left(z; \frac{c_A}{c_B}\right)} d[F^2(z)]. \end{aligned}$$

Therefore, $E(x_A)$ is linear in v . One can similarly establish that $E(x_B)$ is also linear in v . Furthermore,

$$\mu_A = \frac{1}{2c_B} \int_1^\infty \frac{1}{k\left(c; \frac{c_A}{c_B}\right)} d[F^2(c)] \quad \text{and} \quad \mu_B = \frac{1}{2c_A} \int_1^\infty \frac{1}{k^{-1}\left(c; \frac{c_A}{c_B}\right)} d[F^2(c)].$$

Hence, this setup satisfies (1).

Second, we prove (2). When $c_A = c_B$, (13) solves as $k(c) = c$, so

$$p_A = \int_1^{+\infty} (1 - F(k(c))) f(c) dc = \int_1^{+\infty} (1 - F(c)) f(c) dc = \frac{1}{2} \int_1^{+\infty} 2(1 - F(c)) f(c) dc = \frac{1}{2},$$

as the argument under the integral is the pdf of the minimum of two i.i.d. random variables.

When $c_A > c_B$, Lemma 5 implies $k(c) > c$, so

$$p_A = \int_1^{+\infty} (1 - F(k(c))) f(c) dc < \int_1^{+\infty} (1 - F(c)) f(c) dc = \frac{1}{2}.$$

Symmetrically, the above displayed inequality reverses for $c_A < c_B$. Hence, this setup satisfies (2).

Third, we prove (3). If $c_A = c_B$, then $x_B(c) = x_A(c)$, so

$$v \cdot \mu_A = \int_1^\infty x_A(c) dF(c) = \int_1^\infty x_B(c) dF(c) = v \cdot \mu_B,$$

hence $\mu_A = \mu_B$.

Consider now $c_A > c_B$. Lemma 5 implies $k(c) > c$ and $x_B(c) > x_A(c)$, so

$$v \cdot \mu_A = \int_1^\infty x_A(c) dF(c) < \int_1^\infty x_B(c) dF(c) = v \cdot \mu_B,$$

hence $\mu_A < \mu_B$. Symmetrically, if $c_A < c_B$, then $\mu_A > \mu_B$. Hence, this setup satisfies (3).

8 Conclusions

FLP innovate the classic individualistic contests (where the same two players fight against each other in all battles) by assuming that each battle is fought by a different player who bears only her battle’s cost: a team contest. Besides applying widely, this novel setup has a number of 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, page 2134): sequential and (partially) simultaneous contests yield the same TE . As illustrated in the stylized example of Figure 1, a “more sequential” contest generates shifts of future-battle efforts that keep the average invariant (namely, $(2/3) \cdot 6 + (1/3) \cdot 12 = 8$), and since TE depends only on the average effort, it is unaffected by such shifts.

But there are other quantities of interest besides TE . An especially important one is WE , which, in contrast to TE , does not depend only on the average effort, and hence is affected by the above shifts of efforts. Consequently, FLP’s neutrality result for TE ceases to hold for WE and, more generally, as soon as the winners’ and losers’ efforts are not equally valuable, which is natural in several applications discussed in the Introduction. We further FLP’s analysis and show that their model—with two additional properties commonly satisfied in the literature— can be used to establish that a simultaneous contest is the best temporal structure and a sequential contest is the worst temporal structure when the efforts of the winners are more valuable than those of the losers, thus paralleling in FLP’s team contest the conventional wisdom valid for TE in individualistic contests. In doing so, 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 . On the one hand, our results underscore the need for a cautious application-driven specification of the objective. On the other hand, our results reopen design questions, such as the effect of interim evaluations, that the focus on TE and FLP’s neutrality result had closed.¹⁴ As we showed, these questions can be explored using FLP’s elegant setup and our novel extraction-and-merger technique.

¹⁴The contest’s temporal structure—i.e., whether battles are carried out sequentially or (partially) simultaneously—in the model of FLP can be alternatively interpreted as the contest’s feedback policy (e.g., interim performance evaluations), as they discuss in Section III.A.

Appendix: proofs of lemmas 1-5

Proof of Lemma 1. We only focus on contests that are not already decided — that is, $0 \leq b_i \leq n$ with $i \in \{A, B\}$. At each of the battles in the set n_j , the common valuation of victory for the two players is the probability that that battle is pivotal, i.e., that all other battles result in a tie n to n , which is also the difference between the win probability of a player's team if that player wins her battle, and that if she loses it. So, if the result of the early battles is (b_A, b_B) , then the probability that the battle is pivotal is

$$\binom{2n - (b_A + b_B)}{n - b_A} p_A^{n-b_A} p_B^{n-b_B}.$$

Hence, this is the common valuation of victory at that battle, so (4) follows from (1). \square

Proof of Lemma 2. Using (4) and (6) into (8), we obtain that $WE^T / (n_l - 1)$ equals

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

where ψ is defined in (9) and $\xi(z)$ as follows:

$$\xi(z) \equiv \sum_{b_A=\max\{z-n,0\}}^{\min\{n,z\}} \binom{z}{b_A} \binom{2n-z}{n-b_A}. \quad (15)$$

We now apply Vandermonde's convolution to show that $\xi(z) = \binom{2n}{n}$. Indeed, if $z \leq n$, we

have

$$\xi(z) = \sum_{b_A=0}^z \binom{z}{b_A} \binom{2n-z}{n-b_A} = \binom{2n}{n},$$

while if $z > n$

$$\xi(z) = \sum_{b_A=z-n}^n \binom{z}{b_A} \binom{2n-z}{n-b_A} = \sum_{j=0}^{2n-z} \binom{z}{n-j} \binom{2n-z}{j} = \binom{2n}{n},$$

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

$$\frac{WE^T}{(n_l - 1)p_A^n p_B^n} = \mu_B \binom{2n}{n} + [\mu_A - \mu_B] \psi(l-1). \quad (16)$$

Applying similar steps to (7), we have

$$\frac{WE^{T^{ext}}}{(n_l - 1)p_A^n p_B^n} = \mu_B \binom{2n}{n} + [\mu_A - \mu_B] \psi(l). \quad (17)$$

Comparing (16) and (17) establishes the Lemma. \square

Proof of Lemma 3. Throughout the proof we let $\Delta(z) \equiv \psi(z-1) - \psi(z)$. We show that $p_A \geq p_B \Rightarrow \Delta(z) \geq 0$ 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 recursion applied to (9) gives

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

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

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

Therefore, we have

$$\begin{aligned}
\Delta(z) &= - \sum_{b_A=1}^z (P_A(b_A, z-b_A) - P_A(b_A-1, z-b_A)) \binom{z-1}{b_A-1} \binom{2n-z}{n-b_A} \\
&\quad + \sum_{b_A=0}^{z-1} (P_A(b_A, z-1-b_A) - P_A(b_A, z-b_A)) \binom{z-1}{b_A} \binom{2n-z}{n-b_A}.
\end{aligned} \tag{18}$$

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

$$\begin{aligned}
P_A(b_A, z-b_A) - P_A(b_A-1, z-b_A) &= P_A(b_A, z-b_A) - p_A P_A(b_A, z-b_A) \\
&\quad - p_B P_A(b_A-1, z-b_A+1) \\
&= p_B [P_A(b_A, z-b_A) - P_A(b_A-1, z-b_A+1)] \\
&= p_B \binom{2n+1-z}{n+1-b_A} p_A^{n+1-b_A} p_B^{n-z+b_A} \\
&= \binom{2n+1-z}{n+1-b_A} p_A^{n+1-b_A} p_B^{n+1-z+b_A}.
\end{aligned} \tag{19}$$

Similarly,

$$\begin{aligned}
P_A(b_A, z-1-b_A) - P_A(b_A, z-b_A) &= p_A P_A(b_A+1, z-1-b_A) + p_B P_A(b_A, z-b_A) \\
&\quad - P_A(b_A, z-b_A) \\
&= p_A (P_A(b_A+1, z-1-b_A) - P_A(b_A, z-b_A)) \\
&= p_A \binom{2n+1-z}{n+1-(b_A+1)} p_A^{n+1-(b_A+1)} p_B^{n-z+b_A+1} \\
&= \binom{2n+1-z}{n-b_A} p_A^{n+1-b_A} p_B^{n+1-z+b_A}. \tag{20}
\end{aligned}$$

Thus, using (19) and (20) into (18), we obtain

$$\begin{aligned}
\Delta(z) &= -\psi(z) + \psi(z-1) \\
&= -\sum_{b_A=1}^z \binom{2n+1-z}{n+1-b_A} p_A^{n+1-b_A} p_B^{n+1-z+b_A} \binom{z-1}{b_A-1} \binom{2n-z}{n-b_A} \\
&\quad + \sum_{b_A=0}^{z-1} \binom{2n+1-z}{n-b_A} p_A^{n+1-b_A} p_B^{n+1-z+b_A} \binom{z-1}{b_A} \binom{2n-z}{n-b_A} \\
&= \binom{2n+1-z}{n} \binom{2n-z}{n} p_A^{n+1} p_B^{n+1-z} - \binom{2n+1-z}{n+1-z} \binom{2n-z}{n-z} p_A^{n+1-z} p_B^{n+1} \\
&\quad + \sum_{b_A=1}^{z-1} \binom{2n-z}{n-b_A} \left[\binom{2n+1-z}{n-b_A} \binom{z-1}{b_A} - \binom{2n+1-z}{n+1-b_A} \binom{z-1}{b_A-1} \right] p_A^{n+1-b_A} p_B^{n+1-z+b_A} \\
&= \binom{2n+1-z}{n} \binom{2n-z}{n} (p_A p_B)^{n+1-z} (p_A^z - p_B^z) \\
&\quad + p_A^{n+1} p_B^{n+1-z} \sum_{b_A=1}^{z-1} \binom{2n-z}{n-b_A} \left(\frac{p_B}{p_A} \right)^{b_A} \left[\binom{2n+1-z}{n-b_A} \binom{z-1}{b_A} - \binom{2n+1-z}{n+1-b_A} \binom{z-1}{b_A-1} \right]. \tag{21}
\end{aligned}$$

Recall that we want to show that $p_A \geq p_B \Rightarrow \Delta(z) \geq 0$ for $z \leq n$ and $z \neq 2$. Given (21) and $p_B = 1 - p_A$, it suffices to prove that

$$p_A \geq p_B \implies \sum_{b_A=1}^{z-1} \binom{2n-z}{n-b_A} d(b_A) \left(\frac{1-p_A}{p_A} \right)^{b_A} \geq 0 \tag{22}$$

$$\text{where } d(b_A) \equiv \left[\binom{2n+1-z}{n-b_A} \binom{z-1}{b_A} - \binom{2n+1-z}{n+1-b_A} \binom{z-1}{b_A-1} \right].$$

The sign of $d(b_A)$ in (22) equals

$$\begin{aligned}
& \operatorname{sgn} \left[\binom{2n+1-z}{n-b_A} \binom{z-1}{b_A} \left(1 - \frac{(n+1-z+b_A)b_A}{(n+1-b_A)(z-b_A)} \right) \right] \\
&= \operatorname{sgn} [(n+1-b_A)(z-b_A) - (n+1-z+b_A)b_A] \\
&= \operatorname{sgn} [(z-2b_A)(n+1)] \\
&= \operatorname{sgn}(z-2b_A). \tag{23}
\end{aligned}$$

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 $b_A = z/2$, if any (i.e., z is even). In words, in the polynomial (22), the first coefficients are positive and the last coefficients are negative since in the summation $b_A = \{1, \dots, z-1\}$. Additionally, the number of positive coefficients equals the number of negative ones. 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 $k = \{0, 1, \dots, z/2-1\}$. 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 (22) satisfy $d(1) = -d(z-1) > 0$, $d(2) = -d(z-2) > 0$, and so on until. To visualize this structure of coefficients and its immediate consequence on proving (22), consider $n = 5$. When also $z = 5$, the coefficients are $\{270, 600, -600, -270\}$. When $z = 4$, the coefficients are $\{1260, 0, -1260\}$.¹⁵ Thus, it is clear from this structure of coefficients that, since each of these coefficients is multiplied respectively by $\left\{ \left(\frac{1-p_A}{p_A} \right)^1, \left(\frac{1-p_A}{p_A} \right)^2, \dots \right\}$, 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 (22). 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 (22). This concludes the proof that $p_A \geq p_B \Rightarrow \Delta(z) \geq 0$ for $z \leq n$ and $z \neq 2$.

¹⁵The case $b = 2$ is special in that there is only one coefficient which equals 0. This is why we analyze separately, in **Step 2**, the case of $b = 2$.

Step 2. Consider $z = 2$. The summation in (21) has only one term ($b_A = 1$). However, since in (23) we proved that $\text{sgn}[c(b_A)] = \text{sgn}(z - 2b_A)$, when $z = 2$ and $b_A = 1$ we have that $d(1) = 0$, and thus the summation in (21) equals 0. Therefore,

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

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

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

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

Similarly,

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

So,

$$\begin{aligned}
\Delta(z) &= - \sum_{b_A=z-n}^n \binom{z-1}{b_A-1} \binom{2n-z}{n-b_A} (P_A(b_A, z-b_A) - P_A(b_A-1, z-b_A)) \\
&\quad + \sum_{j=z-n}^n \binom{z-1}{b_A} \binom{2n-z}{n-b_A} (P_A(b_A, z-1-b_A) - P_A(b_A, z-b_A)) \\
&= - \sum_{b_A=z-n}^n \binom{z-1}{b_A-1} \binom{2n-z}{n-b_A} \binom{2n+1-z}{n+1-b_A} p_A^{n+1-b_A} p_B^{n+1-z+b_A} \\
&\quad + \sum_{b_A=z-n}^n \binom{z-1}{b_A} \binom{2n-z}{n-b_A} \binom{2n+1-z}{n-b_A} p_A^{n-b_A+1} p_B^{n-z+b_A+1} \\
&= p_A^{n+1} p_B^{n+1-z} \sum_{b_A=z-n}^n \binom{2n-z}{n-b_A} \left(\frac{p_B}{p_A}\right)^{b_A} \left[\binom{2n+1-z}{n-b_A} \binom{z-1}{b_A} - \binom{2n+1-z}{n+1-b_A} \binom{z-1}{b_A-1} \right].
\end{aligned} \tag{24}$$

where we used (19) and (20).

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

$$p_A \geq p_B \Rightarrow \sum_{b_A=z-n}^n \binom{2n-z}{n-b_A} d(b_A) \left(\frac{1-p_A}{p_A}\right)^{b_A} \geq 0, \tag{25}$$

which is identical to (22), 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 (25) with respect to that of (22) 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 (25) then follows exactly as that of (22). This concludes the proof that $p_A \geq p_B \Rightarrow \Delta(z) \geq 0$ for $z > n$ and $z \neq 2n$.

Step 4. Consider the case $z = 2n$. By (24), we have

$$\Delta(z) = p_A^{n+1} p_B^{n+1-z} \left[\binom{z-1}{n} - \binom{z-1}{n-1} \right] \left(\frac{p_B}{p_A} \right)^{b_A} = 0,$$

therefore the proof of the lemma is complete. \square

Proof of Lemma 4. Using (4) and (6) into (11), we obtain that WE^d/n_2 equals

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

where ψ is defined in (9) and ξ is defined in (15). As showed in the proof of Lemma 2, ξ is constant and equal to $\binom{2n}{n}$.

Thus, we obtain

$$\frac{WE^T}{n_2 p_A^n p_B^n} = \mu_B \binom{2n}{n} + [\mu_A - \mu_B] \psi(n_1). \quad (26)$$

Applying similar steps to (12), we have

$$\begin{aligned} \frac{WE^{Tmer}}{n_2} &= P_A(0,0) x_A(0,0) + P_B(0,0) x_B(0,0) \\ &= \binom{2n}{n} p_A^n p_B^n [P_A(0,0) \mu_A + P_B(0,0) \mu_B] \\ &= \binom{2n}{n} p_A^n p_B^n [P_A(0,0) (\mu_A - \mu_B) + \mu_B], \end{aligned}$$

so that

$$\frac{WE^{T^{mer}}}{n_2 p_A^n p_B^n} = \mu_B \binom{2n}{n} + [\mu_A - \mu_B] \psi(0). \quad (27)$$

Comparing (26) and (27) establishes the Lemma. \square

Proof of Lemma 5. Start by recasting the differential equation (13) as

$$k' \left(c; \frac{c_A}{c_B} \right) = \left(\frac{c_A}{c_B} \right) \frac{c \cdot f(c)}{k \left(c; \frac{c_A}{c_B} \right) \cdot f \left(k \left(c; \frac{c_A}{c_B} \right) \right)}. \quad (28)$$

By the initial condition $k \left(1; \frac{c_A}{c_B} \right) = 1$, which must hold for any c_A/c_B , we see that

$$k' \left(1; \left(\frac{c_A}{c_B} \right)_H \right) = \left(\frac{c_A}{c_B} \right)_H > \left(\frac{c_A}{c_B} \right)_L = k' \left(1; \left(\frac{c_A}{c_B} \right)_L \right).$$

So in a right-neighborhood of $c = 1$ we have

$$k \left(c; \left(\frac{c_A}{c_B} \right)_H \right) > k \left(c; \left(\frac{c_A}{c_B} \right)_L \right).$$

Suppose now by contradiction that there exists some $c_t > 1$ with

$$k \left(c_t; \left(\frac{c_A}{c_B} \right)_H \right) = k \left(c_t; \left(\frac{c_A}{c_B} \right)_L \right) = k_t.$$

As $k \left(c; \left(\frac{c_A}{c_B} \right)_L \right)$ must have “caught up” to $k \left(c; \left(\frac{c_A}{c_B} \right)_H \right)$ this requires

$$k' \left(c_t; \left(\frac{c_A}{c_B} \right)_H \right) \leq k' \left(c_t; \left(\frac{c_A}{c_B} \right)_L \right). \quad (29)$$

But, given the contradiction hypothesis, (28) gives

$$k' \left(c_t; \left(\frac{c_A}{c_B} \right)_H \right) = \left(\left(\frac{c_A}{c_B} \right)_H \right) \frac{c \cdot f(c)}{k_t \cdot f(k_t)} > \left(\left(\frac{c_A}{c_B} \right)_L \right) \frac{c \cdot f(c)}{k_t \cdot f(k_t)} = k' \left(c_t; \left(\frac{c_A}{c_B} \right)_L \right). \quad (30)$$

Inequality (29) contradicts the extremes of inequality (30). \square

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.** 2004. "Two-player asymmetric contests with ratio-form contest success functions." *Economic Inquiry*, 42: 679-689.
- [3] **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.
- [4] **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.
- [5] **Baye, Michael R., Dan Kovenock, Casper G. De Vries.** 1996. "The all-pay auction with complete information." *Economic Theory*, 8(2): 291-305.
- [6] **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.
- [7] **Decamps, Ambroise, Changxia Ke, and Lionel Page.** 2018. "How success breeds success." Working Paper No. 053. QUT Business School.
- [8] **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.
- [9] **Fu, Qiang, Jingfeng Lu, and Yue Pan.** 2015. "Team Contests with Multiple Pairwise Battles." *American Economic Review*, 105(7): 2120-40.
- [10] **Fu, Qiang, and Zenan Wu.** 2018. "Contests: Theory and Topics." to appear in the *Oxford Research Encyclopedia of Economics*.
- [11] **Häfner, Samuel.** 2017. "A tug-of-war team contest." *Games and Economic Behavior*, 104: 372-391.
- [12] **Hillman, Arye L., and John G. Riley.** 1989. "Politically contestable rents and transfers." *Economics & Politics*, 1(1): 17-39.
- [13] **Klumpp, Tilman and Mattias K. Polborn.** 2006. "Primaries and the New Hampshire effect." *Journal of Public Economics*, 90(6-7): 1073-1114.

- [14] **Konrad, Kai A.** 2009. *Strategy and Dynamics in Contests*. Oxford: Oxford University Press.
- [15] **Konrad, Kai A., and Dan Kovenock.** 2009. “Multi-battle contests.” *Games and Economic Behavior*, 66: 256-274.
- [16] **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.
- [17] **Malueg, David A., and Andrew J. Yates.** 2005. “Equilibria and comparative statics in two-player contests.” *European Journal of Political Economy*, 21: 738-752.
- [18] **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.
- [19] **Serena, Marco.** 2017. “Quality Contests.” *European Journal of Political Economy*, 46: 15-25.