Biasing Dynamic Contests Between Ex-Ante Symmetric Players*

Stefano Barbieri†, Marco Serena‡

October 2, 2020

Abstract

We consider a best-of-three Tullock contest between two ex-ante symmetric players. An effort-maximizing designer commits to a vector of three biases (advantages or disadvantages), one per match. When the designer can choose victory-*dependent* biases (i.e., biases that depend on the record of matches won by players), the effort-maximizing biases eliminate the momentum effect, leaving players equally likely to win each match and the overall contest. Instead, when the designer can only choose victory-*independent* biases, the effort-maximizing biases alternate advantages in the first two matches and leave players not equally likely to win the overall contest. Therefore, in the victory-independent optimal contest, ex-ante symmetric players need not be treated identically, though a coin flip may restore ex-ante symmetry. We analyze several extensions of our basic model, including generalized Tullock contests, ex-ante asymmetric players, best-of-five contests, and winner's effort maximization.

Keywords: dynamic contests, bias, momentum effect.
JEL classification: C72, D72, D74, D81.

*Previous versions of this paper circulated under the name "On the Suboptimality of Perfectly Leveling the Playing Field in Dynamic Contests" and "Biasing Unbiased Dynamic Contests." We are pleased to acknowledge useful comments by three referees of this journal and by Mikhail Drugov, Christian Ewerhart, Jörg Franke, Qiang Fu, Kai Konrad, Dan Kovenock, László Kozma, Jingfeng Lu, Tore Nilssen and Anil Yildizparlak. We would like to thank participants at the 2018 CBESS Conference on Contests (University of East Anglia), at the 30^{th} Tax Day (Max Planck Institute for Tax Law and Public Finance) and the Games and Contests Workshop (Wageningen University). 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

In a contest, players exert costly efforts to win a prize. We call a one-shot contest between players X and Y "unbiased" if swapping the efforts of X and Y implies swapping their probabilities of winning; that is, X and Y are treated equally.[1] Unbiased contests are prevalent in the literature, in part because of the "common wisdom" that unbiased one-shot contests maximize total effort when players are symmetric.[2] In fact, asymmetries between players due to their exogenous characteristics lead to lower aggregate effort in many commonly used setups because of the so-called discouragement effect: "a weaker player, either with higher unit costs of effort or a lower value of winning, finds it relatively unprofitable to try to beat the stronger player and, consequently, cuts back on his costly expenditure. This, in turn, may allow the stronger player to bid more passively as well when compared to a contest in which he faces a player of similar strength" (Dechenaux, Kovenock, and Sheremeta, 2015; p. 622-623). Indeed, the common wisdom has a parallel for asymmetric players; if X is stronger than Y, effort maximization is achieved by giving a disadvantage to X so as to restore a level playing field, and each player has the same probability of winning in equilibrium. This result appears, for instance, in Proposition 2 of Franke (2012) which is stated for a setup similar to ours, but static. It is, in fact, intuitive that a level playing field maximizes competition, and therefore aggregate effort, in a one-shot setup.

We ask whether the common wisdom that leads to the optimality of unbiased contests between two ex-ante symmetric players extends from static to dynamic contests: with endogenous biases in *each* of the matches that make up the overall contest, should one mirror the common wisdom for static contests and induce a level playing field in the overall contest?[3] In general, when moving from a static to a dynamic setting, an important force arises: the *momentum effect*.[4] For instance, a player's loss in the first match of the game gives her a one-match disadvantage in the second

[1] This property is called "anonymity" in the seminal axiomatization of Skaperdas (1996). Our usage of "unbiased" follows Drugov and Ryvkin (2017).

[2] We borrow the term "common wisdom" from Drugov and Ryvkin (2017, p. 118): "The 'common wisdom' prevailing in the literature that it is optimal not to bias the contest when players are symmetric (and thus it is optimal to 'level the playing field' when players are different) has an obvious intuitive appeal." A non-exhaustive list of papers showing that symmetric players should be treated identically includes: Dukerich et al. (1990); Schotter and Weigelt, (1992); Baye, Kovenock, and De Vries (1993) and (1996); Baik (1994); Gradstein (1995); Clark and Riis (2000); Stein (2002); Nti (2004); Fu (2006); Fain (2009); Epstein et al. (2011); Franke (2012); Franke et al. (2013); Lee (2013), along with references in Drugov and Ryvkin (2017). The technical details vary from model to model though they have the key underlying intuition in common. Serena (2017) shows that the common wisdom holds not only for maximization of total effort, but also winner's effort. This common wisdom *does have* exceptions that we discuss in the literature review, with special attention to Drugov and Ryvkin (2017).

[3] Henceforth, we reserve the word "contest" for the overall dynamic competition, and use the "match" for a component competition of a contest.

[4] Early and seminal contributions on the momentum effect appear in Harris and Vickers (1985, 1987) and Budd, Harris and Vickers (1993). For more recent contributions see, for instance, Klumpp and Polborn (2006), Konrad (2009) and Konrad and Kovenock (2009). The momentum effect is sometimes called "avalanche effect" (e.g., Beviá and Corchón, 2013), and sometimes "momentum/discouragement effect" (Feng and Lu, 2018). We use momentum and momentum/discouragement effect interchangeably.

match, and when a contestant is sufficiently disadvantaged, competition—and thus efforts—suffers.[5] Indeed, players in the second match are typically asymmetric, even though they are symmetric to start with, as one player is one match ahead of her rival. Thus, as we investigate the common wisdom in our dynamic setting, we ask whether it is beneficial to mitigate the momentum/discouragement effect by means of biased matches, and if so, how. Furthermore, we investigate whether it is the case that, for biases that maximize total effort in our dynamic setting, the overall ex-ante probability of winning is the same for the two *ex-ante* symmetric players.

Many real-life contests have the two features we capture, namely; (i) dynamics, so that the victory of the contest occurs only if a player wins a sufficient number of matches, and (ii) varying biases as the contest progresses, so that in each match a player may have or may be given a competitive advantage, or a pre-existing competitive advantage may be endogenously mitigated. Perhaps the most natural application for our model is sports, where it is often the case that competitions unfold over time through a series of matches where players' advantages vary. The related literature recognizes the importance of effort maximization in sports.[6] For instance, Feng and Lu (2018), working on a contest design question related to ours and using a model with ex-ante symmetric contestants where effort maximization is the goal, also use sports as their leading example.[7] And even more directly, "organizers of athletic or artistic competitions often need to maximize average performance (or some related measure) in order to thrill audiences" (Moldovanu and Sela, 2001; p. 543-544).[8] In sports, the identity of the player who is advantaged by biases is often endogenous in each match. For instance, the location of each match is endogenous and generates a home-field advantage, which is the benefit that the home team has over the visiting team because of psychological effects (for instance, the effect of supporting fans on home and away teams and referees), physiological effects (the advantage home teams have playing near home in familiar situations, or the disadvantages away teams suffer from travelling), or strategic effects (such as the home team batting second in baseball).[9] In addition, the *extent* to which the designer has control over the size of the biases varies. In fact, the home-field advantage can be mitigated by introducing instant replays to weaken the referees' discretion or by reserving a certain proportion of tickets for the visiting team. Also, the winner of the previous round is sometimes advantaged (e.g., the pool player who shoots a ball into a table's pocket makes the next move) and sometimes disadvantaged (the soccer team that scores a goal does not kick off next).

[5] For empirical evidence of momentum effect see, e.g., Malueg and Yates (2010).

[6] The other prominent objective in sports is the selection of the "better" player as the winner of the competition. As we focus on the competition between two ex-ante symmetric players, this objective is not relevant for our setup.

[7] Furthermore, Feng and Lu (2018) in their footnote 2 cite Gradstein and Konrad (1999), stating that "contest structures result from the careful consideration of a variety of objectives, one of which is to maximize the effort of contenders."

[8] In the extension of Section 6.4, we consider the alternative of winner's effort maximization, as some organizers of sport contests may alternatively be interested in, for instance, having the world record broken during the competition.

[9] Empirical evidence abounds, for instance, in soccer (Nevill and Holder, 1999) and for the NFL (Vergin and Sosik, 1999). Jamieson (2010) provides a general meta-study of home-field advantages.

To account for such varying control over biases across applications, we consider two alternative settings. First, we let the designer have "full" control over biases (Section 4), in the sense that biases can be contingent on the outcome of previous matches; for instance, if X wins today, she will be given a disadvantage tomorrow, otherwise an advantage. This is the case, for instance, in badminton contests which typically have a best-of-three structure where the winner of each match serves at the beginning of the next match (see, Badminton World Federation, 2019).[10] In this setup with victory-dependent biases, we find that the effort-maximizing designer leaves the first and third matches unbiased, and biases the second match in favor of the loser of the first match, to completely compensate for her disadvantage of lagging one match behind. Such structure of biases leaves players equally likely to win each match and the entire contest. In this sense, we conclude that, when the designer can tailor biases to the outcome of previous matches, the common wisdom of equalizing players' equilibrium winning probabilities *extends* from a static setting to our dynamic one. This result contributes to the understanding of whether, in subsequent matches, one should favor the winner or the loser of early matches. In particular, this result contrasts with the "favor-the-leader" result in the literature that we discuss below.

Second, in the perhaps more realistic case of limited control (Section 5), the designer cannot tailor biases to the outcome of previous matches; e.g., whoever plays at home today plays away tomorrow, regardless of who wins today. This is the case, for instance, in volleyball contests which typically have a best-of-five structure where one team begins play by serving in the first and last sets, and the other team begins play by serving in the other sets, regardless of who wins each set (FIVB, 2020).[11] In this setup with victory-independent biases, we find that the effort-maximizing designer does not treat ex-ante symmetric players identically; in particular, rather than leaving each match unbiased, the optimal biases favor one player first and the other later. And the tie-break is not unbiased. Moreover, these biases do not balance out ex-ante; that is, in the bias configuration that maximizes total effort, the overall ex-ante probability of winning is *not* the same for the two ex-ante symmetric players. To clearly identify the forces underlying our results, we add a further intuitive restriction on the designer's choice of biases by analyzing "alternating contests" in which the biases in the first and second match have the same magnitude and favor one player in the first match and the other in the second, whereas the tie-break is left unbiased. We show that, even in this more restrictive setup, ex-ante symmetric players are not treated identically, because biases are introduced in the optimum and the overall ex-ante probability of victory is not identical for the two players.

To understand the intuition behind the optimality of biasing dynamic contests between ex-ante symmetric players, we focus on the simpler intuition behind the optimality of biasing alternating contests. Recall that, in unbiased dynamic contests between ex-ante symmetric players, early

[10] Note that, in badminton, serving is found to be a disadvantage in doubles and an advantage in singles by Bialik (2016), who analyzes data from the 2016 Olympics.

[11] In volleyball, serving is often seen as a disadvantage (e.g., Bialik, 2016).

victories distort future matches so that the laggard gives up and the front-runner eases up. It turns out that this momentum/discouragement effect is mitigated by alternating advantages. Despite creating an asymmetry in the first match—thus reducing first-match efforts—alternating advantages tend to reduce asymmetries in the second match, since the player most likely lagging one match behind in the second match is given an advantage—thus increasing second-match expected efforts. Most importantly, alternating first- and second-match advantages increases the probability that the game will reach a tie-break—thus increasing third-match expected efforts.[12] We show that the second- and third-match positive effects of alternating advantages on efforts overcome the first-match negative effect. We conclude that in the optimal bias configuration, the two ex-ante symmetric players need not be treated identically and one player is more likely to win the overall contest. A similar intuition applies for the optimal biases even if one does not restrict attention to alternating contests. In this sense, the common wisdom discussed above *does not extend* from a static to our dynamic setting if the designer cannot tailor biases to the outcome of previous matches (i.e., for victory-independent biases). Of course, with symmetric players, ex-ante symmetric treatment can be restored by a fair coin flip that decides who should receive the advantage, as it happens often in sports.[13]

Our two main results described above—one with victory-dependent biases (Section 4), and one with victory-independent biases (Section 5)—are derived if the designer maximizes expected total effort in a best-of-three contest in which, in each match, the winner determination process is characterized by the Tullock contest success function with discriminatory parameter $r = 1$ (Tullock, 1980). In Section 6, we show that our two main results are (locally) robust to considering several extensions: discriminatory parameter $r \neq 1$, ex-ante asymmetric players, best-of-five contest, and maximization of expected winner's effort, rather than total effort.[14]

Literature. There are two relevant strands of the literature that are usually—but not exclusively—kept apart; dynamic contests and biased contests. In *dynamic* contests,[15] one important theoretical contribution is that of Klumpp and Polborn (2006); they model the US primaries as a best-of-n contest between two candidates, where the battlefield in each state takes the form of a Tullock contest. As in Harris and Vickers (1985, 1987), they find that the outcome of the very first match creates an asymmetry between ex-ante symmetric players that is endogenously carried over to later periods. This momentum boosts efforts in the first matches and makes the later matches less relevant. This finding resembles the so-called "New Hampshire effect"; candidates who win early primaries are

[12] In fact, the third match is virtually certain to occur if an arbitrarily high advantage is given in the first period to a player and in the second period to her rival.

[13] However, coin flips are sometimes viewed as "unfair" in sports, and several proposals to avoid leaving such an important role for luck have been presented (see, for instance, the auction approach of Che and Hendershott, 2008).

[14] We present a detailed description of the changes that each extension introduces in Section 6.

[15] One of the early contributions to dynamic Tullock contests with multiplicative biases is Leininger (1993), who is mostly interested in players' choice of order of moves.

more likely to win later primaries, too.[16] Similar findings appear in Konrad and Kovenock (2009), who model matches as all-pay auctions, and in Ferrall and Smith (1999), who adopt the rank-order tournaments of Lazear and Rosen (1981). Malueg and Yates (2006) generalize Klumpp and Polborn's (2006) results to a more general symmetric contest success function and derive results for a three-match contest assuming the existence of a pure strategy equilibrium.[17] All the above models, while more general than ours in many other respects, do not consider the question of how to bias *match-by-match* the overall contest to maximize total effort.[18]

There is an extensive literature on *biases* in static contests, where biases are defined as tilting the probability of winning in favor of one (or more) contestants, for given efforts.[19] The common wisdom typically drawn from this strand of the literature is the optimality of an unbiased contest if the two players are symmetric (see Footnote 2). A recent prominent exception to this common wisdom for static contests is presented in Drugov and Ryvkin (2017);[20] they characterize very general properties of the contest success function and of the cost of effort that determine whether a biased or unbiased contest between two ex-ante symmetric players is optimal. In the standard Tullock contest with linear costs of effort, multiplicative bias, and static competition, Drugov and Ryvkin (2017), among other results, show that the common wisdom of the optimality of unbiased contests is not robust to dropping the assumption of multiplicative biases. In contrast, we maintain multiplicative biases, but investigate the common wisdom in a dynamic, rather than static, setting.

We are not the first to consider biases in dynamic contests (see, for instance, Meyer, 1991, 1992; Lizzeri et al., 1999, 2002; Höffler and Sliwka, 2003; Aoyagi, 2010; Ederer, 2010).[21] A common finding is the "favor-the-leader" result; for example, Meyer (1991, 1992) shows that biasing the second match in favor of whoever performed better in the first match tends to be beneficial for the principal in terms of better information and of larger efforts, as one obtains only a second-

[16] A victory in the New Hampshire primary increases a candidate's expected share of total primary votes by 26.6% (Mayer, 2004).

[17] Empirical tests of theoretical predictions with sports data is provided for best-of-three contests by Malueg and Yates (2010) and for best-of-n by Ferrall and Smith (1999). In the experimental literature, a test using best-of-three Tullock contests is provided by Mago et al. (2013).

[18] For instance, Klumpp and Polborn (2006) allow for biases, but do not analyze the effort-maximizing set of biases. Konrad and Kovenock (2009) allow for asymmetries and show that in this case headstarts (a situation in which the number of component contests that need to be won to secure overall victory varies across players with different values for victory) could be used to generate maximal expected effort.

[19] Early contributions where the bias enters the probability of victory is, for instance, in the corruption models of Lien; in particular, Lien (1990), building on Lien (1986), considers a two-briber game where bribes are denoted by B_i, and the corrupt official determines the winner "not only by the amount of bribe, but by some other considerations such as friendship." Friendships are captured by a parameter $\alpha > 0$, such that the winner is determined by the ranking between αB_1 and B_2. Our biases are modeled essentially in the same way.

[20] Other exceptions to the common wisdom can be derived from an extension of the model to an ex-ante heterogenous n-player setting (e.g., Franke et al., 2013), to a private information setting (e.g., Pérez-Castrillo and Wettstein, 2016), and to the maximization of the probability of a high-ability winner (e.g., Kawamura and Moreno de Barreda, 2014). In the present paper, we keep the standard two-player complete information setting under effort-maximization.

[21] Another tool that could also affect the contest outcome is the extra information that may or may not be disclosed as the dynamic contest unravels, so that the designer induces a certain dynamic inference by players who may therefore be discouraged or encouraged.

order negative effect of bias in the second period, but a first-order positive effect of bias in the first period. More recently, Clark et al. (2012), Möller (2012), Beviá and Corchón (2013), Esteve-Gonzáles (2016), and Klein and Schmutzler (2017) analyze a two-period setting, assuming that if a player wins (or exerts higher effort) in the first period she will have a certain *advantage* in the second period—modeled as a lower marginal cost of effort, or as a favorable bias in the contest success function.[22] There are several differences between the setups of those papers and ours. For one, with victory-dependent biases, we allow for second-period advantages or disadvantages to the first-period winner. But the key difference in generating our contrasting result is that, unlike in the two-match setting of the above papers and of Meyer (1991, 1992), in our best-of-three setting the length of the game depends on the outcome of previous matches: contestants play a tie-break only if they are tied after two matches. We discuss this further in Section 4.

Related to our setup with victory-dependent biases is Feng and Lu (2018), who analyze the effort-maximizing allocation of prizes among battles in a three-battle contest with a more general prize structure than we consider. In contrast, our model endogenizes biases rather than the allocation of prizes among matches and allows biases to differ across matches and to treat players differently. Feng and Lu's "main insight is that mitigating the momentum/discouragement effect is essential for effort-maximizing prize design in dynamic multi-battle contests" (Feng and Lu, 2018; p. 82). Similarly, our results call for the complete elimination of the momentum/discouragement effect, in accordance with the insight of Feng and Lu, when we consider victory-dependent biases. Furthermore, we highlight that whether or not biases are victory dependent plays a crucial role in addressing the momentum/discouragement effect.

Finally, our main findings share some common intuition with the "suspense" literature (e.g., Chan, Courty and Hao, 2009; Ely, Frankel and Kamenica, 2015). Suspense is defined by Chan et al. (2009) as valuing contestants' efforts more in a close race; they show that rank-order schemes depending on who wins are more preferred than schemes depending linearly on the final score when preference for suspense increases. Suspense is defined by Ely et al. (2015) as the variance of next period's beliefs; they analyze how to reveal information over time to a Bayesian audience valuing suspense. Our model does not have private information. But in a way related to the suspense literature, our vector of optimal biases keeps the contest sufficiently "open" over time, so as not to deter competition and efforts, which are of value to the designer.[23]

[22] An exception is Ridlon and Shin (2013), who find that, in a setup with ex-ante asymmetry in players' abilities, if "abilities are sufficiently different, favoring the first-period loser in the second period increases the total effort over both periods. However, if abilities are sufficiently similar, [...] total effort increases the most in response to a handicapping strategy of favoring the first-period winner" (Ridlon and Shin, 2013; p. 768).

[23] Konrad and Kovenock (2009) in their Footnote 5 also point out the relation to the notion of suspense.

2 Model

Two risk-neutral and ex-ante symmetric players, X and Y, play in a best-of-three contest. That is, they play three matches at most, and the first player who wins two matches is the contest winner and obtains a prize of value $V > 0$. In Figure 1, we draw the structure of the best-of-three contest we describe. The game begins at node $(0,0)$, where no player has won a match; here, the effort of X is denoted as $x^{(0,0)} \geq 0$ and that of Y as $y^{(0,0)} \geq 0$. If X wins, the game then moves to node $(1,0)$; here, the effort of X is denoted as $x^{(1,0)} \geq 0$ and that of Y as $y^{(1,0)} \geq 0$. If instead Y wins the first match in node $(0,0)$, the game then moves to node $(0,1)$; here, the effort of X is denoted as $x^{(0,1)} \geq 0$ and that of Y as $y^{(0,1)} \geq 0$. If the first two matches are won by the same player the game ends, otherwise the game reaches node $(1,1)$ and the third match is played; here, the effort of X is denoted as $x^{(1,1)} \geq 0$ and that of Y as $y^{(1,1)} \geq 0$. For $i, j \in \{0,1\}$, if node (i,j) is reached, players simultaneously choose efforts $(x^{(i,j)}, y^{(i,j)}) \in [0,\infty)^2$, and the probability of victory of player X in that match depends on the contest technology as follows:

$$p_X^{(i,j)}(x^{(i,j)}, y^{(i,j)}) = \begin{cases} \frac{\alpha_{(i,j)} x^{(i,j)}}{\alpha_{(i,j)} x^{(i,j)} + y^{(i,j)}} & \text{if } (x^{(i,j)}, y^{(i,j)}) \neq (0,0)\,, \\ \frac{1}{2} & \text{if } (x^{(i,j)}, y^{(i,j)}) = (0,0)\,, \end{cases} \tag{1}$$

where $\alpha_{(i,j)} \in (0,\infty)$ and $p_Y^{(i,j)}(x^{(i,j)}, y^{(i,j)}) = 1 - p_X^{(i,j)}(x^{(i,j)}, y^{(i,j)})$.[24] The contest technology (1) is axiomatized in Clark and Riis (1998) and micro-founded in Jia (2008). When $\alpha_{(i,j)} = 1$, (1) boils down to the contest technology proposed by Tullock (1980). We refer to $\alpha_{(i,j)} > 1$ ($\alpha_{(i,j)} < 1$) as an *advantage* (*disadvantage)* given to X in node (i,j). The vector of biases $\{\alpha_{(0,0)}, \alpha_{(1,0)}, \alpha_{(0,1)}, \alpha_{(1,1)}\}$ is commonly known at the beginning of the game. The marginal cost of effort equals 1 for both players and there is complete information. In Figure 1, we specify the payoffs gross of effort costs; net payoffs simply subtract the cost of effort.

We analyze how the expected total effort (henceforth, TE) varies with the biases. We define as "optimal" the contest that maximizes TE, which is

$$\begin{aligned} TE \equiv \left(x^{(0,0)} + y^{(0,0)}\right) + p_X^{(0,0)}\left(x^{(1,0)} + y^{(1,0)}\right) + p_Y^{(0,0)}\left(x^{(0,1)} + y^{(0,1)}\right) \\ + \left(p_X^{(0,0)} p_Y^{(1,0)} + p_Y^{(0,0)} p_X^{(0,1)}\right)\left(x^{(1,1)} + y^{(1,1)}\right). \end{aligned} \tag{2}$$

For simplicity, in (2) we omitted the arguments of probabilities, as we will do throughout the paper whenever this does not yield confusion.

[24] In Section 6.1 we consider the generalization of (1)

$$p_X^{(i,j)}(x^{(i,j)}, y^{(i,j)}) = \frac{\alpha_{(i,j)}\left(x^{(i,j)}\right)^r}{\alpha_{(i,j)}\left(x^{(i,j)}\right)^r + \left(y^{(i,j)}\right)^r},$$

with $r \in (0,\infty)$.

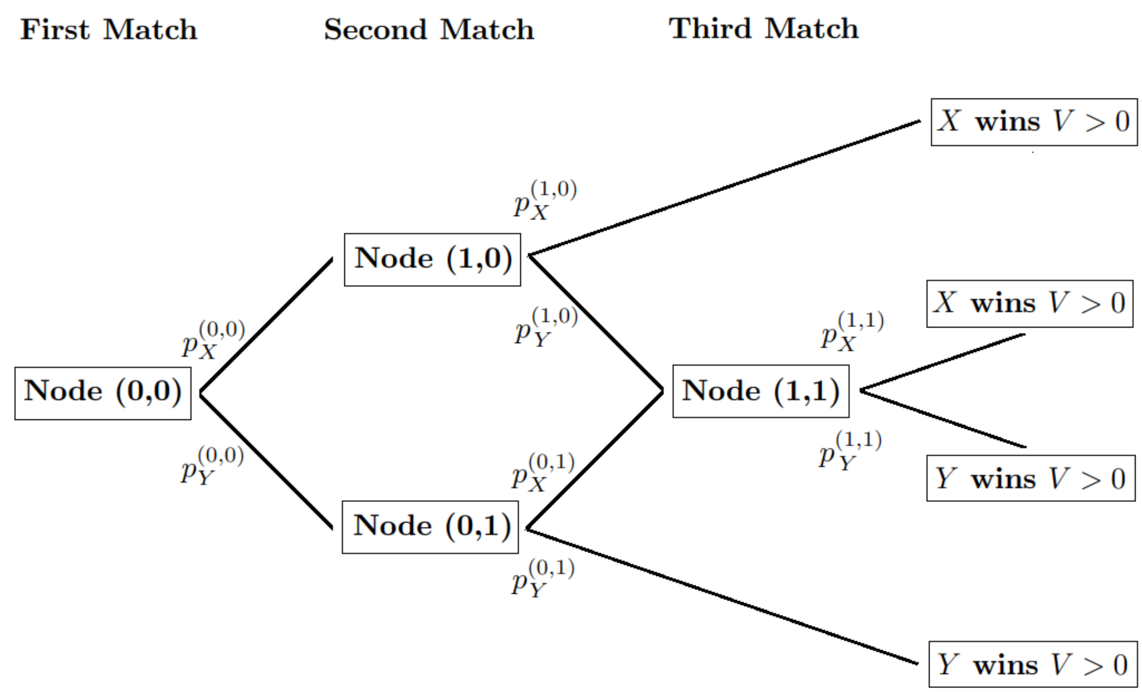

Figure 1: Nodes, matches, and payoffs gross of effort costs in a best-of-three contest between players X and Y.

Structure of the paper. In Section 3, we provide general preliminary results, well-known in the literature, that apply to each node. In Section 4, biases are victory-*dependent*; that is, a possibly different α is chosen for each node, so that four biases $\{\alpha_{(0,0)}, \alpha_{(1,0)}, \alpha_{(0,1)}, \alpha_{(1,1)}\}$ are chosen in order to maximize TE. In Section 5, instead, biases are victory-*independent*; that is, α cannot be conditioned on the outcome of the previous matches, so that three biases $\{\alpha_1, \alpha_2, \alpha_3\}$ are chosen in order to maximize TE; in the first match (i.e., node $(0,0)$) player X is given bias α_1, in the second match (i.e., node $(1,0)$ or $(0,1)$) she is given bias α_2, and in the third match (i.e., node $(1,1)$), if played, she is given bias α_3. The extra constraint of victory-independent biases is inspired by the applications, as discussed in the Introduction. We conclude by considering several extensions: Section 6.1 considers the generalization of (1) described in Footnote 24, Section 6.2 considers ex-ante asymmetric players, Section 6.3 considers a best-of-five contest, and Section 6.4 considers the maximization of expected winner's effort, rather than total effort. Proofs are in Appendix A for sections 3-5, and in Appendix B for the extensions in Section 6.1.

3 Preliminaries

Denoting with u_X^W the expected (continuation) payoff of player X if she wins, and with u_X^L her expected (continuation) payoff if she loses, her payoff u_X in a general node with bias α reads

$$u_X = \frac{\alpha x}{\alpha x + y} u_X^W + \left(1 - \frac{\alpha x}{\alpha x + y}\right) u_X^L - x = \frac{\alpha x}{\alpha x + y}\left(u_X^W - u_X^L\right) + u_X^L - x.$$

Defining the "effective prize spread" as $\Delta u_X \equiv u_X^W - u_X^L$, we obtain

$$u_X = \frac{\alpha x}{\alpha x + y} \Delta u_X + u_X^L - x.$$

For player Y, we use a similar notation and obtain

$$u_Y = \frac{y}{\alpha x + y} \Delta u_Y + u_Y^L - y.$$

The equilibrium is uniquely identified by the FOCs, which give the typical property[25]

$$y = \frac{\Delta u_Y}{\Delta u_X} x;$$

[25] Whenever confusion does not arise, we do not differentiate notation between equilibrium levels and generic variables.

the equilibrium efforts are

$$x = \frac{\alpha \left(\Delta u_X\right)^2 \left(\Delta u_Y\right)}{\left(\alpha \Delta u_X + \Delta u_Y\right)^2}, \tag{3}$$

$$y = \frac{\alpha \Delta u_X \left(\Delta u_Y\right)^2}{\left(\alpha \Delta u_X + \Delta u_Y\right)^2}; \tag{4}$$

the equilibrium probabilities of victory

$$p_X = \frac{\alpha \Delta u_X}{\alpha \Delta u_X + \Delta u_Y}, \tag{5}$$

$$p_Y = \frac{\Delta u_Y}{\alpha \Delta u_X + \Delta u_Y}; \tag{6}$$

and the equilibrium payoffs

$$u_X = \Delta u_X p_X \left(1 - p_Y\right) + u_X^L, \tag{7}$$

$$u_Y = \Delta u_Y p_Y \left(1 - p_X\right) + u_Y^L. \tag{8}$$

Summing (3) and (4), and using (5) and (6), we obtain the following property, which we use repeatedly in our proofs,

$$x + y = \left(\Delta u_X + \Delta u_Y\right) \cdot p_X p_Y. \tag{9}$$

The common wisdom that in a *static* contest players should be left equally likely to win in equilibrium is apparent in (9); for given Δu_X and Δu_Y, $x + y$ is maximized by choosing α such that $p_X = p_Y = 1/2$ (recall that $p_X + p_Y = 1$).

In the following lemma, we consider the game players play after all biases become known. We apply the above analysis for each node, $(0,0)$, $(1,0)$, $(0,1)$, and $(1,1)$, substituting the appropriate continuation value. We then obtain that, in the model described in Section 2, TE in (2) can be expressed as a function of the equilibrium probabilities of victory in each node as stated in the following lemma.

Lemma 1 *Consider a best-of-three Tullock contest between two ex-ante symmetric players. The equilibrium probabilities of victory for X in each node are recursively determined as a function of the vector of biases $\{\alpha_{(0,0)}, \alpha_{(1,0)}, \alpha_{(0,1)}, \alpha_{(1,1)}\}$ as follows:*

$$p_X^{(1,1)} = \left(1 + \frac{1}{\alpha_{(1,1)}}\right)^{-1}, \tag{10}$$

$$p_X^{(0,1)} = \left(1 + \frac{1}{\alpha_{(0,1)}} \frac{1 + p_Y^{(1,1)}}{p_X^{(1,1)}}\right)^{-1}, \tag{11}$$

$$p_X^{(1,0)} = \left(1 + \frac{1}{\alpha_{(1,0)}} \frac{p_Y^{(1,1)}}{1 + p_X^{(1,1)}}\right)^{-1}, \tag{12}$$

and

$$p_X^{(0,0)} = \left(1 + \frac{1}{\alpha_{(0,0)}} \frac{\left(p_Y^{(0,1)}\right)^2 p_X^{(1,1)} \left(1 + p_Y^{(1,1)}\right) + p_X^{(1,0)} \left(p_Y^{(1,1)}\right)^2 \left(1 + p_Y^{(1,0)}\right)}{\left(p_X^{(1,0)}\right)^2 p_Y^{(1,1)} \left(1 + p_X^{(1,1)}\right) + \left(p_X^{(1,1)}\right)^2 p_Y^{(0,1)} \left(1 + p_X^{(0,1)}\right)}\right)^{-1}, \tag{13}$$

where $p_Y^{(i,j)} = 1 - p_X^{(i,j)}$ *for any* $i, j \in \{0, 1\}$*. Furthermore, in equilibrium,* TE *in (2) satisfies*

$$\begin{aligned} \frac{TE}{2V} = & \left(p_X^{(1,0)} p_Y^{(1,1)} + p_X^{(1,1)} p_Y^{(0,1)} - p_X^{(1,1)} p_Y^{(1,1)} \left(p_X^{(0,1)} p_Y^{(0,1)} + p_X^{(1,0)} p_Y^{(1,0)}\right)\right) p_X^{(0,0)} p_Y^{(0,0)} \\ & + p_X^{(0,0)} p_Y^{(1,1)} p_X^{(1,0)} p_Y^{(1,0)} + \left(1 - p_X^{(0,0)}\right) p_X^{(0,1)} p_Y^{(0,1)} p_X^{(1,1)} \\ & + \left(p_X^{(0,0)} p_Y^{(1,0)} + p_Y^{(0,0)} p_X^{(0,1)}\right) p_X^{(1,1)} p_Y^{(1,1)}. \end{aligned} \tag{14}$$

Proof. See Appendix A. ■

By changing the biases $\{\alpha_{(0,0)}, \alpha_{(1,0)}, \alpha_{(0,1)}, \alpha_{(1,1)}\}$, one can generate different winning probabilities at each node and therefore affect TE. As $2V$ is independent of biases, in what follows we often refer to the function $\tau(A, B, C, D)$ defined as

$$\begin{aligned} \tau(A, B, C, D) \equiv & \left(D(1 - D) B^2 + D^2 C (1 - C) + D (1 - C)^2 + B (1 - D)^2\right) A (1 - A) \\ & + A (1 - D) B (1 - B) + (1 - A) DC (1 - C) \\ & + (A (1 - B) + (1 - A) C) D (1 - D). \end{aligned} \tag{15}$$

Using (14), one can easily verify that $\tau\left(p_X^{(0,0)}, p_X^{(1,0)}, p_X^{(0,1)}, p_X^{(1,1)}\right) = TE/(2V)$. Therefore, maximizing TE is equivalent to maximizing τ.

4 Victory-dependent biases

The problem of maximizing TE when a possibly different α is chosen at each node has four choice variables: $\{\alpha_{(0,0)}, \alpha_{(1,0)}, \alpha_{(0,1)}, \alpha_{(1,1)}\} \in \mathbb{R}^4_{>0}$. As we can see from (5), with victory-dependent biases one can generate any probability of victory between 0 and 1. Therefore, maximizing TE is equivalent to

$$\max \tau\left(p_X^{(0,0)}, p_X^{(1,0)}, p_X^{(0,1)}, p_X^{(1,1)}\right),$$

where the choice variables are the four probabilities $\left\{p_X^{(0,0)}, p_X^{(1,0)}, p_X^{(0,1)}, p_X^{(1,1)}\right\} \in (0, 1)^4$ and there are no other constraints. We obtain the following result.

Proposition 1 *Consider a best-of-three Tullock contest between two ex-ante symmetric players. With victory-dependent biases, the point* $\{\alpha_{(0,0)}, \alpha_{(1,0)}, \alpha_{(0,1)}, \alpha_{(1,1)}\} = \{1, 1/3, 3, 1\}$ *is the unique global maximum for* TE *in* $\mathbb{R}^4_{>0}$. *Thus, in the victory-dependent optimal contest* $p_X^{(i,j)} = 1/2$ *for any* $i, j \in \{0, 1\}$.

Proof. See Appendix A. ■

The "asymmetric" nodes $(1,0)$ and $(0,1)$, where one player is leading by one match, are affected by the momentum/discouragement effect typically present in dynamic contests without endogenous biases. With endogenous biases, Proposition 1 shows that it is optimal to give the laggard an advantage in these asymmetric nodes; namely, if X loses the first match, the game reaches node $(0,1)$ and X is given an advantage of $\alpha_{(0,1)} = 3$, while if X wins the first match, the game reaches node $(1,0)$ and X is given a disadvantage of $\alpha_{(1,0)} = 1/3$. These values of $\alpha_{(0,1)}$ and $\alpha_{(1,0)}$ counteract the unbalanced competition due to the momentum/discouragement effect and leave players equally likely to win the second match in equilibrium, as Proposition 1 shows. The "symmetric" nodes $(0,0)$ and $(1,1)$, where players have identical continuation values and have won an identical number of matches, are optimally left unbiased by setting $\alpha_{(0,0)} = \alpha_{(1,1)} = 1$. Therefore, the unique global maximum of Proposition 1 leaves players equally likely to win each match and the entire contest; this is the victory-dependent optimal contest.

The result in Proposition 1 may at first seem to be an intuitive extension of the common wisdom of static contest. However, further reflection reveals it to be surprising. Intuitively, if today's victory grants an advantage tomorrow, a player fights fiercely today so as to require less effort to win tomorrow. This simple intuition has been extensively analyzed by the literature. The general finding, stemming from Meyer (1992), is that the second match should be biased in favor of the winner of the first match in order to increase total effort because, "...starting with no bias, the introduction of a small amount in favor of the first-period winner generates a first-order increase in first-period incentives, but only a second-order reduction in second-period incentives..." (see Meyer, 1992; p. 167). Ridlon and Shin (2013), when analyzing competition between two employees of asymmetric abilities, find an analogous result if abilities are sufficiently similar; "total effort increases the most in response to a handicapping strategy of favoring the first-period winner" (Ridlon and Shin, 2013; p. 1). There are several differences between the setups of those papers and ours.[26] But the key difference in generating our contrasting result is that in our setup the length

[26] "Favor-the-leader" results are not universally true in all contest models, and thus it is informative to briefly describe the setups of Meyer (1992) and Ridlon and Shin (2013). In Meyer (1992), two agents, indexed by k, exert effort a_k each, and agent k's per-period output is given by $x_k = f(a_k, s) + \varepsilon_k$, where s is a common shock affecting the production of both agents and ε_k is the individual-specific shock. The function f is strictly increasing and weakly concave in a_k. The per-period winner is agent i (j) if $x_i + c > (<) x_j$, where c is the bias i's output. Ridlon and Shin (2013) consider a two-period model, where in each period two players compete in a Tullock contest for a fixed prize v, with multiplicative biases, like ours. The designer does not know the identity of the contestants (who is

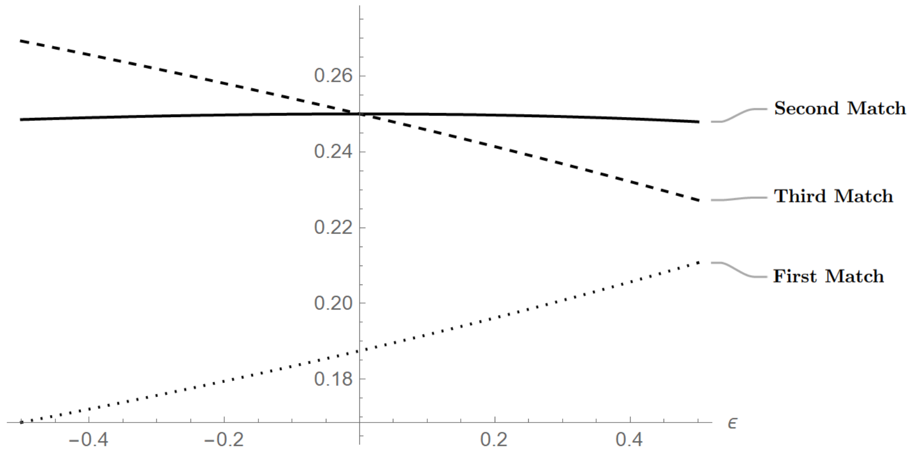

Figure 2: Equilibrium efforts per match as a function of ε when $\alpha_{(0,0)} = \alpha_{(1,1)} = 1$, $\alpha_{(0,1)} = 3 - \varepsilon$, and $\alpha_{(1,0)} = 1/(3 - \varepsilon)$.

of the game depends on the results of previous matches: contestants play a tie-break only if they are tied. In contrast, all two matches are always played in Meyer (1992). This is very important for us, as we explain below.

The intuition behind the "favor-the-leader" literature would suggest, in our setup, leaving some small advantage in the second match to the winner of the first match by setting $\alpha_{(0,1)} = 3 - \varepsilon$ with some small $\varepsilon > 0$ and symmetrically $\alpha_{(1,0)} = 1/(3 - \varepsilon)$, which would result in $p_X^{(0,1)} < \frac{1}{2}$ and $p_X^{(1,0)} > \frac{1}{2}$. In Figure 2, we plot the efforts of each match, weighted by the corresponding probability of reaching such match, as a function of ε. One can see that, in our setup, setting $\varepsilon > 0$ would result in a first-order increase in first-period efforts and a negligible decrease in second-period efforts, in line with the favor-the-leader literature. However, in our *best-of-three* setup, a new force arises: setting $\varepsilon > 0$ has an extra negative first-order effect due to a lower probability of reaching node $(1, 1)$. Proposition 1 shows that all first-order effects balance out and, when considering second-order effects, we find that the beneficial effect of setting $\varepsilon = 0$ in terms of probability of reaching node $(1, 1)$ overwhelms the favor-the-leader increase in first-period efforts that setting $\varepsilon > 0$ generates.

One may also wonder why the victory-dependent optimal contest does not give the loser of the first match a greater than half probability of winning the second match; in fact, with favor-the-loser biases, one restores second-match competition, which is beneficial to efforts. One can see the answer by setting $\varepsilon < 0$ in the above favor-the-leader exercise. In Figure 2, we show that, in

strong and who is weak), but the outcome of the first period serves as a source of noisy information for the designer about players' identities. The designer precommits to a bias, which she assigns in the second period to one of the two players after having observed the outcome of the first period, and thus having updated her belief about who is the strong player. They find that, if abilities are sufficiently similar, total effort increases the most in response to a favor-the-leader policy.

our setup, setting $\varepsilon < 0$ would result in a first-order decrease in first-period efforts, a negligible change in second-period efforts, and a first-order increase in the third-period expected effort, as the probability of reaching the third period decreases in ε. As above, Proposition 1 shows that all first-order effects balance out and, when considering second-order effects, we find that the detrimental effect in first-period efforts dominates the beneficial effect of a greater probability of reaching the third period that setting $\varepsilon < 0$ generates.[27]

Thus, the common wisdom for static contests, namely to leave players equally likely to win in equilibrium, extends to each node and to the entire contest in our dynamic model.[28] However, this is clearly possible because the designer can tailor biases to the outcome of previous matches, so as to keep competition fierce at all nodes and counteract the momentum/discouragement effect. In the remainder of the paper we characterize and discuss the victory-*in*dependent optimal contest: the optimal vector of biases when the designer cannot tailor biases to the outcome of previous matches.

5 Victory-independent biases

Under victory-independent biases, the model has the extra constraint $\alpha_{(0,1)} = \alpha_{(1,0)}$ with respect to Section 4. More generally, biases depend on matches only, not on nodes, so that $\alpha_{(i,j)} = \alpha_{i+j+1}$ with $i, j \in \{0, 1\}$; that is, α_{i+j+1} is given to player X in the $(i+j+1)^{th}$-match *regardless* of the outcome of previous matches. Thus, the vector of biases that can be used to maximize TE is $\{\alpha_1, \alpha_2, \alpha_3\}$: one bias per match. Proposition 1 shows that, with victory-dependent optimal biases, $\alpha_{(0,1)} \neq \alpha_{(1,0)}$. Hence, we obtain the following result.

Corollary 1 *Consider a best-of-three Tullock contest between two ex-ante symmetric players. Victory-independent biases are not optimal in the larger class of contests with victory-dependent biases.*

Furthermore, in a contest with victory-independent biases, it is clearly impossible to induce an equal equilibrium probability of winning across players at each node of the contest. An easy way to see this is that, since α_2 cannot depend on the outcome of the first match, one can set α_2 so as to achieve at most one between $p_X^{(1,0)} = 1/2$ and $p_X^{(0,1)} = 1/2$. In this sense, the momentum/discouragement effect cannot be completely eliminated, as with victory-dependent biases (Section 4), but can be at most mitigated.

An equal equilibrium probability of winning across players at each node implies an equal equilibrium probability of winning the entire contest. Since the former can no longer be achieved,

[27] In Figure 2, summing the efforts of the three matches, the graph of TE would be inverse U-shaped in ε with maximum at $\varepsilon = 0$, consistent with the finding of Proposition 1.

[28] Our main question is whether the common wisdom of the optimality of unbiased contests between two symmetric players extends to a *dynamic* contest. But it is also interesting to see whether the related result for asymmetric players discussed in the Introduction, i.e., that one should bias a static contest in favor of the weaker player to ensure that each player has the same probability of victory in equilibrium (see e.g., Proposition 2 in Franke, 2012), also holds in our dynamic model. In Section 6.2, we show that the answer crucially depends on the form of the asymmetry between players.

it is natural to ask whether the latter remains optimal, and in particular whether the victory-independent optimal contest is achieved at $\alpha_1 = \alpha_2 = \alpha_3 = 1$, which is what we call a *fully unbiased* contest. We find that a fully unbiased contest does not maximize TE.

Proposition 2 *Consider a best-of-three Tullock contest between two ex-ante symmetric players. With victory-independent biases, the fully unbiased contest $\{\alpha_1, \alpha_2, \alpha_3\} = \{1, 1, 1\}$—thus, $p_X^{(0,0)} = \frac{1}{2}, p_X^{(1,0)} = \frac{3}{4}, p_X^{(0,1)} = \frac{1}{4}, p_X^{(1,1)} = \frac{1}{2}$—does not maximize TE in $\mathbb{R}^3_{>0}$. Therefore, the fully unbiased contest is not the victory-independent optimal contest.*

Proof. See Appendix A. ■

The main conclusion of the above proposition is that the fully unbiased contest is not optimal within the family of contests with victory-independent biases. However, one can also show that the fully unbiased contest does satisfy the FOCs for the maximization of TE. Hence, this underscores the complexity of the problem, since the FOCs are not sufficient to fully characterize the maximum, but they only provide necessary conditions.[29]

Below, we first explain the intuition behind the result in Proposition 2, and then provide an analytical characterization of the contest with victory-independent optimal biases.

Intuition. We provide the intuition behind the non-optimality of a fully unbiased contest by showing that TE increases as we move away from a fully unbiased contest to a particular structure of biases, which we call *alternating contest*. In an alternating contest, one player has an advantage in the first match, her rival an advantage of the same magnitude in the second, and the tie-break is unbiased; that is, $\{\alpha_1, \alpha_2, \alpha_3\} = \{\alpha, 1/\alpha, 1\}$. We illustrate that, within the family of alternating contests, moving away from $\alpha = 1$ and setting, for instance, $\alpha = 2$ increases TE. One can follow the mathematics behind the reasoning below with the help of the two figures of Appendix C, one depicted for $\alpha = 1$ and the other for $\alpha = 2$. What happens when moving from $\alpha = 1$ to $\alpha = 2$?

In the *first* match, since players are ex-ante symmetric, setting $\alpha = 2$ creates an asymmetry that unbalances the first match, hence reducing efforts. This parallels the common wisdom for static contests. Thus, setting $\alpha = 2$ has a *negative* effect on first-match efforts. In particular, the decrease in first-match efforts is $\frac{21}{64} - \frac{26}{81} \simeq 0.007$.

In the *second* match, setting $\alpha = 2$ gives player X an advantage in the first match and thus she will most likely be the winner of the first match, but the second match is biased against her. Hence, the effect of the second-match bias against her is more likely to attenuate than enhance her lead; in other words, the second-match bias is more likely to help the second-match laggard rather than the second-match leader. Thus, setting $\alpha = 2$ has a *positive* effect on second-match efforts.

[29] In numerical analysis, one of the commonly used methods is Newton's method, as described, for instance, in Judd (1998, Ch. 4.3, p. 103–104), who warns to check whether the Hessian matrix is definite. This warning is very important in our case as Proposition 2 demonstrates, since the fully unbiased contest satisfies the FOCs but it is not a maximum.

In the *third* match, setting $\alpha = 2$ increases the probability of reaching the third match at all, as opposed to $\alpha = 1$; this is easy to see for extreme alternating biases ($\alpha \to \infty$), where the third match is reached with certainty. Total effort increases with the probability of reaching the third match. Thus, setting $\alpha = 2$ has a *positive* effect on third-match efforts. In particular, the beneficial effect on efforts of increasing the probability of reaching the third match is $\frac{1}{7} - \frac{1}{8} \simeq 0.018$.

All in all, we can neglect the second-match positive effect on efforts, as the third-match positive effect alone (0.018) suffices to overcome the first-match negative effect (0.007). In words, when moving from a fully unbiased contest to an alternating contest ($\alpha = 2$), the beneficial effect on the efforts to increase the probability of reaching the third match overcomes the decrease in first-match efforts due to the lack of balance in the first match.

While setting $\alpha = 2$ suffices to show that a fully unbiased contest ($\alpha = 1$) can be improved upon, Proposition 3 below shows that the TE-maximizing contest in the family $\{\alpha, 1/\alpha, 1\}$ has $\alpha^* \in (4.2, 4.3)$. Furthermore, while Proposition 2 shows that the fully unbiased contest $\{\alpha_1, \alpha_2, \alpha_3\} = \{1, 1, 1\}$ is not optimal, it remains of interest the direction in which one could move to improve upon the fully unbiased contest, and Proposition 3 shows that one such direction is precisely that of an alternating contest.

Proposition 3 *Consider a best-of-three Tullock contest between two ex-ante symmetric players. With victory-independent biases and within the family of alternating contests* $\{\alpha_1, \alpha_2, \alpha_3\} = \{\alpha, 1/\alpha, 1\}$*, let* $TE(\alpha)$ *be* TE *as a function of* $\alpha \in (0, \infty)$*. Setting* $\alpha = 1$ *(the fully unbiased contest) gives a local minimum for* $TE(\alpha)$*. There is a unique* α^* *such that* α^* *and* $1/\alpha^*$ *are the only two global maximizers of* $TE(\alpha)$*. Furthermore,* α^* *is found as the unique solution larger than* 1 *of* $\partial TE(\alpha) / \partial \alpha = 0$*, and we find that* $\alpha^* \in (4.2, 4.3)$*. Finally, the only value of* α *such that the ex-ante probability of victory are identical across players is* $\alpha = 1$*; thus, the optimal alternating contest gives different ex-ante probabilities of victory across symmetric players.*

Proof. See Appendix A. ■

Numerical simulations show that the probability of reaching the third match sharply increases (from 0.25 to 0.41) as one moves from $\alpha = 1$ to $\alpha = \alpha^*$, in line with the intuition discussed above. An immediate consequence of Proposition 3 is the following:

Corollary 2 *Consider a best-of-three Tullock contest between two ex-ante symmetric players. A fully unbiased contest is not optimal in the class of alternating contests.*

In the remainder of this section, we provide analytical and numerical features of the victory-independent optimal contest, which is the best-of-three contest with victory-independent biases chosen to maximize TE.

Victory-independent optimal contest: analytical characterization. We proceed by reformulating the problem in terms of $p_X^{(0,0)}, p_X^{(1,0)}, p_X^{(0,1)}$, and $p_X^{(1,1)}$ as done in Section 4. Since $\alpha_{(0,1)} = \alpha_{(1,0)}$, in contrast to Section 4, there is a constraint that $p_X^{(0,1)}$ and $p_X^{(1,0)}$ in (11) and (12) must satisfy, as described in the following lemma.

Lemma 2 *Consider* $p_X^{(0,1)}$ *and* $p_X^{(1,0)}$ *in* (11) *and* (12). *If* $\alpha_{(0,1)} = \alpha_{(1,0)}$, *then*

$$p_X^{(1,0)} = p_X^{(0,1)} \frac{\left(1 + p_X^{(1,1)}\right)\left(1 + p_Y^{(1,1)}\right)}{2p_X^{(0,1)} + p_X^{(1,1)} p_Y^{(1,1)}}. \tag{16}$$

Proof. See Appendix A. ■

Therefore, the victory-independent optimal contest is the solution to problem (P) below:

$$\max \tau\left(p_X^{(0,0)}, p_X^{(1,0)}, p_X^{(0,1)}, p_X^{(1,1)}\right) \quad \text{s.t. } (16), \tag{P}$$

where the choice variables are the four probabilities $\left\{p_X^{(0,0)}, p_X^{(1,0)}, p_X^{(0,1)}, p_X^{(1,1)}\right\} \in (0,1)^4$. Naturally, if the optimal solution of (P) exists, then it must satisfy the FOCs, but as we discussed above, the FOCs are not sufficient. To make further progress toward the characterization of the optimal solution of (P), let τ^* be the value of τ at the optimal solution of (P), if it exists, and consider

$$p_X^{(0,0)} = \frac{457}{657},\ p_X^{(0,1)} = \frac{37}{451}, \text{ and } p_X^{(1,1)} = \frac{313}{730},$$

along with the resulting $p_X^{(1,0)}$ obtained from (16) and the above displayed values. Simple substitution of this feasible point into $\tau(\cdot)$ shows that

$$\tau^* > \frac{25}{76}. \tag{17}$$

In a sequence of lemmas, we use (17) to reach contradictions so as to characterize properties of the optimal solution of (P).

Lemma 3 *Problem* (P) *admits a solution. At the optimal solution of* (P), *we have*[30]

$$\frac{1}{5} < p_X^{(0,0)} < \frac{4}{5},\ \frac{1}{5} < p_X^{(1,1)} < \frac{4}{5}. \tag{18}$$

Proof. See Appendix A. ■

[30] Throughout this section, when no confusion arises, we omit the star to denote equilibrium quantities.

Lemma 4 *At the optimal solution of* (P)*, we have*

1) $0 = \partial\tau/\partial p_X^{(0,0)}$*, leading to*

$$p_X^{(0,0)} = \frac{1}{2}\frac{\left(p_Y^{(1,1)}\right)^2 p_X^{(1,0)}\left(1+p_Y^{(1,0)}\right) + \left(p_Y^{(0,1)}\right)^2 p_X^{(1,1)}\left(1+p_Y^{(1,1)}\right)}{p_X^{(1,0)}p_Y^{(1,1)}\left(1-p_X^{(1,1)}p_Y^{(1,0)}\right) + p_X^{(1,1)}p_Y^{(0,1)}\left(1-p_X^{(0,1)}p_Y^{(1,1)}\right)}. \tag{19}$$

2) $p_X^{(1,0)} \geq \frac{1}{2}\frac{(1+p_Y^{(0,0)})p_Y^{(1,1)}}{1-p_X^{(1,1)}p_Y^{(0,0)}}$.

3) $p_X^{(0,1)} \leq \frac{1}{2}\frac{(1+p_Y^{(1,1)})p_Y^{(0,0)}}{1-p_Y^{(1,1)}p_X^{(0,0)}}$.

4) $p_X^{(1,1)} - \frac{1}{2}$ *and* $\partial\tau/\partial p_X^{(1,1)}$ *have the same sign.*

Proof. See Appendix A. ■

Lemma 3 and Lemma 4 characterize $p_X^{(i,j)}$, while Lemma 5 below characterizes the connection between α_1, α_2, α_3 and $p_X^{(0,0)}$, $p_X^{(1,0)}$, $p_X^{(0,1)}$, $p_X^{(1,1)}$ at the optimal solution of (P).

Lemma 5 *Consider the optimal solution of* (P) *and the corresponding optimal values of* α_1*,* α_2*,* α_3*. We have*

1) $p_X^{(0,0)} \gtreqless \frac{1}{2} \Leftrightarrow \alpha_1 \gtreqless 1$

2) $p_X^{(1,0)} \gtreqless \frac{1+p_X^{(1,1)}}{2} \Leftrightarrow \alpha_2 \gtreqless 1$

3) $p_X^{(0,1)} \gtreqless \frac{p_X^{(1,1)}}{2} \Leftrightarrow \alpha_2 \gtreqless 1$

4) $p_X^{(1,1)} \gtreqless \frac{1}{2} \Leftrightarrow \alpha_3 \gtreqless 1$

Proof. See Appendix A. ■

We can use Lemma 3, Lemma 4, and Lemma 5 to show that the optimal solution of (P) involves alternating advantages in the first two matches, which is one of the salient features of the example $\{\alpha_1, \alpha_2, \alpha_3\} = \{2, 1/2, 1\}$ that we discussed above when illustrating the intuition.

Proposition 4 *Consider the optimal solution of* (P)*. The implied values of* α_1 *and* α_2 *are such that* $(\alpha_1 - 1)(\alpha_2 - 1) < 0$*. Hence, the first match is biased in favor of a player and the second in favor of her rival.*

Proof. See Appendix A. ■

Another salient feature of the example $\{\alpha_1, \alpha_2, \alpha_3\} = \{2, 1/2, 1\}$ that we discussed above when illustrating the intuition is that the tie-break is unbiased. This assumption was made to simplify the intuition behind Proposition 2; however, this is not true of the optimal solution of (P), as the following proposition shows.

Proposition 5 *Consider the optimal solution of* (P) *. The implied values of* α_3 *is different from* 1*. Hence, the tie-break is biased.*

Proof. See Appendix A. ■

An immediate consequence of Proposition 5 is the following:

Corollary 3 *Consider a best-of-three Tullock contest between two ex-ante symmetric players. An alternating contest is not optimal in the class of contests with victory-independent biases.*

Next, we prove that the ex-ante probabilities of victory differ across players in the optimal solution of (P).

Proposition 6 *In the victory-independent optimal contest solving* (P)*, the ex-ante probabilities of victory are different across players.*

Proof. See Appendix A. ■

Victory-independent optimal contest: numerical simulations. Propositions 4-6 analytically characterize some key properties of the victory-independent optimal contest. Numerical simulations show that the victory-independent optimal contest makes all matches biased in the following way:

$$\{\alpha_1, \alpha_2, \alpha_3\} \approx \{5.22, 0.33, 0.75\},$$

in line with the analytical findings of Proposition 4 and Proposition 5.

Thus, it is optimal to give a large (approx. 5) advantage to player X in the first match, and balance it out with a medium (approx. 3) advantage to player Y in the second match and a small (approx. 4/3) advantage to player Y in the third match, if necessary. Moreover, one can show that the optimal alternating contest already attains roughly 81% of the improvement achieved by the victory-independent optimal contest over the fully unbiased contest. Reasonably, in the optimal alternating contest $\alpha \approx 4.21$, which is in-between the first-match advantage given to X (approx. 5) and the second-match advantage given to Y (approx. 3) of the victory-independent optimal contest.

Next, we analyze the equilibrium winning probabilities in the victory-independent optimal contest. Evaluating p_X at each node, we obtain

$$p_X^{(0,0)} \approx 0.696,\ \ p_X^{(0,1)} \approx 0.082,\ p_X^{(1,0)} \approx 0.450,\ p_X^{(1,1)} \approx 0.429.$$

In $(0,0)$, on the one hand, the optimal bias yields a substantial departure from $p_X = 0.5$. On the other hand, it is less than what would happen without dynamics; in fact, if we apply the optimal α_1 to a one-shot contest, we obtain $p_X \approx 5.22/6.22 \approx 0.839$. Instead, with dynamics, one needs to account for the future advantages of Y. In $(0,1)$, it is not at all surprising to obtain a small p_X. In fact, two effects point in the same direction: Y is both one match ahead and advantaged by the bias. Instead, in $(1,0)$, while Y is still advantaged by the bias, she is lagging one match behind. The bias favoring Y more than compensates for her disadvantage of lagging one match behind — i.e., $p_X^{(1,0)} < 1/2$. In the last node $(1,1)$, the stakes are the same for X and Y. Thus, what accounts for $p_X^{(1,1)}$ being different than $1/2$ is solely the mechanical effect of $\alpha_3 \neq 1$. Combining the above findings, we calculate that the ex-ante probability of victory for player X is 0.488, in line with the analytical finding of Proposition 6.

6 Extensions

6.1 Generalized Tullock contests

In this section, we consider "generalized" Tullock contests, i.e., we replace the contest technology (1) with

$$p_X^{(i,j)}(x^{(i,j)}, y^{(i,j)}) = \begin{cases} \frac{\alpha_{(i,j)}\left(x^{(i,j)}\right)^r}{\alpha_{(i,j)}\left(x^{(i,j)}\right)^r + \left(y^{(i,j)}\right)^r} & \text{if } (x^{(i,j)}, y^{(i,j)}) \neq (0,0) \\ \frac{1}{2} & \text{if } (x^{(i,j)}, y^{(i,j)}) = (0,0)\,, \end{cases} \tag{20}$$

where $r \in (0, \infty)$ (Tullock, 1980). The contest technology (1) corresponds to $r = 1$. In what follows, we discuss separately the cases $r \in (0, 1)$, $r \in (2, \infty)$, and $r \in (1, 2]$, as the nature of equilibrium is different in each case: the equilibrium is in pure strategies for $r \in (0, 1)$, in mixed strategies for $r \in (2, \infty)$, and in semi-mixed or pure strategies for $r \in (1, 2]$ *depending* on the biases chosen by the designer. Furthermore, for $r \in (1, \infty)$, a further adjustment to (20) in case of a tie with zero efforts is needed, as we discuss below.

6.1.1 Generalized Tullock contest with $r \in (0, 1)$

In this section, we begin by asking whether Proposition 1 and Proposition 2 carry over to a contest technology (20) with $r \in (0, 1)$. With victory-dependent biases, we find that Proposition 1 carries over $\forall r \in (0, 1)$. Formally, we obtain:

Proposition 7 *Consider a best-of-three generalized Tullock contest between two ex-ante symmetric players with* $r \in (0,1)$. *With victory-dependent biases, the point* $\{\alpha_{(0,0)}, \alpha_{(1,0)}, \alpha_{(0,1)}, \alpha_{(1,1)}\} = \left\{1, \left(\frac{2-r}{2+r}\right)^r, \left(\frac{2+r}{2-r}\right)^r, 1\right\}$ *is the unique global maximum for* TE *in* $\mathbb{R}^4_{>0}$ *for every* $r \in (0,1)$. *Thus, in the victory-dependent optimal contest* $p_X^{(i,j)} = 1/2$ *for any* $i, j \in \{0,1\}$.

Proof. See Appendix B. ■

With victory-independent biases, we find that Proposition 2 carries over only if r is sufficiently close to 1. Formally, we obtain:

Proposition 8 *Consider a best-of-three generalized Tullock contest between two ex-ante symmetric players with* $r \in (0,1)$. *With victory-independent biases, there exists* $\hat{r} \in (0,1)$ *such that, for every* $r \in (\hat{r}, 1)$, *the fully unbiased contest* $\{\alpha_1, \alpha_2, \alpha_3\} = \{1,1,1\}$ *does not maximize* TE *in* $\mathbb{R}^3_{>0}$. *Therefore, if* $r \in (\hat{r}, 1)$, *then the fully unbiased contest is not the victory-independent optimal contest.*

Proof. See Appendix B. ■

Numerical results show that $\hat{r} \approx 0.826581$, and that, if $r < 0.801$, then the fully unbiased contest is the victory-independent optimal contest.[31] In fact, when r is small, the momentum/discouragement effect loses quantitative relevance, and thus dynamically leveling the playing field becomes less attractive. Therefore, an interesting similarity between our results and those in Feng and Lu (2018) arises for an r sufficiently close to 0. Since the momentum/discouragement effect is small in this case, we find no benefit from the introduction of biases. Similarly, in this case Feng and Lu (2018) find no benefit from diverting resources away from a prize for winning all matches in favor of prizes for winning individual matches: "When discriminatory power r stays low, the momentum effect is weaker; therefore, there is no need to provide battle prizes to mitigate the momentum effect for effort elicitation" (see Feng and Lu, 2018, p. 83).

In Proposition 3, we showed that moving from $\{\alpha_1, \alpha_2, \alpha_3\} = \{1,1,1\}$ in the direction implied by an alternating contest $\{\alpha_1, \alpha_2, \alpha_3\} = \{\alpha, 1/\alpha, 1\}$ increases TE for α in a neighborhood of 1. We now generalize to $r \in (0,1)$ this result.

Proposition 9 *Consider a best-of-three generalized Tullock contest between two ex-ante symmetric players and* $r \in (0,1)$. *With victory-independent biases and within the family of alternating contests* $\{\alpha_1, \alpha_2, \alpha_3\} = \{\alpha, 1/\alpha, 1\}$, *let* $TE(\alpha)$ *be* TE *as a function of* $\alpha \in (0,\infty)$. *There exists* $\tilde{r} \in (0,1)$ *such that, for every* $r \in (\tilde{r}, 1)$, *setting* $\alpha = 1$ *(the fully unbiased contest) gives a local minimum for* $TE(\alpha)$. *Therefore, if* $r \in (\tilde{r}, 1)$, *then the fully unbiased contest is not optimal even within the class of alternating contests.*

[31] When $r \in [0.801, \hat{r}]$, numerical results show that the fully unbiased contest is a local, but not global, maximum.

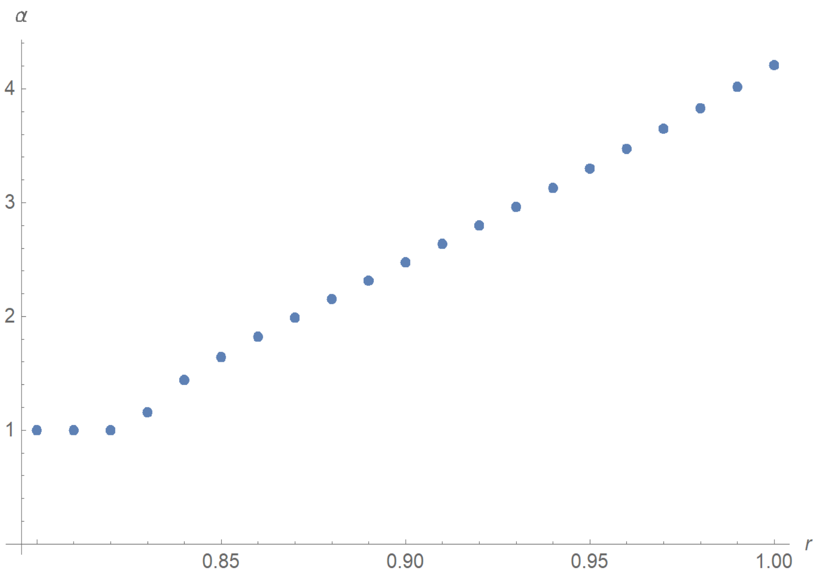

Figure 3: Optimal α within the family of alternating contests as a function of $r \in [0.8, 1]$.

Finally, within the family of alternating contests, we numerically compute the optimal α^* as a function of $r \in (0, 1]$. We find the threshold value $\tilde{r} \approx 0.828$ of Proposition 9, such that $\alpha^* = 1$ for every $r \in (0, \tilde{r})$, and for larger values of r we find that α^* is an increasing function of r, as expected. For $r = 1$, we find, as discussed in Proposition 3, that $\alpha^* \approx 4.21$. We plot α^* as a function of r in Figure 3, where we discretize the grid of $r \in [0.8, 1]$ into 21 values, of steps 0.01 each, and for each such r we computed numerical simulations to derive the optimal α^* in an alternating contest.

6.1.2 Generalized contest with $r \in (2, \infty)$

In this section, we ask whether Proposition 1 and Proposition 2 carry over to a contest technology (20) with $r \in (2, \infty)$. In doing so, we build heavily on the elegant characterization results by Ewerhart (2017). In a nutshell, we find that when $r \in (2, \infty)$ there is a continuum of optimal biases, both for the victory-dependent and victory-independent setups. The optimal biases we characterized in Proposition 1 belongs to this continuum. And it turns out that the fully unbiased contest, which is not optimal when $r = 1$ as shown in Proposition 2, also belongs to this continuum of optimal biases.

In what follows, it will often happen as the analysis develops that, in a match, player $j \in \{X, Y\}$ fights for an effective prize $\Delta u_j = 0$ whereas her rival fights for a strictly positive effective prize. It is well known that this situation may generate issues with equilibrium existence. To sidestep this problem, we make the following key assumption that modifies the contest technology in (20) :

$$\text{If } \Delta u_i > \Delta u_j = 0 \text{ with } i, j \in \{X, Y\} \text{ and } i \neq j \text{, then } p_i(0,0) = 1. \tag{A}$$

In words, we assume that the player fighting for the strictly positive prize wins with probability 1 if both players exert 0 effort, so that both players exerting no effort is the only equilibrium. Assumption (A) is common in the literature and it is also made, for instance, in Konrad and Kovenock (2009) for an analogous reason (see the paragraph directly after their Proposition 1).[32]

As for victory-dependent biases, we find that when $r \in (2, \infty)$ there is a continuum of victory-dependent optimal biases. The vector $\{\alpha_{(0,0)}, \alpha_{(1,0)}, \alpha_{(0,1)}, \alpha_{(1,1)}\} = \{1, 1/3, 3, 1\}$ of Proposition 1 is now only one among the continuum of optimal biases. Formally, we obtain the following result.

Proposition 10 *Consider a best-of-three generalized Tullock contest between two ex-ante symmetric players with $r \in (2, \infty)$ and assumption (A). With victory-dependent biases, TE is maximized if and only if the quadruple $\{\alpha_{(0,0)}, \alpha_{(1,0)}, \alpha_{(0,1)}, \alpha_{(1,1)}\}$ satisfies the following condition:*

$$\begin{cases} \alpha_{(0,0)}^{1/r}\left(1 - \alpha_{(1,1)}^{1/r}\right) - \left(\alpha_{(0,0)}^{1/r} - 1\right)\alpha_{(1,0)}^{1/r} = 0 & \text{if } \alpha_{(1,1)} < 1, \\ \alpha_{(0,0)} = 1 & \text{if } \alpha_{(1,1)} = 1, \\ \alpha_{(0,0)}^{1/r}\left(1 - \alpha_{(1,1)}^{1/r}\right) - \left(\alpha_{(0,0)}^{1/r} - 1\right)\alpha_{(0,1)}^{1/r} = 0 & \text{if } \alpha_{(1,1)} > 1. \end{cases} \tag{21}$$

Therefore, there is a continuum of victory-dependent optimal biases. Furthermore, full-rent extraction is achieved by all victory-dependent optimal biases.

Proof. See Appendix B. ■

As for victory-independent biases, we also find that when $r \in (2, \infty)$ there is a continuum of victory-independent optimal biases. Formally, we obtain the following result.

Proposition 11 *Consider a best-of-three generalized Tullock contest between two ex-ante symmetric players with $r \in (2, \infty)$ and assumption (A). With victory-independent biases, TE is maximized if and only if the triple $\{\alpha_1, \alpha_2, \alpha_3\}$ satisfies*

$$\begin{cases} \alpha_1^{1/r}\left(1 - \alpha_3^{1/r}\right) - \left(\alpha_1^{1/r} - 1\right)\alpha_2^{1/r} = 0 & \text{if } \alpha_3 \neq 1, \\ \alpha_1 = 1 & \text{if } \alpha_3 = 1. \end{cases} \tag{22}$$

Therefore, there is a continuum of victory-independent optimal biases. Furthermore, full-rent extraction is achieved by all victory-independent optimal biases.

[32] Konrad and Kovenock (2009), in their Footnote 11, also present a limiting argument to justify assumption (A).

Proof. See Appendix B. ■

Note first that (21) and (22) are very similar as, in the optimum, TE is not affected by whether biases are victory dependent or victory independent. In fact, when α_3 or $\alpha_{(1,1)}$ is smaller than 1, then efforts in node $(0,1)$ are always 0 as the player who has an advantage in the tie-break will also be one match ahead in the second match, and hence $\alpha_{(0,1)}$ plays no role in TE by assumption (A). Similarly, when α_3 is greater than 1, efforts in $(1,0)$ are 0, and hence $\alpha_{(1,0)}$ plays no role by (A). Therefore, imposing the extra constraint of victory-independent biases ($\alpha_2 = \alpha_{(1,0)} = \alpha_{(0,1)}$) does not decrease TE. In what follows, we present the intuition behind propositions 10 and 11 by focusing on the continuum of victory-independent optimal biases with $\alpha_3 \leq 1$ characterized in (22). (The intuition for the other cases, namely victory-dependent and $\alpha_3 > 1$, is analogous.) In particular, we discuss two families of optimal biases satisfying condition (22); one with $\alpha_3^* = 1$ and one with $\alpha_2^* = 1$.

Intuition. A first family of optimal biases that condition (22) includes is $\alpha_1^* = 1$, $\alpha_2^* \in (0,\infty)$, $\alpha_3^* = 1$, and $r \in (2,\infty)$. This nests the case of $\{\alpha_1^*, \alpha_2^*, \alpha_3^*\} = \{1,1,1\}$, which we showed was *not* optimal when $r = 1$ in Proposition 2. When $\alpha_1^* = \alpha_3^* = 1$, the equilibrium payoff of the third match is 0 for both players, so that whoever is one match ahead in the second match wins for sure with no effort by (A) (as her rival is fighting for an effective prize of 0), and hence the only match where efforts are non-negative is the first one, which is unbiased. Consequently, full rent extraction is obtained, as is well-known to be the case, for instance, in a one-shot symmetric all-pay auction ($r \to \infty$). The above discussion applies to every maximizer in the family $\{1, \alpha_2^*, 1\}$ with $\alpha_2^* \in (0,\infty)$.

A second family of optimal biases that condition (22) includes is one with $\alpha_1^* \neq 1$, $\alpha_2^* = 1$, and $\alpha_3^* \neq 1$. Consider, for instance, $r = 3$ and $\{\alpha_1^*, \alpha_2^*, \alpha_3^*\} = \{8, 1, 1/8\}$, which satisfies (22). In node $(1,1)$, efforts are $1/4$ for both players and X has payoff 0, whereas Y has payoff $1/2$. In node $(1,0)$, efforts are $1/4$ for X and $1/8$ for Y, X wins with probability $3/4$ and has a payoff of $1/2$, and Y has a payoff of 0. In node $(0,1)$, both efforts are 0, Y wins for sure, and the payoff of X is 0 and that of Y is 1. In node $(0,0)$, efforts are $1/4$ for X and $1/2$ for Y, and players are equally likely to win. Hence, the full rent extraction is due to the fact that $3/4$ of total effort is exerted in node $(0,0)$ while the remaining $1/4$ necessary to reach full-rent extraction is exerted in node $(1,0)$, which is reached with probability $1/2$ and where total effort is $3/8$, and in node $(1,1)$, which is reached with probability $1/8$ and where total effort is $1/2$.

6.1.3 Generalized contest with $r \in (1,2]$

In this section, we ask whether Proposition 1 and Proposition 2 carry over to a contest technology (20) with $r \in (1,2]$. In a nutshell, we find that our two results are locally robust; namely when r is sufficiently close to 1, Proposition 1 and Proposition 2 carry over.

For $r \in (1,2]$, the equilibrium is in semi-mixed or pure strategies *depending on* the asymmetry of the match, which is affected by biases $\alpha's$, which are a choice variable for the designer. This complicates the analysis. The equilibrium characterization that follows is based on the combination of the equilibria characterizations of pure strategy (e.g., Nti, 1999) and semi-mixed strategy equilibrium (Wang, 2010). Uniqueness of the semi-mixed strategy equilibrium is proved by Feng and Lu (2017).

For $r \in (1,2]$, the pure-strategy equilibrium found by FOCs is unique. The characterization is identical to that for $r \in (0,1)$ which we developed in Lemma 6 of Appendix B. In particular, a necessary and sufficient condition for existence of a pure-strategy equilibrium is

$$\begin{cases} \left(\frac{\Delta u_X}{\Delta u_Y}\right)^r \leq \frac{1}{\alpha(r-1)} & \text{if} \quad \alpha^{1/r}\Delta u_X \geq \Delta u_Y, \\ \left(\frac{\Delta u_Y}{\Delta u_X}\right)^r \leq \frac{\alpha}{r-1} & \text{if} \quad \alpha^{1/r}\Delta u_X \geq \Delta u_Y. \end{cases}$$

Using (5), (7), and (8), this can be equivalently restated in terms of the equilibrium probability of victory p_X as

$$\begin{cases} 1 - rp_X \geq 0 & \text{if} \quad p_X \geq \frac{1}{2}, \\ 1 - r(1-p_X) \geq 0 & \text{if} \quad p_X \leq \frac{1}{2}. \end{cases}$$

In other words, taking as given that in the contest the pure-strategy equilibrium is played, a designer that chooses α to implement p_X can do so if and only if $p_X \in \left[1-\frac{1}{r}, \frac{1}{r}\right]$. If the designer wants to implement more extreme values of p_X, namely $p_X > \frac{1}{r}$ (corresponding to Case I below) or $p_X < 1-\frac{1}{r}$ (corresponding to Case II below), then players must use the semi-mixed equilibrium strategy described in Wang (2010).

Adapting Wang's notation to ours and using Wang's equilibrium characterization in Proposition 3, we obtain two cases.

- **Case I**: If $\left(\frac{\Delta u_X}{\Delta u_Y}\right)^r \geq \frac{1}{\alpha(r-1)}$, then there is a unique equilibrium which is in semi-mixed strategies. In equilibrium, X always participates with effort

$$x = \frac{\Delta u_Y}{\alpha^{\frac{1}{r}}}(r-1)^{-\frac{1}{r}}\left(1-\frac{1}{r}\right),$$

while Y stays inactive with probability $1-\tilde{p}^y$ and exerts effort $\Delta u_Y(1-1/r)$ with probability $\tilde{p}^y$, where

$$\tilde{p}^y = \frac{\Delta u_Y}{\alpha^{\frac{1}{r}}\Delta u_X}(r-1)^{-\frac{1}{r}}.$$

Hence, X wins with probability

$$\begin{aligned} p_X &= 1 - \frac{\Delta u_Y}{\alpha^{\frac{1}{r}} \Delta u_X} (r-1)^{-\frac{1}{r}} + \frac{\Delta u_Y}{\alpha^{\frac{1}{r}} \Delta u_X} (r-1)^{-\frac{1}{r}} \left(\frac{\Delta u_Y^r (r-1)^{-1} \left(1-\frac{1}{r}\right)^r}{\Delta u_Y^r (r-1)^{-1} \left(1-\frac{1}{r}\right)^r + \Delta u_Y^r \left(1-\frac{1}{r}\right)^r} \right) \\ &= 1 - \frac{\Delta u_Y}{\alpha^{\frac{1}{r}} \Delta u_X} (r-1)^{-\frac{1}{r}} \left(1-\frac{1}{r}\right), \qquad (23) \end{aligned}$$

and has an equilibrium payoff of

$$\begin{aligned} u_X &= \Delta u_X - \frac{\Delta u_Y}{\alpha^{\frac{1}{r}}} (r-1)^{-\frac{1}{r}} \left(1-\frac{1}{r}\right) - \frac{\Delta u_Y}{\alpha^{\frac{1}{r}}} (r-1)^{-\frac{1}{r}} \left(1-\frac{1}{r}\right) + u_X^L \\ &= \Delta u_X - \frac{2\Delta u_Y}{\alpha^{\frac{1}{r}}} (r-1)^{-\frac{1}{r}} \left(1-\frac{1}{r}\right) + u_X^L, \end{aligned}$$

while Y's equilibrium payoff is u_Y^L. Therefore, we can recast total expected effort and utilities in terms of p_X as follows:

$$\begin{aligned} x+y &= x + \tilde{p}^y \Delta u_Y (1-1/r) = (\Delta u_X + \Delta u_Y)(1-p_X), \qquad (24) \\ u_X &= \Delta u_X (2p_X - 1) + u_X^L, \\ u_Y &= u_Y^L. \end{aligned}$$

- **Case II**:

  If $\left(\frac{\Delta u_Y}{\Delta u_X}\right)^r \geq \frac{\alpha}{r-1}$, then there is a unique equilibrium which is in semi-mixed strategies. In equilibrium, Y always participates with effort

  $$y = \alpha^{1/r} \Delta u_X (r-1)^{-1/r} \left(1-\frac{1}{r}\right),$$

  while X stays inactive with probability $1-\tilde{p}^x$ and exerts effort $\Delta u_X \left(1-\frac{1}{r}\right)$ with probability $\tilde{p}^x$, where

  $$\tilde{p}^x = \frac{\alpha^{\frac{1}{r}} \Delta u_X}{\Delta u_Y} (r-1)^{-\frac{1}{r}}.$$

  Hence, X wins with probability

  $$p_X = \frac{\alpha^{\frac{1}{r}} \Delta u_X}{\Delta u_Y} (r-1)^{-\frac{1}{r}} \left(1-\frac{1}{r}\right), \qquad (25)$$

  and has a payoff of u_X^L, while Y has a payoff of

  $$\Delta u_Y - 2\alpha^{\frac{1}{r}} \Delta u_X (r-1)^{-\frac{1}{r}} \left(1-\frac{1}{r}\right) + u_Y^L.$$

Therefore, we can recast total expected effort and utilities in terms of p_X as follows:

$$\begin{aligned} x+y &= \tilde{p}^x x + y = (\Delta u_X + \Delta u_Y)\, p_X, & (26)\\ u_X &= u_X^L, & (27)\\ u_Y &= \Delta u_Y (1-2p_X) + u_Y^L. \end{aligned}$$

If neither Case I nor Case II apply, then the pure-strategy equilibrium (e.g., Nti, 1999) is the unique one. It is described at the beginning of Appendix B.

We first discuss the victory-dependent optimal contest with $r \in (1,2]$. We obtain the following result.

Proposition 12 *Consider a best-of-three generalized Tullock contest between two ex-ante symmetric players with $r \in (1,2]$ and assumption (A). With victory-dependent biases, there exists $\check{r} \in \left(1, 2\left(\sqrt{3}-1\right)\right]$ such that, if $r \in (1,\check{r}]$, then $p_X^{(i,j)} = 1/2$ for any $i,j \in \{0,1\}$ in the victory-dependent optimal contest. In contrast, if $r > 2\left(\sqrt{3}-1\right)$, then $p_X^{(i,j)} = 1/2$ for any $i,j \in \{0,1\}$ does* not *maximize TE.*

Proof. See Appendix B. ■

Proposition 12 shows that the result with $r = 1$ in Proposition 1 is locally robust; when $r \in (1,2]$ is sufficiently close to 1, Proposition 1 carries over. However, we do not identify the global maximum for every value of r in $(1,2]$, as for $r > \check{r}$ finding the overall maximum becomes computationally intensive. The reason is that in each of the four nodes, the designer may induce, with her choice of biases, an equilibrium that is in pure strategies, in quasi-mixed strategies in which the equilibrium payoff of Y is zero, or in quasi-mixed strategies in which the equilibrium payoff of X is zero, and hence one has to work through 81 (3^4) possible cases.

As for victory-independent biases, we also find that the result with $r = 1$ in Proposition 2 is locally robust; when $r \in (1,2]$ is sufficiently close to 1, Proposition 2 carries over.

Proposition 13 *Consider a best-of-three generalized Tullock contest between two ex-ante symmetric players with $r \in (1,2]$ and assumption (A). Let r' be the unique solution in the interval $(1,2)$ of*

$$1 + \left(\frac{2-r}{2+r}\right)^r = r,$$

which yields $r' \approx 1.1935$. With victory-independent optimal biases, if $r \leq r'$, then the fully unbiased contest $\alpha_1 = \alpha_2 = \alpha_3 = 1$ does not maximize TE in $\mathbb{R}^3_{>0}$.

Proof. See Appendix B. ■

6.2 Ex-ante asymmetric players

In the results presented so far, we have focused on establishing whether the common wisdom that unbiased static contests between ex-ante symmetric players maximize TE extends to our dynamic setup. Clearly, asymmetries between players at the outset are an important departure from this setup. But, as discussed in the Introduction, a counterpart of the common wisdom is that one should bias a static contest between asymmetric players in favor of the weaker one to ensure that each player has the same probability of victory in equilibrium (see e.g., Proposition 2 in Franke, 2012).

In this section on asymmetric players, we show several results. First, we show that how asymmetry is modeled matters. Indeed, for victory-dependent biases, we show that—depending on how asymmetry is modeled—optimal biases may or may not leave each player equally likely to win each match, in contrast with the common wisdom for static contests. Second, for alternating contests, we show that the introduction of an arbitrarily small amount of asymmetry with respect to the case of symmetric players is sufficient to break the result in Proposition 3 that there are two global maxima; instead, only one global maximum is obtained. Therefore, in contrast with what happens with symmetric players, flipping a fair coin to determine which player should first be advantaged reduces TE. Third, we provide numerical simulations that confirm, for "small" asymmetries, our finding in Proposition 3 regarding the ex-ante probability of victory: the optimal alternating contest exacerbates differences in the ex-ante probability of victory across players as compared to a designer that introduces no (further) biases.

There are two commonly adopted ways to formally analyze asymmetries in the contest literature. A **first** common way is the one employed in Lien (1990) or in Section 5.2 of Klumpp and Polborn (2006), i.e., players are asymmetric in terms of their efficiency of efforts. In our setup, this is formalized by changing the contest technology in (1) to

$$p_X^{(i,j)}\left(x^{(i,j)}, y^{(i,j)}\right) = \frac{\alpha_{(i,j)}\beta x^{(i,j)}}{\alpha_{(i,j)}\beta x^{(i,j)} + y^{(i,j)}}. \tag{28}$$

In words, the effort of player X is multiplied by a commonly-known efficiency parameter $\beta > 0$ in all matches of the contest, and ex-ante asymmetries arise if $\beta \neq 1$. Hence, $\beta = 1.1$ could be interpreted as player X being ex-ante 10% more efficient than player Y in all matches. With this form of asymmetry, according to the definition of bias we use, players have an exogenous bias to begin with. With this form of asymmetry, our main results go through with minor modifications in both victory-dependent and victory-independent setups. Indeed, if we define $\tilde{\alpha}_{(i,j)} \equiv \alpha_{(i,j)}\beta$, the problem of maximizing TE by choosing the vector of $\tilde{\alpha}'s$ is equivalent to that of choosing the vector of $\alpha's$ in the ex-ante symmetric model of Section 2. Thus, in Proposition 1, the optimal $p_X^{(i,j)}$ all remain equal to 1/2 so that $\left\{\tilde{\alpha}_{(0,0)}, \tilde{\alpha}_{(1,0)}, \tilde{\alpha}_{(0,1)}, \tilde{\alpha}_{(1,1)}\right\} = \{1, 1/3, 3, 1\}$ and $\left\{\alpha_{(0,0)}, \alpha_{(1,0)}, \alpha_{(0,1)}, \alpha_{(1,1)}\right\} =$

$\{1/\beta, 1/(3\beta), 3/\beta, 1/\beta\}$. Similarly, the characterization of the optimal solution of problem (P) in Section 5 is unchanged in terms of $p_X^{i,j}$ and simply substitutes $\tilde{\alpha}_{(i,j)}$ for $\alpha_{(i,j)}$. This is because problem (P) is itself unchanged since the constraint is unaffected by β. Proposition 2 shows that the fully unbiased contest is not optimal even when players are ex-ante symmetric, therefore it becomes less surprising to find that the fully unbiased contest is not optimal when players are ex-ante asymmetric and we do not perform a formal analysis.

Focusing now on alternating contests, an important difference with the symmetric setup is that, as soon as $\beta \neq 1$, the total effort obtained with bias α is different from that obtained with bias $1/\alpha$. This ends up implying that the introduction of a small exogenous bias $\beta \neq 1$ breaks the indifference between α^* and $1/\alpha^*$ in Proposition 3, as shown in the following result.

Proposition 14 *Consider a best-of-three Tullock contest between two players in which the contest technology in each match is* (28). *With victory-independent biases and within the family of alternating contests* $\{\alpha_1, \alpha_2, \alpha_3\} = \{\alpha, 1/\alpha, 1\}$*, let* $TE(\alpha, \beta)$ *be* TE *as a function of* $\alpha \in (0, \infty)$ *and* $\beta \in (0, \infty)$*. Note that* $TE(\alpha, 1)$ *is strictly maximized by the same* α^* *and* $1/\alpha^*$ *identified in Proposition 3 (recall* $\alpha^* \in (4.2, 4.3)$*). We obtain*

$$0 < \left.\frac{\partial TE(\alpha, \beta)}{\partial \beta}\right|_{(\alpha^*, 1)} = -\left.\frac{\partial TE(\alpha, \beta)}{\partial \beta}\right|_{(1/\alpha^*, 1)}. \tag{29}$$

Therefore, there exists $\bar{\beta} > 1$ *such that, if* $1 < \beta < \bar{\beta}$*, then* $\arg\max_\alpha TE(\alpha, \beta)$ *is unique and belongs to a neighborhood of* α^**. Similarly, there exists* $\hat{\beta} < 1$ *such that, if* $\hat{\beta} < \beta < 1$*, then* $\arg\max_\alpha TE(\alpha, \beta)$ *is unique and belongs to a neighborhood of* $1/\alpha^*$*.*

Proof. See Appendix B. ■

With the asymmetry introduced by setting $\beta \neq 1$, in line with Proposition 14, one should not expect more than one maximizer for $TE(\alpha, \beta)$, as we in fact show in the numerical results in the first two columns of Table 1. As can be seen in Table 1, it turns out that the optimal α more than compensates for the asymmetry in β. In fact, when X is stronger than Y ($\beta > 1$), in the unique optimal alternating contest, player X is *less* likely to win than player Y. Furthermore, for β in a neighborhood of 1, the ex-ante probability of victory of X is closer to 1/2 in the fully unbiased contest (i.e., $\alpha_1 = \alpha_2 = \alpha_3 = 1$) than in the optimal alternating contest, as shown in Table 1. Thus, with small asymmetries, the optimal alternating contest pushes the ex-ante probability of victory in favor of the *weak* player, but well beyond what is needed to restore a level playing field; in fact, the weak player has an equilibrium probability of winning roughly twice as large as that of the strong player.

β	α^*	Ex-ante Win Prob. of X Optimal Bias	Ex-ante Win Prob. of X Fully Unbiased Contest
1/1.1	0.219177	0.645898	0.437185
1/1.05	0.231045	0.663162	0.467731
1	$\{0.237649, 4.20789\}$	$\{0.685230, 0.31477\}$	0.5
1.05	4.32817	0.336838	0.532269
1.1	4.56252	0.354102	0.562815

Table 1. Optimal α and ex-ante probabilities of victory of player X in the optimal alternating and fully unbiased contests for different values of β.

A **second** common way to introduce asymmetries in the model is to let the two players have asymmetric valuations of victory and marginal costs; V_X, c_X and V_Y, c_Y for players X and Y, respectively (see, e.g., Franke, 2012).

With victory-dependent biases, a qualitative departure arises from our earlier results. Our earlier results with $V_X = V_Y = V$ and $c_X = c_Y = 1$ (see Proposition 1) show that the victory-dependent optimal biases induce an identical probability of victory across players in each node, including node $(1,0)$ and node $(0,1)$ where one player is one match ahead, and hence players are asymmetric. This is similar to the canonical static result where the optimal biases induce an identical ex-ante probability of victory across players regardless of the initial asymmetries between players (see e.g., Proposition 2 in Franke, 2012). However, we show below that, in the case of a best-of-three contest, the canonical static result does not extend; if $(V_X/c_X) \neq (V_Y/c_Y)$, then setting $p_X^{(0,0)} = p_X^{(1,0)} = p_X^{(0,1)} = p_X^{(1,1)} = 1/2$ is not optimal.

Proposition 15 *Consider a best-of-three Tullock contest between two ex-ante asymmetric players; player X's valuation is V_X and her marginal cost of effort is c_X, and player Y's valuation is V_Y and her marginal cost of effort is c_Y. With victory-dependent biases, if $(V_X/c_X) \neq (V_Y/c_Y)$, then TE is* not *maximized by setting $p_X^{(i,j)} = 1/2$ for any $i, j \in \{0,1\}$.*

Proof. See Appendix B. ■

Fixing $V_Y = c_X = c_Y = 1$, even if we move away from $V_X = 1$, the numerical results in Table 2 show that the victory-dependent optimal contest gives an ex-ante probability of victory closer to 1/2 than in the fully unbiased contest for victory-dependent biases (i.e., $\alpha_{(0,0)} = \alpha_{(1,0)} = \alpha_{(0,1)} = \alpha_{(1,1)} = 1$). Thus, with small asymmetries, the optimal alternating contest pushes the ex-ante probability of victory in favor of the *weak* player, but it stops short of what is needed to restore a level playing field.

V_X	Ex-ante Win Prob. of X Optimal Bias	Ex-ante Win Prob. of X Fully Unbiased Contest
1/1.1	0.494899	0.437185
1/1.05	0.497387	0.467731
1	0.5	0.5
1.05	0.502613	0.532269
1.1	0.505101	0.562815

Table 2. Ex-ante probabilities of victory of player X in the victory-dependent optimal and fully unbiased contests for different values of V_X, fixing $V_Y = c_X = c_Y = 1$.

Turning now to the optimal alternating contest discussed above, but for asymmetries modeled in values and costs (V_X/c_X and V_Y/c_Y) rather than in β, it is also the case that the total effort obtained with bias α is different from that obtained with bias $1/\alpha$. Just as we showed above for $\beta \neq 1$, with the asymmetry introduced by setting $(V_X/c_X) \neq (V_Y/c_Y)$, one should not expect more than one maximizer for $TE(\alpha)$, as we in fact find in the first two columns of Table 3, which is the analogue to Table 1 for asymmetries modeled in valuations.

As can be seen in Table 3, it turns out that the optimal α more than compensates for the asymmetry in valuations. In fact, when $(V_X/c_X) > (V_Y/c_Y)$, i.e., X is "stronger," in the optimal alternating contest, player X is *less* likely to win than player Y. Furthermore, for (V_X/c_X) and (V_Y/c_Y) sufficiently close to each other, the ex-ante probability of victory of X is closer to 1/2 in the fully unbiased contest (i.e., $\alpha_1 = \alpha_2 = \alpha_3 = 1$) than in the optimal alternating contest. The two findings discussed in this paragraph are qualitatively similar to the ones we found in the model where asymmetries are modeled through β in the contest technology (28).

V_X	α^*	Ex-ante Win Prob. of X Optimal Bias	Ex-ante Win Prob. of X Fully Unbiased Contest
1/1.1	0.219767	0.645484	0.437185
1/1.05	0.231499	0.662877	0.467731
1	$\{0.237649, 4.20789\}$	$\{0.685230, 0.31477\}$	0.5
1.05	4.31967	0.337123	0.532269
1.1	4.55027	0.354516	0.562815

Table 3. Optimal α and ex-ante probabilities of victory of player X in the optimal alternating and fully unbiased contests for different values of V_X, fixing $V_Y = c_X = c_Y = 1$.

6.3 Best-of-five

Best-of-five competitions are often observed in volleyball, squash, table-tennis for teams, some tennis tournaments (e.g., Grand Slam, Olympics final), some baseball tournaments (e.g., the Division Series of the MLB post-season round), and some basketball tournaments (e.g., the championship series of the Women's NBA). In this section, we consider how propositions 1 and 2 extend to a best-of-five, rather than best-of-three, contest. We also provide an analytical result for best-of-n contests. Extending to more than three matches substantially enlarges the strategy space of the designer.

For victory-dependent biases, rather than choosing four biases, the designer chooses nine biases: $\{\alpha_{(0,0)}, \alpha_{(1,0)}, \alpha_{(0,1)}, \alpha_{(1,1)}, \alpha_{(2,0)}, \alpha_{(2,1)}, \alpha_{(1,2)}, \alpha_{(0,2)}, \alpha_{(2,2)}\}$. Therefore, we treat the problem numerically and find that the global maximum is reached when $\alpha_{(0,0)} = \alpha_{(1,1)} = \alpha_{(2,2)} = 1$, $\alpha_{(1,0)} = \alpha_{(2,1)} = 1/3$, $\alpha_{(2,0)} = 1/9$, $\alpha_{(0,1)} = \alpha_{(1,2)} = 3$, and $\alpha_{(0,2)} = 9$.[33] This structure of biases mirrors the one obtained analytically for the best-of-three contest, where we additionally learn that, if one player is two matches ahead of the rival, it is optimal to give an advantage of 9 to the rival, so as to perfectly level the playing field. In fact, with the optimal vector of biases, both players are equally likely to win each match and thus the overall contest.

For victory-independent biases, rather than choosing three biases, the designer chooses five biases: α_1, α_2, α_3, α_4, and α_5. First, for an alternating-advantage best-of-five contest $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5\} = \{\alpha, 1/\alpha, \alpha, 1/\alpha, 1\}$, numerical simulations show that TE is maximized at $\alpha = 9.171$ and, for a victory-independent optimal best-of-five contest, numerical simulations show that $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5\} \approx \{0.057, 11.374, 0.204, 3.128, 1.414\}$ maximizes TE. Second, besides the numerical simulations, we can also analytically confirm that a fully unbiased contest is not optimal by fixing $\alpha_3 = \alpha_4 = \alpha_5 = 1$ and asking what vector $\{\alpha_1, \alpha_2\}$ maximizes TE. We proceed in two steps. First, repeatedly using (5)-(8), we compute

$$\begin{aligned}
u_X^{(2,0)} &= \frac{30\,073}{30\,976}V, \quad u_Y^{(2,0)} = \frac{1}{30\,976}V \\
u_X^{(1,1)} &= \frac{23}{128}V, \quad u_Y^{(1,1)} = \frac{23}{128}V, \\
u_X^{(0,2)} &= \frac{1}{30\,976}V, \quad u_Y^{(0,2)} = \frac{30\,073}{30\,976}V.
\end{aligned} \tag{30}$$

Second, we derive the following result for a best-of-n Tullock contest.

Proposition 16 *Consider a best-of-n Tullock contest between two ex-ante symmetric players. With victory-independent biases, let $\alpha_3 = ... = \alpha_n = 1$ and calculate $u_X^{(2,0)}, u_Y^{(2,0)}, u_X^{(1,1)}$, and $u_Y^{(1,1)}$. If*

[33] We do not treat this problem analytically, although the numerical simulations detect an intuitive vector of optimal biases. This is because of the well-known complexities already noted by Klumpp and Polborn (2006; p. 1084) for Tullock contests: "a general closed form solution for the SGPE strategies and payoffs is desirable [..]. Unfortunately, such a solution is difficult to obtain for all but a very small number of districts."

$u_X^{(2,0)} + u_Y^{(2,0)} \geq 2\left(u_X^{(1,1)} + u_Y^{(1,1)}\right)$, *then the fully unbiased contest (i.e.,* $\alpha_1 = ... = \alpha_n = 1$*) does not maximize* TE.

Proof. See Appendix B. ■

The inequality in Proposition 16 is satisfied in a best-of-five contest, as can be seen in (30). Therefore, we conclude the following.

Corollary 4 *In a best-of-five Tullock contest, the fully unbiased contest* $\alpha_1 = ... = \alpha_5 = 1$ *does not maximize* TE.

The result of Proposition 16 is in line with the result of Proposition 2. In fact, we conjecture that the non-optimality of an unbiased contest between ex-ante symmetric players carries over to best-of-n contests, since the same logic we explored appears to hold when comparing the vector of victory-independent biases $\{\alpha, 1/\alpha, 1, ..., 1\}$ and $\{1, 1, 1, ..., 1\}$ because of two effects; the symmetry of the continuation values from the third match onwards, and the higher probability of reaching nodes where players have the same number of victories. In fact, in all numerical simulations in Table 4, the hypothesis $u_X^{(2,0)} + u_Y^{(2,0)} \geq 2\left(u_X^{(1,1)} + u_Y^{(1,1)}\right)$ of the best-of-n result of Proposition 16 is satisfied for n greater than 5.[34]

n	7	9	..	19	..	99
$u_X^{(2,0)} + u_Y^{(2,0)} = u_X^{(0,2)} + u_Y^{(0,2)}$	0.98795	0.98824	..	0.98767	..	0.98767
$u_X^{(1,1)} + u_Y^{(1,1)}$	0.36184	0.36882	..	0.36769	..	0.36769

Table 4. Equilibrium utilities in a best-of-n Tullock contest with $\alpha_3 = ... = \alpha_n = 1$.

However, the structure of the victory-independent optimal contest in a best-of-n contest is not a priori clear, and thus is left to future research (see Footnote 33).

6.4 Winner's effort maximization

The interest in the total effort or the expected winner's effort crucially depends on the specific application one has in mind. In sport contests, where the audience might find a lackluster performance of the teams disappointing, total effort maximization is a suitable objective. In contrast, in sport contests where the organizer cares about having the world record broken, expected winner's effort maximization is a suitable objective (see also Serena, 2017).

[34] The limit value for $n \to \infty$ in Table 4 is consistent with Table 1 in Klumpp and Polborn (2006), where v_J^{seq} as $J \to \infty$ is approximately 0.184, because $u_X^{(1,1)} = u_Y^{(1,1)}$.

The expected winner's effort (WE) is defined as follows

$$
\begin{aligned}
WE \equiv \quad & p_X^{(0,0)}\left(p_X^{(1,0)} \cdot \left(x^{(0,0)}+x^{(1,0)}\right)+p_Y^{(1,0)} p_X^{(1,1)} \cdot \left(x^{(0,0)}+x^{(1,0)}+x^{(1,1)}\right)\right) \\
& +p_Y^{(0,0)} p_X^{(0,1)} p_X^{(1,1)} \cdot \left(x^{(0,0)}+x^{(0,1)}+x^{(1,1)}\right) \\
& +p_Y^{(0,0)}\left(p_Y^{(0,1)} \cdot \left(y^{(0,0)}+y^{(0,1)}\right)+p_X^{(0,1)} p_Y^{(1,1)} \cdot \left(y^{(0,0)}+y^{(0,1)}+y^{(1,1)}\right)\right) \\
& +p_X^{(0,0)} p_Y^{(1,0)} p_Y^{(1,1)} \cdot \left(y^{(0,0)}+y^{(1,0)}+y^{(1,1)}\right)
\end{aligned} \tag{31}
$$

The top two lines of WE capture player X's overall effort considering all instances when she wins, and the bottom two lines do the same for player Y.

We first discuss the WE-maximizing victory-dependent biases. Recall that the TE-maximizing victory-dependent biases are $\{\alpha_{(0,0)}, \alpha_{(1,0)}, \alpha_{(0,1)}, \alpha_{(1,1)}\} = \{1, 1/3, 3, 1\}$. Numerical simulations show that the WE-maximizing victory-dependent biases are

$$\{\alpha_{(0,0)}, \alpha_{(1,0)}, \alpha_{(0,1)}, \alpha_{(1,1)}\} \approx \{1, 0.424, 2.358, 1\},$$

where $\alpha_{(1,0)} = 1/\alpha_{(0,1)}$, and hence players are ex-ante equally likely to win, as in the TE-maximizing vector of biases. However, in contrast with the TE-maximizing vector of biases, now, in the second match, the leader is left with an equilibrium probability of winning the second match greater than 1/2 (approximately 0.56). This is because, in the second match, the leader exerts much greater effort than the laggard (approximately 3 times greater), and as WE only values the effort of the player who ends up winning the contest, increasing the win probability of the player who exerts greater effort is beneficial to WE. This explains why $\alpha_{(1,0)} \approx 0.424$, closer to 1 than the TE-maximizing $\alpha_{(1,0)} = 1/3$, which would perfectly level the playing field.

Next, we discuss the WE-maximizing victory-independent biases. Recall that the TE-maximizing victory-independent biases are $\{\alpha_1, \alpha_2, \alpha_3\} \approx \{5.22, 0.33, 0.75\}$. Numerical simulations show that the WE-maximizing victory-independent biases are

$$\{\alpha_1, \alpha_2, \alpha_3\} \approx \{2.14, 0.52, 0.89\}.$$

The qualitative structure of an alternating contest found under TE-maximization carries over to WE-maximization. However, biases are now "milder": the optimal $\alpha's$ are now closer to 1 in every match. The intuition behind their optimality is similar to that for victory-dependent biases explained above; now, with milder biases, in the second match, the most-likely leader (player X) has a probability of winning the second match of approximately 0.59, while with TE-optimal biases it was 0.45. As the leader exerts significantly more effort than the laggard, it is crucial for WE to boost the leader's probability of winning with a milder advantage given to the laggard. (Such an advantage is not too low, otherwise the effect of losing the competition would prevail and would jeopardize efforts.)

Finally, in line with the above intuition of optimal $\alpha's$ being "milder" under WE-maximization, we find analytically that, within the family of alternating contests, the WE-maximizing α^*_{WE} is closer to 1 than the TE-maximizing α^* identified in Proposition 3.

Proposition 17 *Consider a best-of-three Tullock contest between two ex-ante symmetric players. With victory-independent biases and within the family of alternating contests* $\{\alpha_1, \alpha_2, \alpha_3\} = \{\alpha, 1/\alpha, 1\}$*, let* $WE(\alpha)$ *be* WE *as a function of* $\alpha \in (0, \infty)$*. The fully unbiased contest* $\alpha = 1$ *gives a local minimum for* $WE(\alpha)$*. There is a unique* α^*_{WE} *such that* α^*_{WE} *and* $1/\alpha^*_{WE}$ *are the only two global maximizers of* $WE(\alpha)$*. Furthermore,* α^*_{WE} *is found as the unique solution larger than 1 of* $\partial WE(\alpha)/\partial\alpha = 0$*, and we find that* $\alpha^*_{WE} \in (1.9, 2)$*. Finally, the only value of* α *such that the ex-ante probabilities of victory are identical across players is* $\alpha = 1$*; thus, the* WE*-maximizing alternating contest gives different ex-ante probabilities of victory across players.*

Proof. See Appendix B. ■

7 Conclusions

We analyze the effort-maximizing biases in a best-of-three Tullock contest. The first contribution of the paper (Section 4) is to show that the common wisdom of the optimality of unbiased contest between ex-ante symmetric players extends from a static to our dynamic setup when the effort-maximizing designer *can* tailor biases to the outcome of previous matches; that is, by giving a player a different advantage or disadvantage tomorrow according to whether she wins or loses today. Specifically, we characterize the victory-dependent optimal contest and show that it eliminates the well-known "momentum effect" and leaves the two ex-ante symmetric players equally likely to win each match, and therefore the entire contest; the common wisdom of the optimality of unbiased contest between ex-ante symmetric players extends. An interesting counterpoint arises when players are ex-ante asymmetric in their valuations; we find that leaving the two ex-ante asymmetric players equally likely to win at each match is not optimal.

The second contribution of the paper (Section 5) is to show that a fully unbiased contest is not optimal when the effort-maximizing designer *cannot* tailor biases to the outcome of previous matches. We characterize the victory-independent optimal contest and show that biases favor one player in the first match and the other in the second, and the tie-break is biased.[35] At the optimum,

[35] With symmetric players, ex-ante symmetric treatment can be restored by a fair coin flip that decides who should receive the advantage, as it happens often in sports.

the two ex-ante symmetric players are not equally likely to win in equilibrium, both at each match and in the overall contest; hence, the two ex-ante symmetric players need not be treated identically, as biasing the contest between ex-ante symmetric players increases efforts. The fact that the fully unbiased contest is not optimal extends to: i) the maximization of the winner's—rather than total—effort, ii) best-of-n contests, as long as a simple sufficient condition is satisfied, and iii) to matches that are modeled as generalized Tullock contests—i.e., with discriminatory parameter $r \neq 1$—as long as r is sufficiently close to 1.

Our analytical and numerical results depict a bigger picture that could be rephrased as follows; a fully unbiased contest is not optimal in the class of alternating contests (Corollary 2), which is not optimal in the class of contests with victory-independent biases (Corollary 3), which is not optimal in the class of contests with victory-dependent biases (Corollary 1).

To the best of our knowledge, our study is the first investigation of the common wisdom of optimality of unbiased contests in a best-of-three setup, with biases that can differ across matches. Being the first investigation, the present paper leaves room for future extensions. First, the families of biases we analyze (alternating, victory-independent, and victory-dependent) are motivated by real-life applications, but they are not exhaustive; biases could depend on, rather than the outcome of matches, players' direct actions, such as efforts themselves, or transfers to the designer. Second, we assumed that the designer has full control over the size of biases; an interesting extension for applications is that of exogenous-value biases allocated by the designer to either player X or player Y.

APPENDIX

A Appendix A: Proofs of results in sections 3-5

In all proofs below, for simplicity we use this notation:

$$\begin{array}{c} A = p_X^{(0,0)}, B = p_X^{(1,0)}, C = p_X^{(0,1)},\ D = p_X^{(1,1)} \\ 1 - A = p_Y^{(0,0)}, 1 - B = p_Y^{(1,0)}, 1 - C = p_Y^{(0,1)},\ 1 - D = p_Y^{(1,1)} \end{array} \tag{NOT}$$

Proof of Lemma 1. We specialize the general analysis of Section 3, i.e. equations (3)–(8), to each node.

Node (1,1). Since (1,1) is the last match, $u_X^W = u_Y^W = V$ and $u_X^L = u_Y^L = 0$. Thus,

$$\Delta u_X^{(1,1)} = \Delta u_Y^{(1,1)} = V, \Delta u_X^{(1,1)} + \Delta u_Y^{(1,1)} = 2V,$$

and

$$D = \left(1 + \frac{1}{\alpha_{(1,1)}}\right)^{-1}, \tag{32}$$

which using (NOT) verifies (10). Finally, we have

$$u_X^{(1,1)} = VD^2, u_Y^{(1,1)} = V(1-D)^2.$$

Node (1,0). Recall that at (1,0), if player X wins the game ends, otherwise the game moves to node (1,1). Thus,

$$\Delta u_X^{(1,0)} = V - u_X^{(1,1)}, \Delta u_Y^{(1,0)} = u_Y^{(1,1)}, \Delta u_X^{(1,0)} + \Delta u_Y^{(1,0)} = 2(1-D)V,$$

and

$$B = \left(1 + \frac{1}{\alpha_{(1,0)}}\left(\frac{1-D}{1+D}\right)\right)^{-1} \tag{33}$$

which using (NOT) verifies (12). Finally, we have

$$\begin{aligned} u_X^{(1,0)} &= B^2\left(V - u_X^{(1,1)}\right) + u_X^{(1,1)} = B^2 V(1-D)(1+D) + VD^2, \\ u_Y^{(1,0)} &= (1-B)^2 u_Y^{(1,1)} = (1-B)^2 V(1-D)^2. \end{aligned}$$

Node (0,1). Proceeding as for node (1,0), we obtain

$$\Delta u_X^{(0,1)} = u_X^{(1,1)},\ \Delta u_Y^{(0,1)} = V - u_Y^{(1,1)},\ \Delta u_X^{(0,1)} + \Delta u_Y^{(0,1)} = 2DV,$$

and

$$C=\left(1+\frac{1}{\alpha_{(0,1)}}\frac{2-D}{D}\right)^{-1}, \tag{34}$$

which using (NOT) verifies (11). Finally, we have

$$\begin{aligned} u_X^{(0,1)} &= C^2u_X^{(1,1)}=C^2VD^2, \\ u_Y^{(0,1)} &= (1-C)^2\left(V-u_Y^{(1,1)}\right)+u_Y^{(1,1)}=(1-C)^2VD(2-D)+V(1-D)^2. \end{aligned}$$

Node (0,0). Proceeding as for nodes (0,1) and (1,0), we obtain

$$\Delta u_X^{(0,0)}=u_X^{(1,0)}-u_X^{(0,1)}=VB^2(1-D)(1+D)+VD^2(1-C)(1+C),$$

$$\Delta u_Y^{(0,0)}=u_Y^{(0,1)}-u_Y^{(1,0)}=V(1-C)^2D(2-D)+VB(1-D)^2(2-B),$$

and

$$\Delta u_X^{(0,0)}+\Delta u_Y^{(0,0)}=2V\left(B(1-D)+D(1-C)-D(1-D)\left(C(1-C)+B(1-B)\right)\right).$$

Therefore,

$$A=\left(1+\frac{1}{\alpha_{(0,0)}}\frac{(1-C)^2D(2-D)+B(1-D)^2(2-B)}{B^2(1-D)(1+D)+D^2(1-C)(1+C)}\right)^{-1}, \tag{35}$$

which using (NOT) verifies (13).

Moving now to total effort in equilibrium, using (9) into (2) we obtain

$$\begin{aligned} TE=&\left(\Delta u_X^{(0,0)}+\Delta u_Y^{(0,0)}\right)A(1-A)+A\left(\Delta u_X^{(1,0)}+\Delta u_Y^{(1,0)}\right)B(1-B) \\ &+(1-A)\left(\Delta u_X^{(0,1)}+\Delta u_Y^{(0,1)}\right)C(1-C)+\left(A(1-B)+(1-A)C\right)\left(\Delta u_X^{(1,1)}+\Delta u_Y^{(1,1)}\right)D(1-D), \end{aligned}$$

and substituting the values of $\Delta u_X^{(i,j)}+\Delta u_Y^{(i,j)}$ for any $i,j\in\{0,1\}$ determined above, we obtain

$$\begin{aligned} TE=&\left(2V\left(B(1-D)+D(1-C)-D(1-D)\left(C(1-C)+B(1-B)\right)\right)\right)A(1-A) \\ &+A\left(2(1-D)V\right)B(1-B)+(1-A)(2DV)C(1-C)+\left(A(1-B)+(1-A)C\right)(2V)D(1-D). \end{aligned}$$

This last displayed equation, using (NOT), verifies (14). ■

Proof of Proposition 1. Recall (NOT). By the discussion around equation (15), maximizing TE by choosing biases is equivalent to solving the "original problem":

$$\max_{A,B,C,D}\tau(A,B,C,D),$$

by choosing $\{A, B, C, D\} \in (0,1)^4$. In fact, as it can be seen in (5) and (6), the designer cannot induce, with her choice of biases, an equilibrium probability of victory in a match which is 0 or 1 because $\Delta u_X, \Delta u_Y > 0$ in every match. Since $\tau(A, B, C, D)$ is a polynomial, the "relaxed problem" in which $\{A, B, C, D\} \in [0,1]^4$ admits a solution by Weierstrass' theorem. We now show that the solution of the relaxed problem is interior, so the original and relaxed problems have the same solution.

Note that $\tau(1/2, 1/2, 1/2, 1/2) = 11/32$. We now show that the optimal $A \in (0,1)$. Indeed,

$$\tau(0, B, C, D) = DC(2 - D - C) \leq \frac{2}{3}\frac{2}{3}\left(2 - \frac{2}{3} - \frac{2}{3}\right) = \frac{8}{27} < \frac{11}{32},$$

and

$$\tau(1, B, C, D) = (1 - B)(1 - D)(B + D) \leq \left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{3}\right)\left(\frac{1}{3} + \frac{1}{3}\right) = \frac{8}{27} < \frac{11}{32}.$$

Similarly, we can establish that the optimal $D \in (0,1)$. Fix now any $(A, D) \in (0,1)^2$ and consider τ only as a function of B and C, a function we denote with $\tau^{AD}(B, C)$. This function is strictly concave; indeed, its Hessian matrix is the following negative definite matrix:

$$\begin{bmatrix} -2A(1-D)(1-(1-A)D) & 0 \\ 0 & -2(1-A)D(1-A(1-D)) \end{bmatrix}.$$

Therefore, setting the gradient of $\tau^{AD}(B, C)$ to zero, we see that $\tau^{AD}(B, C)$ is maximized at

$$\begin{cases} B = \beta(A, D) \equiv \frac{1}{2}\frac{(2-A)(1-D)}{1-D(1-A)} \in (0,1), \\ C = \gamma(A, D) \equiv \frac{1}{2}\frac{(1-A)(2-D)}{1-A(1-D)} \in (0,1). \end{cases} \tag{36}$$

Therefore, the optimal solution to the relaxed problem is interior. We now define

$$\tilde{\tau}(A, D) \equiv \tau^{AD}(\beta(A, D), \gamma(A, D)) = \tau(A, \beta(A, D), \gamma(A, D), D).$$

By concavity of τ^{AD}, we see that $\tau(A, B, C, D) \leq \tilde{\tau}(A, D)$. Straightforward algebra then shows

$$\tilde{\tau}(A, D) = \frac{1}{4}(T_1(A, D) + T_2(A, D)), \tag{37}$$

where

$$T_1(A, D) \equiv A(1-A) + D(1-D) + 1 - 2(1-A)A(1-D)D, \tag{38}$$

$$T_2(A, D) \equiv (A + D - 1)^2\left(\frac{A(1-D)}{1 - D(1-A)} + \frac{D(1-A)}{1 - A(1-D)} - 1\right). \tag{39}$$

Furthermore, we see that $T_2(A,D) \leq 0$ for any $(A,D) \in (0,1)^2$, as

$$\begin{aligned} \frac{A(1-D)}{1-D(1-A)} + \frac{D(1-A)}{1-A(1-D)} - 1 &= (A)\frac{1-D}{1-D+AD} + (1-A)\frac{D}{1-A+AD} - 1 \\ &\leq A + (1-A) - 1 \\ &= 0. \end{aligned}$$

Using the change of variable $x \equiv A(1-A)$ and $y \equiv D(1-D)$, we can rewrite the RHS of (38) as $x+y-2xy+1$, which is strictly increasing in x and y, as both $x \leq \frac{1}{4}$ and $y \leq \frac{1}{4}$. Therefore, $A = D = \frac{1}{2}$ is a strict unique maximum for $T_1(A,D)$ and, since $T_2(A,D) \leq 0$ and $T_2\left(\frac{1}{2}, \frac{1}{2}\right) = 0$, $A = D = \frac{1}{2}$ is a maximum for $T_2(A,D)$. Therefore, $A = D = \frac{1}{2}$ is the unique maximum of $\tilde{\tau}(A,D)$. Since $\beta(1/2,1/2) = \gamma(1/2,1/2) = 1/2$, there is a unique global maximum for $\tau(A,B,C,D)$ at $\{A^*,B^*,C^*,D^*\} = \{1/2,1/2,1/2,1/2\}$. Recalling (NOT), this implies $p_X^{(i,j)} = 1/2$ for any $i,j \in \{0,1\}$. Using (10)-(13), we obtain $\{\alpha_{(0,0)}, \alpha_{(1,0)}, \alpha_{(0,1)}, \alpha_{(1,1)}\} = \{1, 1/3, 3, 1\}$. ■

Proof of Proposition 2. This proposition can be analytically proved as a special case of the Proof of Proposition 8 setting $r = 1$. ■

Proof of Proposition 3. Recall (NOT). From (14) with (32)-(35) and using the definitions of $\alpha's$ in an alternating contest (i.e., $\alpha_{(0,0)} = \alpha$, $\alpha_{(1,0)} = \alpha_{(0,1)} = 1/\alpha$, and $\alpha_{(1,1)} = 1$), then we can characterize

$$TE(\alpha) = \frac{3\alpha P_1(\alpha)}{(\alpha+1)^2(\alpha+3)^2(3\alpha+1)^2\left(3\alpha^4+26\alpha^3+166\alpha^2+26\alpha+3\right)^2} + \frac{1}{2},$$

where

$$\begin{aligned} P_1(\alpha) \equiv\; & 81\alpha^{12} + 1431\alpha^{11} + 13968\alpha^{10} + 76167\alpha^9 + 284447\alpha^8 + 516514\alpha^7 \\ & + 623232\alpha^6 + 516514\alpha^5 + 284447\alpha^4 + 76167\alpha^3 + 13968\alpha^2 + 1431\alpha + 81. \end{aligned}$$

As $TE(1) = 41/64$ and $\lim_{\alpha\to\infty} TE(\alpha) = \lim_{\alpha\to 0} TE(\alpha) = 1/2$, then $TE(\alpha)$ admits a global maximum in the interval $\alpha \in (0,\infty)$. The critical points of $TE(\alpha)$ are characterized by $\frac{\partial TE(\alpha)}{\partial\alpha} = 0$, with

$$\frac{\partial TE(\alpha)}{\partial\alpha} = \frac{3(1-\alpha)P(\alpha)}{(\alpha+1)^3(\alpha+3)^3(3\alpha+1)^3\left(3\alpha^4+26\alpha^3+166\alpha^2+26\alpha+3\right)^3}, \tag{40}$$

where

$$
\begin{aligned}
P(\alpha) \equiv\; & 729\alpha^{18}+17010\alpha^{17}+181521\alpha^{16}+1007424\alpha^{15}+3158964\alpha^{14}-161928\alpha^{13}-54275324\alpha^{12} \\
& -357517888\alpha^{11}-954820258\alpha^{10}-1282837972\alpha^{9}-954820258\alpha^{8}-357517888\alpha^{7} \\
& -54275324\alpha^{6}-161928\alpha^{5}+3158964\alpha^{4}+1007424\alpha^{3}+181521\alpha^{2}+17010\alpha+729.
\end{aligned} \tag{41}
$$

As P is continuous and $P(1)<0$, then $\alpha=1$ is a local minimum for $TE(\alpha)$, since $\frac{\partial TE(\alpha)}{\partial \alpha}<0$ in a left neighborhood of $\alpha=1$ and $\frac{\partial TE(\alpha)}{\partial \alpha}>0$ in a right neighborhood of $\alpha=1$.

Furthermore, besides $\alpha=1$, $\frac{\partial TE(\alpha)}{\partial \alpha}=0$ has at most two positive roots by the Descartes' rule of signs applied to the polynomial $P(\alpha)$ in (41). As $TE(\alpha)=TE(1/\alpha)$, we can focus on $\alpha>1$. One can further show that $P(4.2)<0$ and $P(4.3)>0$, and hence α^*, which is the solution of $P(\alpha)=0$ for $\alpha>1$, satisfies $\alpha^*\in(4.2,4.3)$. Furthermore, α^* is a maximum since $\frac{\partial TE(\alpha)}{\partial \alpha}>0$ in a left neighborhood of α^* and $\frac{\partial TE(\alpha)}{\partial \alpha}<0$ in a right neighborhood of α^* (see (40)). Similar considerations yield that $1/\alpha^*$ is also a maximum. As $TE(\alpha)=TE(1/\alpha)$, both α^* and $1/\alpha^*$ are global maxima, and there cannot be any other solutions by the Descartes' rule of signs.

Finally, we show that the only value of α such that, in an alternating contest, the ex-ante probabilities of victory are identical across players is $\alpha=1$. In an alternating contest, as $\alpha_3=1$ and $D=1/2$, simple algebra shows that the ex-ante probabilities of victory are identical across players if and only if

$$AB=(1-A)(1-C).$$

Furthermore, using (35), in an alternating contest we have $\alpha_1=1/\alpha_2=1/\alpha$, thus

$$A=\frac{3B^2+\left(1-C^2\right)}{3B^2+(1-C^2)+\alpha\left(3(1-C)^2+B(2-B)\right)}.$$

We then have the ex-ante probabilities of victory are identical across players if and only if

$$
\begin{aligned}
& \left(3B^2+\left(1-C^2\right)\right)B=\alpha\left(3(1-C)^2+B(2-B)\right)(1-C) \\
\iff\; & \left(\left(\frac{3\alpha}{3\alpha+1}\right)^2+\frac{2\alpha+3}{(\alpha+3)^2}\right)\frac{1}{3\alpha+1}=\frac{1}{\alpha+3}\left(\left(\frac{3}{\alpha+3}\right)^2+\frac{(3\alpha+2)}{(3\alpha+1)^2}\alpha\right) \\
\iff\; & \frac{9\alpha^2(\alpha+3)^2+(2\alpha+3)(3\alpha+1)^2}{3\alpha+1}=\frac{9(3\alpha+1)^2+(3\alpha+2)(\alpha+3)^2\alpha}{\alpha+3} \\
\iff\; & 4\alpha(\alpha-1)\left(9(\alpha-1)^2+16\alpha\right)=0 \\
\iff\; & \alpha=1.
\end{aligned}
$$

■

Proof of Lemma 2 . Recalling (NOT), $p_X^{(0,1)}$ and $p_X^{(1,0)}$ in (11) and (12) are B and C described in (33) and (34). If $\alpha_{(0,1)} = \alpha_{(1,0)} = \alpha_2$, (33) and (34) read as

$$B = \frac{1}{1 + \frac{1}{\alpha_2}\frac{1-D}{1+D}}, \; C = \frac{1}{1 + \frac{1}{\alpha_2}\frac{2-D}{D}}.$$

Solving these two displayed equalities for α_2 and matching solutions gives

$$B = C\frac{(1+D)(2-D)}{2C - D(1-D)}. \tag{42}$$

Using (NOT), (42) establishes (16). ■

Proof of Lemma 3. To see that problem (P) admits a solution, note that the "relaxed problem" where $\left\{p_X^{(0,0)}, p_X^{(1,0)}, p_X^{(0,1)}, p_X^{(1,1)}\right\} \in [0,1]^4$ admits a solution by Weierstrass' theorem. Therefore, if the solution of the relaxed problem is interior, (P) and the relaxed problem have the same solution. Recall the notation in (NOT). Note that the optimal $A \in (0,1)$. Indeed,

$$\tau(0,B,C,D) = DC(2-D-C) \leq \frac{2}{3}\frac{2}{3}\left(2-\frac{2}{3}-\frac{2}{3}\right) = \frac{8}{27} < \frac{25}{76} < \tau^*,$$

and

$$\tau(1,B,C,D) = (1-B)(1-D)(B+D) \leq \left(1-\frac{1}{3}\right)\left(1-\frac{1}{3}\right)\left(\frac{1}{3}+\frac{1}{3}\right) = \frac{8}{27} < \frac{25}{76} < \tau^*.$$

Similarly, we can establish that the optimal $D \in (0,1)$. Fix now any $(A,D) \in (0,1)^2$. We now show that $(B,C) \in (0,1)^2$. First, note that by the constraint of problem (P), which by (NOT) is (16), $B = 1 \Leftrightarrow C = 1$ and $B = 0 \Leftrightarrow C = 0$. Second, $\tau(A,1,1,D) = (1-A)(1-D)(A+D)$, which is maximized by $A = D = \frac{1}{3}$ for a value of $\left(\frac{2}{3}\right)^3$. Third, $\tau(A,0,0,D) = AD(2-A-D)$, which is maximized by $A = D = \frac{2}{3}$ for a value of $\left(\frac{2}{3}\right)^3$. Therefore, since $\tau^* > \frac{25}{76} > \left(\frac{2}{3}\right)^3$ (see (17)), the optimal solution of the relaxed problem is in the interior of $[0,1]^4$, so a solution to (P) exists.

From here, we show that $(A,D) \in (1/5, 4/5)^2$. Recall the quantities defined in the Proof of Proposition 1, and in particular $\tilde{\tau}(A,D)$, $T_1(A,D)$ and $T_2(A,D)$, as defined in (37), (38) and (39), respectively. Then, consider a subset S of $[0,1]^2$. As $\tau(A,B,C,D) \leq \tilde{\tau}(A,D)$, if we show that for any $(A,D) \in S$ we have $\tilde{\tau}(A,D) \leq \frac{25}{76}$, then by (17) the optimal solution of (P) must be such that $(A,D) \notin S$.

We now show that, if $A \geq \frac{4}{5}$ and $A + D \geq 1$, then $\tilde{\tau}(A,D) \leq \frac{25}{76}$. First, note that by $A \geq \frac{4}{5} > \frac{1}{2}$, we have $\frac{\partial T_1(A,D)}{\partial A} < 0$. Second, we now establish that

$$\begin{aligned} \frac{\partial T_2(A,D)}{\partial A} &= 2\frac{T_2(A,D)}{A+D-1} + (A+D-1)^2 \frac{\partial}{\partial A}\left(\frac{A(1-D)}{1-D(1-A)} + \frac{D(1-A)}{1-A(1-D)} - 1\right) \\ &< 0. \end{aligned}$$

Recall that $A+D \geq 1$ and, as established in the Proof of Proposition 1, $T_2(A,D) \leq 0$. If we also have $D > \frac{1}{2}$, we can write

$$\begin{aligned}
\frac{\partial}{\partial A}\left(\frac{A(1-D)}{1-D(1-A)}+\frac{D(1-A)}{1-A(1-D)}-1\right) &= -\frac{\left(1-A(1-A)-(A-D)^2\right)\left(A\left(D^2+(1-D)^2\right)-(1-D)^2\right)}{(1-D(1-A))^2(1-A(1-D))^2} \\
&\leq -\frac{\left(1-\frac{4}{25}-\left(1-\frac{1}{2}\right)^2\right)\left(\frac{4}{5}\left(D^2+(1-D)^2\right)-(1-D)^2\right)}{(1-D(1-A))^2(1-A(1-D))^2} \\
&= -\frac{\frac{59}{100}\cdot\frac{1}{5}(1+D)(3D-1)}{(1-D(1-A))^2(1-A(1-D))^2} \\
&< 0,
\end{aligned}$$

therefore we obtain $\frac{\partial T_2(A,D)}{\partial A} < 0$. To analyze the region where $1-A \leq D \leq \frac{1}{2}$, one can calculate that

$$\frac{\partial T_2(A,D)}{\partial A} = -(A+D-1)\frac{w(A,D)}{(1-D(1-A))^2(1-A(1-D))^2},$$

where

$$\begin{aligned}
w(A,D) \equiv\; & 3(1-A)^3 - D\left(9-16A+2A^2\right)(1-A)^2 + 2D^2(1-A)\left(4-18A+19A^2-4A^3\right) \\
& +2D^3A(2-A)\left(6A^2-10A+5\right) - D^4\left(12A^3-10A^2-A-6A^4+3\right) - D^5(2A-1).
\end{aligned}$$

We now show $w(A,D) > 0$. Note that

$$\frac{\partial^5 w(A,D)}{\partial D^5} = -(2A-1) < 0,$$

therefore $\frac{\partial^3 w(A,D)}{\partial D^3}$ is concave in D for fixed A. As

$$\left.\frac{\partial^3 w(A,D)}{\partial D^3}\right|_{D=1-A} = 12(2A-1)\left(1+4A(1-A)-6A^2(1-A)^2\right) > 0,$$

and

$$\left.\frac{\partial^3 w(A,D)}{\partial D^3}\right|_{D=\frac{1}{2}} = 3(2A-1)(7-20A(1-A)) > 0,$$

then $\frac{\partial^3 w(A,D)}{\partial D^3} > 0$ for any (A,D) with $A \geq \frac{4}{5}$ and $1-A \leq D \leq \frac{1}{2}$. So $\frac{\partial^2 w(A,D)}{\partial D^2}$ is increasing in D for fixed A. Since

$$\left.\frac{\partial^2 w(A,D)}{\partial D^2}\right|_{D=1-A} = 8A(1-A)\left(2-2A(1-A)-9(A(1-A))^2\right) > 0,$$

then $\frac{\partial^2 w(A,D)}{\partial D^2} > 0$ for any (A, D) with $A \geq \frac{4}{5}$ and $1 - A \leq D \leq \frac{1}{2}$. So $\frac{\partial w(A,D)}{\partial D}$ is increasing in D for fixed A. Now note that

$$\left.\frac{\partial w\left(A,D\right)}{\partial D}\right|_{D=1-A} = 3(1-A)^2 A^2 \left(2A-1\right)\left(5+4A\left(1-A\right)\right) > 0,$$

therefore $\frac{\partial w(A,D)}{\partial D} > 0$ for any (A, D) with $A \geq \frac{4}{5}$ and $1 - A \leq D \leq \frac{1}{2}$. So $w\left(A,D\right)$ is increasing in D for fixed A. Since

$$w\left(A, 1-A\right) = 6A^3\left(A+1\right)\left(2-A\right)\left(1-A\right)^3 > 0,$$

we have established that $w\left(A,D\right) > 0$ for any (A, D) with $A \geq \frac{4}{5}$ and $1 - A \leq D \leq \frac{1}{2}$, and therefore also in this region we have $\frac{\partial T_2(A,D)}{\partial A} < 0$, thus concluding the proof that if $A \geq \frac{4}{5}$ and $A + D \geq 1$, then $\tilde{\tau}\left(A,D\right)$ is decreasing in A.

Now note that, if $A \geq \frac{4}{5}$ and $\frac{1}{5} \leq D \leq 1$, we have

$$\tilde{\tau}\left(A,D\right) \leq \tilde{\tau}\left(\frac{4}{5}, D\right) - \frac{25}{76} = -\frac{389 - 361D - 4476D^2 + 5909D^3 + 4883D^4}{1900(5-D)(1+4D)},$$

and one can show the numerator of the fraction displayed above is strictly positive, so $\tilde{\tau}\left(\frac{4}{5}, D\right) < \frac{25}{76}$ for any $D \geq \frac{1}{5}$. Finally, we show that for $1 - A \leq D < \frac{1}{5}$, we also obtain $\tilde{\tau}\left(A,D\right) < \frac{25}{76}$. Indeed, we have

$$\begin{aligned} \tilde{\tau}\left(A,D\right) &\leq \tilde{\tau}\left(1-D,D\right) && \text{(by } \partial T_1/\partial A < 0\text{)} \\ &= \tfrac{1}{4} T_1\left(1-D,D\right) && \text{(by } T_2\left(1-D,D\right) = 0\text{)} \\ &= \tfrac{1}{4}\left(1 + 2D\left(1-D\right) - 2\left(D\left(1-D\right)\right)^2\right), \end{aligned}$$

which is a quadratic in $D\left(1-D\right)$, strictly increasing in $D\left(1-D\right) < \frac{1}{4}$ and therefore strictly increasing in D for $D \leq \frac{1}{5}$; furthermore we have $\tilde{\tau}\left(\frac{4}{5}, \frac{1}{5}\right) = \frac{793}{500} < \frac{25}{76}$. To sum up, if $A \geq \frac{4}{5}$ and $A + D \geq 1$, then we have shown that $\tilde{\tau}\left(A,D\right) < \frac{25}{76}$.

We now extend the result to other regions of the space $[0,1]^2$ using symmetry of $\tilde{\tau}$. By $\tilde{\tau}\left(x,y\right) = \tilde{\tau}\left(y,x\right)$, the above argument also implies that if $D \geq \frac{4}{5}$ and $A + D \geq 1$, then $\tilde{\tau}\left(A,D\right) < \frac{25}{76}$. By $\tilde{\tau}\left(x,y\right) = \tilde{\tau}\left(1-x, 1-y\right)$, we obtain that if $A \leq \frac{1}{5}$ and $A + D \leq 1$, then $\tilde{\tau}\left(A,D\right) < \frac{25}{76}$. And finally by $\tilde{\tau}\left(x,y\right) = \tilde{\tau}\left(1-y, 1-x\right)$, we conclude that, if $D \leq \frac{1}{5}$ and $A + D \leq 1$, then $\tilde{\tau}\left(A,D\right) < \frac{25}{76}$. All in all, we see that if $(A, D) \in S = [0,1]^2 \setminus \left[\frac{1}{5}, \frac{4}{5}\right]^2$, then $\tilde{\tau}\left(A,D\right) < \frac{25}{76}$. Since $\tau\left(A,B,C,D\right) \leq \tilde{\tau}\left(A,D\right)$ for any (A, B, C, D), we conclude by (17) that the optimal solution to (P) must have $(A, D) \in \left(\frac{1}{5}, \frac{4}{5}\right)^2$. Recalling (NOT), this is (18). $\blacksquare$

Proof of Lemma 4. Recall (NOT). By (18), $(A, D) \in \left(\frac{1}{5}, \frac{4}{5}\right)^2$.

To see part 1), note that

$$\begin{aligned}\frac{\partial \tau}{\partial A} &= \left(D(1-D)B^2 + D^2C(1-C) + D(1-C)^2 + B(1-D)^2\right)(1-2A) \\ &\quad + (1-D)B(1-B) - DC(1-C) + (1-B-C)D(1-D),\end{aligned}$$

so $\frac{\partial \tau}{\partial A} = 0$ solved for A gives

$$A = \frac{1}{2}\frac{(1-D)^2 B(1-B) + (1-C)^2 D(1-D) + D(1-C)^2 + B(1-D)^2}{D(1-D)B^2 + D^2C(1-C) + D(1-C)^2 + B(1-D)^2}, \tag{43}$$

which using (NOT) is (19).

To see parts 2) and 3), recall the definitions of β and γ in (36) and note that, using (NOT), part 2) is $B \geq \beta(A, D)$ and part 3) is $C \leq \gamma(A, D)$. We now prove that either one of $C > \gamma(A, D)$ or $B < \beta(A, D)$ makes it impossible to solve the FOC of (P) for B and C. To do so, we first require a little more notation. Using the constraint of problem (P), which by (NOT) is (42), we see that

$$\begin{aligned} B &\gtreqless \beta(A, D) \Longleftrightarrow C \gtreqless \gamma^l \equiv \frac{1}{2}\frac{D(1-D)^2(2-A)}{D(1-D)^2 + A\left(1 + D + D^2(1-D)\right)}, \\ C &\gtreqless \gamma(A, D) \Longleftrightarrow B \gtreqless \beta^h \equiv \frac{1}{2}\frac{(2-D)^2(1+D)(1-A)}{2 - D^2 - A(D^3 - 2D^2 + 2)}. \end{aligned} \tag{44}$$

In other words, the values for β^h and γ^l are derived from $\gamma(A, D)$ and $\beta(A, D)$ using (42). Letting the Lagrangean of (P) be

$$L \equiv \tau(A, B, C, D) - \mu\left(B - C\frac{(1+D)(2-D)}{2C + D - D^2}\right), \tag{45}$$

then the FOC for B and C are

$$0 = \frac{\partial L}{\partial B} \Longleftrightarrow \frac{\partial \tau}{\partial B} - \mu = 0, \text{ and } 0 = \frac{\partial L}{\partial C} \Longleftrightarrow \frac{\partial \tau}{\partial C} + \mu\frac{(1+D)(2-D)(1-D)D}{(2C + D - D^2)^2} = 0.$$

Clearly, the above displayed equations imply that $\frac{\partial \tau}{\partial B}$ and $\frac{\partial \tau}{\partial C}$ must have different signs. Since

$$\begin{aligned} \frac{\partial \tau}{\partial C} &= \left(D^2(1-2C) - 2D(1-C)\right)A(1-A) + (1-A)D(1-2C) + (1-A)D(1-D), \\ \frac{\partial^2 \tau}{\partial C^2} &= -2D(1-A)(1-A(1-D)) < 0, \text{ and } \frac{\partial \tau}{\partial C}(A, B, \gamma(A, D), D) = 0, \end{aligned}$$

we obtain

$$C \gtreqqless \gamma\left(A,D\right) \Rightarrow \frac{\partial \tau}{\partial C} \lesseqqgtr 0. \tag{46}$$

Similarly, since

$$\begin{aligned}
\frac{\partial \tau}{\partial B} &= \left(2D\left(1-D\right)B + \left(1-D\right)^2\right)A\left(1-A\right) + A\left(1-D\right)\left(1-2B\right) - AD\left(1-D\right), \\
\frac{\partial^2 \tau}{\partial B^2} &= -2A\left(1-D\right)\left(1-D\left(1-A\right)\right) < 0, \text{ and } \frac{\partial \tau}{\partial B}\left(A,\beta\left(A,D\right),C,D\right) = 0,
\end{aligned}$$

we obtain

$$B \gtreqqless \beta\left(A,D\right) \Rightarrow \frac{\partial \tau}{\partial B} \lesseqqgtr 0. \tag{47}$$

Finally, note that we can establish $\gamma\left(A,D\right) > \gamma^l$ because it is equivalent to

$$2A\left(1-A\right)\left(1+D\left(1-D\right)\right) - D^2\left(1-D\right)^2 > 0,$$

which is true since, using $(A,D) \in \left(\frac{1}{5},\frac{4}{5}\right)^2$ from (18) and (NOT), we obtain

$$2A\left(1-A\right)\left(1+D\left(1-D\right)\right) - D^2\left(1-D\right)^2 > 2\frac{1\cdot 4}{5^2}\left(1+\frac{1\cdot 4}{5^2}\right) - \left(\frac{1}{4}\right)^2 > 0.$$

Therefore, if by contradiction we have $C > \gamma\left(A,D\right)$, then $C > \gamma^l$, and thus by (44) we have $B > \beta\left(A,D\right)$. But then (46) and (47) imply both $\frac{\partial \tau}{\partial C}$ and $\frac{\partial \tau}{\partial B}$ are negative, which is a contradiction. Therefore, at the optimal solution $C \leq \gamma\left(A,D\right)$. Similarly, if $B < \beta\left(A,D\right)$, then $C < \gamma^l < \gamma\left(A,D\right)$. But then (46) and (47) imply that both $\frac{\partial \tau}{\partial C}$ and $\frac{\partial \tau}{\partial B}$ are positive, which is a contradiction. Therefore, at the optimal solution of (P) we must have $B \geq \beta\left(A,D\right)$. Using (NOT) and (36), this establishes parts 2) and 3).

To establish part 4) of the lemma, note that the multiplier μ in (45) must be negative by (47), the previously established $\beta\left(A,D\right) \leq B$, and the first-order condition $\frac{\partial \tau}{\partial B} = \mu$. Consider now the FOC for D calculated from (45). We have

$$0 = \frac{\partial L}{\partial D} \Longleftrightarrow \frac{\partial \tau}{\partial D} = \mu \frac{\partial\left(B - C\frac{(1+D)(2-D)}{2C+D-D^2}\right)}{\partial D} = \frac{4\left(1-C\right)C}{\left(2C+D-D^2\right)^2}\mu\left(\frac{1}{2}-D\right).$$

As $\mu < 0$, $\frac{\partial \tau}{\partial D}$ and $D - \frac{1}{2}$ have the same sign. Using (NOT), this establishes part 4) of the lemma. ■

Proof of Lemma 5. To see part 1), recall that in equilibrium $p_X^{(0,0)}$ is given in (13). Using (NOT), this is (35). Recalling that $\alpha_1 = \alpha^{(1,1)}$, we obtain

$$A = \frac{1}{1 + \frac{1}{\alpha_1}\left(\frac{D(2-D)(1-C)^2 + B(2-B)(1-D)^2}{B^2(1-D^2)+D^2(1-C^2)}\right)}.$$

Matching with the optimal value of A in (43), the optimal α_1 is

$$\alpha_1 = \left(\frac{1 - D\,(2-D)\,C\,(2-C) - (1-B)^2\,(1-D)^2}{D^2\,(1-C^2) + B^2\,(1-D^2)}\right)^2. \tag{48}$$

Note that the numerator of the right-hand side of (48) is positive. Indeed, for fixed B and C, its derivative with respect to D is

$$2\left(B^2 - 2C - 2B + C^2 + 1\right)(1-D),$$

whose sign does not depend on D, so the numerator is monotone in D. Therefore, for any $D \in (0,1)$, we have

$$1 - D\,(2-D)\,C\,(2-C) - (1-B)^2\,(1-D)^2 \geq \min\left\{1 - (1-B)^2, 1 - C\,(2-C)\right\} > 0.$$

So, using (48), $\alpha_1 \geq 1$ if and only if

$$D^2\left(1-C^2\right) + D\,(2-D)\,C\,(2-C) + B^2\left(1-D^2\right) + (1-B)^2\,(1-D)^2 - 1 \leq 0. \tag{49}$$

From (43), we see that $A \geq \frac{1}{2}$ if and only if

$$(1-D)\,B\,(1-B) - DC\,(1-C) + (1-B-C)\,D\,(1-D) \geq 0.$$

Part 1) then follows because this last displayed inequality is the same as (49), as one can calculate that

$$\frac{(1-D)\,B\,(1-B) - DC\,(1-C) + (1-B-C)\,D\,(1-D)}{D^2\,(1-C^2) + D\,(2-D)\,C\,(2-C) + B^2\,(1-D^2) + (1-B)^2\,(1-D)^2 - 1} = -\frac{1}{2}. \tag{50}$$

So numerator and denominator of (50) have opposite signs for any (C,D).

To see parts 2), 3), and 4), note that they follow immediately from (NOT), (32)-(34), $\alpha_1 = \alpha^{(1,0)} = \alpha^{(0,1)}$, and $\alpha_2 = \alpha^{(1,1)}$. ■

Proof of Proposition 4. Without loss of generality, let $\alpha_1 > 1$ and suppose by contradiction that $\alpha_2 \geq 1$ at the optimal solution. By Lemma 5, we have $p_X^{(0,0)} > \frac{1}{2}$, $p_X^{(1,0)} \geq \frac{1+p_X^{(1,1)}}{2}$, and

$p_X^{(0,1)} \geq \frac{p_X^{(1,1)}}{2}$. We also recall that $p_X^{(1,1)} > \frac{1}{5}$ by (18) in Lemma 3.

We proceed with two claims, the proofs of which are immediately below the current proof. The first claim further restricts the possible values of $p_X^{(0,1)}$ and $p_X^{(1,1)}$ at the optimal solution.

Claim 1. *Suppose the optimal solution to* (P) *has* $\alpha_1 > 1$ *and* $\alpha_2 \geq 1$. *Then, the optimal solution must have* $p_X^{(0,1)} < \frac{1}{4}$ *and* $p_X^{(1,1)} < \bar{D}$ *where*

$$\bar{D} \equiv 0.458426. \tag{51}$$

Now we know that since $p_X^{(1,1)} < \frac{1}{2}$, by part 4) of Lemma 4 we must have $\partial\tau/\partial p_X^{(1,1)} < 0$, which can be rewritten as

$$p_X^{(1,1)} > \frac{1}{2} \frac{p_X^{(0,0)} \left(p_Y^{(1,0)}\right)^2 \left(1 + p_Y^{(0,0)}\right) + p_X^{(0,1)} \left(p_Y^{(0,0)}\right)^2 \left(1 + p_Y^{(0,1)}\right)}{p_X^{(0,0)} p_Y^{(1,0)} \left(1 - p_X^{(1,0)} p_Y^{(0,0)}\right) + p_X^{(0,1)} p_Y^{(0,0)} \left(1 - p_X^{(0,0)} p_Y^{(0,1)}\right)}. \tag{52}$$

But (52) is impossible as the second claim shows.

Claim 2. *Consider condition* (52) *in which the value of* $p_X^{(0,0)}$ *is given in* (19) *and the value of* $p_X^{(1,0)}$ *is given by the constraint* (16). *There is no pair* $\left(p_X^{(0,1)}, p_X^{(1,1)}\right)$ *that solves* (52) *for* $\frac{1}{5} < p_X^{(1,1)} < \bar{D}$ *and* $\frac{1}{2} p_X^{(1,1)} \leq p_X^{(0,1)} < \frac{1}{4}$. ■

Proof of Claim 1 in the Proof of Proposition 4.

Recall (NOT). An equivalent expression for the objective function in (P) is

$$\begin{aligned}
\tau(A, B, C, D) \quad = \quad & \frac{1}{2} - \frac{1}{2} D^2 C^2 - \frac{1}{2} (1 - D)^2 (1 - B)^2 - \frac{1}{2} \left(B^2 \left(1 - D^2\right) + D^2 \left(1 - C^2\right)\right) A^2 \\
& - \frac{1}{2} \left(D (2 - D) (1 - C)^2 + B (2 - B) (1 - D)^2\right) (1 - A)^2 .
\end{aligned}$$

By part 1) of Lemma 4, at the optimal solution $\frac{\partial \tau}{\partial A} = 0$, which can be rewritten as

$$\left(B^2 \left(1 - D^2\right) + D^2 \left(1 - C^2\right)\right) A = k = \left(D (2 - D) (1 - C)^2 + B (2 - B) (1 - D)^2\right) (1 - A),$$

where k simply indicates the common value of the left and right extremes of the above displayed inequality. Therefore,

$$\begin{aligned}
\tau^* \quad &= \quad \frac{1}{2} \left(1 - D^2 C^2 - (1 - D)^2 (1 - B)^2 - kA - k (1 - A)\right) \\
&= \quad \frac{1}{2} \left(1 - D^2 C^2 - (1 - D)^2 (1 - B)^2 - \left(B^2 \left(1 - D^2\right) + D^2 \left(1 - C^2\right)\right) A\right),
\end{aligned}$$

where the above is evaluated at the optimal (A,B,C,D) for (P). Hence, we obtain

$$\tau^* \leq \left(\max_{(A,B,C,D)\in[0,1]^4} \theta(A,B,C,D) \quad \text{s.t.} \quad B = C\tfrac{(1+D)(2-D)}{2C+D-D^2} \right), \tag{53}$$

where

$$\theta(A,B,C,D) \equiv \frac{1}{2}\Big(1 - D^2C^2 - (1-D)^2(1-B)^2 - \left(B^2\left(1-D^2\right) + D^2\left(1-C^2\right)\right)A\Big), \tag{54}$$

because we obtain the value τ^* if we evaluate $\theta(A,B,C,D)$ at the optimal solution for (P).

Next, we show that if $\alpha_1 > 1$ and $\alpha_2 \geq 1$, then $C < \frac{1}{4}$ at the optimal solution for (P) by showing that, if $C \geq \frac{1}{4}$, then $\theta(A,B,C,D) \leq \frac{25}{76} < \tau^*$ (recall (17)), and a contradiction to (53) arises.

It is immediate to show that $\frac{\partial}{\partial A}\theta(A,B,C,D) < 0$. Also,

$$\frac{\partial}{\partial C}\theta(A,B,C,D) = D^2C^2A - D^2C < 0.$$

We now show that if $\alpha_1 > 1$ and $\alpha_2 \geq 1$, then $\frac{\partial}{\partial B}\theta(A,B,C,D) < 0$. Indeed,

$$\frac{\partial}{\partial B}\theta(A,B,C,D) = 2(1-D)(1-D-B(1+A-D(1-A))),$$

which is negative if

$$B > \frac{1-D}{1+A-D(1-A)},$$

and this last condition is true by $\alpha_1 > 1$ (i.e., $A > \frac{1}{2}$, see part 1) of Lemma 5) and $\alpha_2 \geq 1$ (i.e., $B \geq \frac{1+D}{2}$, see part 2) of Lemma 5) since we have

$$\begin{aligned}
\frac{1-D}{1+A-D(1-A)} - B &\leq \frac{1-D}{1+A-D(1-A)} - \frac{1+D}{2} \\
&\leq \frac{1-D}{1+\frac{1}{2}-D\left(1-\frac{1}{2}\right)} - \frac{1+D}{2} \\
&= \frac{D^2-6D+1}{2(3-D)} \\
&< 0,
\end{aligned}$$

as $D \in \left(\frac{1}{5}, \frac{4}{5}\right)$ (recall (18)). By the constraint of problem (P), if C increases, then B increases. Therefore, if $C \geq \frac{1}{4}$, then

$$B \geq \frac{1}{4} \cdot \frac{(1+D)(2-D)}{2\left(\frac{1}{4}\right)+D-D^2} = \frac{1}{2}\frac{(1+D)(2-D)}{1+2D(1-D)}.$$

Thus,

$$\begin{aligned}\theta(A,B,C,D)-\frac{25}{76} &\leq \theta\left(\frac{1}{2},\frac{1}{2}\frac{(1+D)(2-D)}{1+2D(1-D)},\frac{1}{4},D\right)-\frac{25}{76}\\ &= -\frac{96-528D+1159D^2-2972D^3+7680D^4-7904D^5+2584D^6}{1216(1+2D(1-D))^2},\end{aligned}$$

and one can show this function is negative for $D\in\left(\frac{1}{5},\frac{4}{5}\right)$ (recall (18)). Therefore, if $C\geq\frac{1}{4}$, then $\theta(A,B,C,D)\leq\frac{25}{76}$, which is the desired contradiction.

We now show, if $\alpha_1>1$ and $\alpha_2\geq1$, then D cannot be larger than $\bar{D}$. Indeed, by $\alpha_2\geq1$ we know that $C\geq\frac{D}{2}$ and $B\geq\frac{1+D}{2}$ (see parts 2) and 3) of Lemma 5), so with the same procedure as above we have

$$\begin{aligned}\theta(A,B,C,D)-\frac{25}{76} &\leq \theta\left(\frac{1}{2},\frac{1+D}{2},\frac{D}{2},D\right)-\frac{25}{76}\\ &= -\frac{5-114D+304D^2-190D^3+38D^4}{304},\end{aligned}$$

and one can show this is strictly negative for $D>\bar{D}$, thus reaching the desired conclusion. ■

Proof of Claim 2 in the Proof of Proposition 4.

Recall (NOT). For ease of exposition, restate condition (52) as $\phi>0$, where

$$\phi(A,B,C,D)\equiv D-\frac{1}{2}\frac{(1-B)^2A(2-A)+C(1-A)^2(2-C)}{A(1-B)(1-B(1-A))+C(1-A)(1-A(1-C))}.$$

We begin by first considering the interval $\frac{1}{5}<D\leq\frac{1}{3}$. Note first that

$$\begin{aligned}\frac{\partial\phi}{\partial A} &= \frac{1}{2}\frac{A^2(1+B)(1-B)^3+C^3(1-A)^2(2-C)+C(1-B)(2A(1-A)(1+B(1-C))-C(1-B))}{(A(1-B)(1-B(1-A))+C(1-A)(1-A(1-C)))^2}\\ &> 0,\end{aligned}$$

since, by $C<\frac{1}{4}$ (Claim 1), $B>\frac{1}{2}$ ($\alpha_2\geq1$, Lemma 5 and $D>\frac{1}{5}$), $\frac{1}{5}<A<\frac{4}{5}$ (18), we have

$$2A(1-A)(1+B(1-C))-C(1-B)>2\frac{4\cdot1}{5^2}-\frac{1}{4}\left(1-\frac{1}{2}\right)>0.$$

Therefore, if $A > \frac{65}{100}$, then, proceeding along the lines of Claim 1, we have

$$\begin{aligned} 2\left(\theta\left(A,B,C,D\right)-\frac{25}{76}\right) &\leq 1-\frac{65}{100}\left(\left(\frac{1+D}{2}\right)^2\left(1-D^2\right)+D^2\left(1-\left(\frac{D}{2}\right)^2\right)\right) \\ &\quad -D^2\left(\frac{D}{2}\right)^2-(1-D)^2\left(1-\left(\frac{1+D}{2}\right)\right)^2-\frac{25}{38} \\ &= -\frac{3268D^2-1026D-2014D^3+266D^4+107}{1520}, \end{aligned}$$

and one can show this is negative for $D \in \left(\frac{1}{5}, \frac{1}{2}\right)$. Therefore, to avoid violating (17), at the optimal solution $A < \frac{65}{100}$.

Since $\frac{\partial \phi}{\partial A} > 0$, we have

$$\phi\left(A,B,C,D\right) < \phi\left(\frac{65}{100}, B, C, D\right) < D - \frac{1}{2}\frac{98C-49C^2-702B+351B^2+351}{49C+91C^2-351B+91B^2+260},$$

and further substituting B with the constraint of problem (P), straightforward algebra shows that

$$D - \frac{1}{2}\frac{98C-702B+351B^2-49C^2+351}{49C-351B+91B^2+91C^2+260} < 0$$

if and only if

$$\begin{aligned} &-196C^3\left(2-C\right)+28C^2\left(21C+26C^2-14\right)D \\ &+(-351+1280C-194C^2+532C^3)D^2+(1222-2488C+1252C^2-728C^3)D^3 \\ &+(-1391+1812C-1030C^2)D^4+4\left(91C^2-151C+130\right)D^5<0. \end{aligned} \tag{55}$$

We now proceed to show that condition (55) is true, and therefore $\phi < 0$ for $D \leq \frac{3}{10}$. In the following paragraph, by "derivative," we mean "partial derivative with respect to D of the left-hand side of (55)," by "increasing" and "decreasing", we mean "increasing in D" and "decreasing in D" for fixed C, and by "quasi-concave" we mean "quasi-concave in D" for fixed C.

Note that for $C \in \left(\frac{1}{10}, \frac{1}{4}\right)$, the 5^{th} derivative with respect to D is positive (recall $C \geq \frac{D}{2} \geq \frac{1}{2}\cdot\frac{1}{5}$). Therefore, the fourth derivative is increasing. Evaluating it at the upper bound $D = \frac{3}{10}$, it equals,

$$24\left(-611+906C-484C^2\right)<0.$$

Therefore, the third derivative is decreasing. Evaluating it at the upper bound $D = \frac{3}{10}$ the third derivative is

$$\frac{12}{5}(52-2143C+859C^2-1820C^3),$$

which is negative for $C \geq \frac{1}{10}$. Therefore, the second derivative is either first increasing and then decreasing, or always decreasing. Therefore, the second derivative is strictly quasi-concave. Evaluating the second derivative at the lower bound $D = \frac{1}{5}$, we obtain it is equal to

$$\frac{2}{25}(2249 + 4344C + 8478C^2 + 2380C^3) > 0$$

and at the upper bound $D = \frac{3}{10}$ it is equal to

$$\frac{6903 - 7190C + 23744C^2 - 6160C^3}{25} > 0.$$

Therefore, the second derivative is always positive. Hence, the left-hand side of (55) is convex in D for fixed C. Evaluating the left-hand side of (55) at $D = \frac{1}{5}$ we obtain

$$\frac{4}{3125}(-4940 + 26564C - 60684C^2 - 202300C^3 + 266875C^4) < 0,$$

and at $D = \frac{3}{10}$ we obtain

$$\frac{7}{50000}(-61425 + 437382C - 776532C^2 - 1338400C^3 + 2960000C^4) < 0.$$

Hence, by convexity we conclude that the left-hand side of (55) is always negative, which concludes the proof that $\phi < 0$ for $D \leq \frac{3}{10}$.

We now consider $D > \frac{3}{10}$. After tedious but straightforward algebra, one can show that condition (52) in which the value of A is given in (43) and the value of B is given by the constraint of problem

(P) is satisfied if and only if $P_L(C,D) > 0$, where

$$
\begin{aligned}
P_L(C,D) \equiv\ & -64C^{12}D^4(3-2D) - 64C^{11}D^4\left(11D - 15D^2 + 6D^3 - 6\right) \\
& +32C^{10}D^2\left(58D^3 - 17D^2 - 4D - 88D^4 + 90D^5 - 65D^6 + 18D^7 + 4\right) \\
& -64C^9D^2\left(68D^3 - 19D^2 - 122D^4 + 144D^5 - 120D^6 + 79D^7 - 38D^8 + 8D^9 + 4\right) \\
& +C^8\left(\begin{array}{c} 288D^{13} - 1712D^{12} + 4864D^{11} - 10272D^{10} + 17560D^9 - 23420D^8 + 24352D^7 \\ -19320D^6 + 10824D^5 - 2076D^4 - 1312D^3 + 96D^2 + 256D + 64 \end{array}\right) \\
& -4C^7(1-D)(2-D)\left(\begin{array}{c} 32D - 40D^2 - 304D^3 - 31D^4 + 1174D^5 - 2278D^6 + 2404D^7 \\ -2119D^8 + 1430D^9 - 784D^{10} + 304D^{11} - 100D^{12} + 24D^{13} + 16 \end{array}\right) \\
& -2C^6D(1-D)\left(\begin{array}{c} 112D + 432D^2 - 2272D^3 + 412D^4 + 8813D^5 \\ -19\,393D^6 + 23\,868D^7 - 20\,794D^8 + 14\,193D^9 - 7785D^{10} \\ +3442D^{11} - 1180D^{12} + 300D^{13} - 60D^{14} + 8D^{15} + 64 \end{array}\right) \\
& +4C^5D^2(1-D)\left(\begin{array}{c} 1676D^3 - 696D^2 - 16D - 100D^4 - 5372D^5 \\ +11\,365D^6 - 13\,588D^7 + 11\,466D^8 - 7348D^9 \\ +3717D^{10} - 1496D^{11} + 468D^{12} - 104D^{13} + 12D^{14} + 32 \end{array}\right) \\
& +C^4D^2(1-D)^2\left(\begin{array}{c} 288D + 312D^2 - 2920D^3 + 3573D^4 + 3668D^5 - 15\,342D^6 + 22\,220D^7 \\ -20\,403D^8 + 13\,768D^9 - 7032D^{10} + 2640D^{11} - 660D^{12} + 80D^{13} - 48 \end{array}\right) \\
& +4C^3D^3(1-D)^3\left(\begin{array}{c} 32D + 128D^2 - 328D^3 + 59D^4 + 791D^5 - 1490D^6 \\ +1549D^7 - 1071D^8 + 510D^9 - 152D^{10} + 20D^{11} - 16 \end{array}\right) \\
& +2C^2D^4(1-D)^4\left(\begin{array}{c} 106D^2 - 8D - 76D^3 - 205D^4 + 550D^5 \\ -602D^6 + 390D^7 - 137D^8 + 18D^9 - 8 \end{array}\right) \\
& +4CD^6(1-D)^5\left(6D + 14D^2 - 35D^3 + 33D^4 - 12D^5 + D^6 - 4\right) \\
& -D^8(2-3D)(2-D)(1-D)^6
\end{aligned}
$$

In what follows, by "derivative," we mean "partial derivative with respect to C", by "proportional" we mean "directly proportional", and by "increasing" and "decreasing", we mean "increasing in C" and "decreasing in C" for fixed D. Similarly to the proof of Proposition 6, we analyze the condition $P_L(C,D) > 0$ by focusing on the proof of the sign of bivariate polynomials (in both C and D). The more standard proofs of the sign of univariate polynomials (either in C, or in D) are omitted for brevity.

The 12^{th} derivative of P_L is proportional to $2D - 3 < 0$, therefore, the 11^{th} derivative of P_L is decreasing. Evaluating at the lower bound $C = \frac{D}{2}$ the 11^{th} derivative, we obtain it is proportional to

$$2 - 9D + 6D^2,$$

which is strictly negative for $D > \frac{3}{10}$. Therefore, the 11^{th} derivative is negative, and hence the 10^{th} derivative is decreasing, in the relevant range.

Evaluating at the upper bound $C = \frac{1}{4}$ (see Claim 1) the 10^{th} derivative, we obtain a positive polynomial in D. Hence, the 9^{th} derivative is increasing, in the relevant range. Evaluating at the upper bound $C = \frac{1}{4}$ the 9^{th} derivative, we obtain a univariate polynomial in D which is strictly negative for $D > \frac{3}{10}$. Hence the 8^{th} derivative is decreasing, in the relevant range. Evaluating at the upper bound $C = \frac{1}{4}$ the 8^{th} derivative, we obtain a positive polynomial in D, and hence the 7^{th} derivative is increasing, in the relevant range. Evaluating at the lower bound $C = \frac{D}{2}$ the 7^{th} derivative, we obtain a polynomial in D,which is strictly positive for $D > \frac{3}{10}$. Hence the 5^{th} derivative is convex.

We now skip over the 6^{th} derivative and determine the behavior of the 5^{th} derivative. Evaluating it at the upper bound $C = 1/4$, the 5^{th} derivative is a polynomial in D, which is negative by $D < \bar{D}$ (see (51)). Evaluating the 5^{th} derivative at the lower bound $C = D/2$, the 5^{th} derivative is also negative by $D < \bar{D}$ (see (51)). Therefore, the 5^{th} derivative is negative in the relevant range because it is convex and it starts and ends negative. So the 4^{th} derivative is decreasing.

Evaluating the 4^{th} derivative at the lower bound $C = \frac{D}{2}$, it is negative for $D > \frac{3}{10}$. Therefore, the 4^{th} derivative is always negative. Evaluating at the lower bound $C = \frac{D}{2}$ the 3^{rd} derivative, we obtain a negative polynomial in D. Therefore the 3^{rd} derivative is always negative. Evaluating at the lower bound $C = \frac{D}{2}$ the 2^{nd} derivative, we obtain a negative polynomial in D. Therefore the 2^{nd} derivative is always negative. Evaluating at the lower bound $C = D/2$ the 1^{st} derivative, we obtain a negative polynomial in D. Therefore the 1^{st} derivative is always negative.

Evaluating P_L at the lower bound $C = \frac{D}{2}$, we obtain it is proportional to

$$-3 + 6D + 12D^2 - 40D^3 + 30D^4 - 12D^5 + 4D^6,$$

which is strictly negative, so P_L is always negative. This concludes the proof that $\phi < 0$ for $D > \frac{3}{10}$. Hence, the proof of the claim is complete. ■

Proof of Proposition 5. Recall (NOT) and the definition of τ in (15). From part 4) of Lemma (5), $\alpha_3 = 1$, which by (32) is equivalent to $D = \frac{1}{2}$, implies $\frac{\partial \tau}{\partial D} = 0$, which gives

$$A + C - 2AB - 2AC - A^2 - C^2 + AB^2 + A^2B + AC^2 + A^2C = 0. \tag{56}$$

By $D = \frac{1}{2}$, Lemma (2) gives

$$B = \frac{9C}{1 + 8C}. \tag{57}$$

Part 1) of Lemma (5) gives

$$\begin{aligned} A &= \frac{1}{2}\frac{\left(1-\frac{1}{2}\right)^2 B\left(1-B\right)+\left(1-C\right)^2\left(\frac{1}{2}\right)\left(1-\left(\frac{1}{2}\right)\right)+\left(\frac{1}{2}\right)\left(1-C\right)^2+B\left(1-\left(\frac{1}{2}\right)\right)^2}{\left(\frac{1}{2}\right)\left(1-\left(\frac{1}{2}\right)\right)B^2+\left(\frac{1}{2}\right)^2 C\left(1-C\right)+\left(\frac{1}{2}\right)\left(1-C\right)^2+B\left(1-\left(\frac{1}{2}\right)\right)^2} \\ &= \frac{1}{2}\left(B-3C+B^2+C^2+2\right)^{-1}\left(2B-6C-B^2+3C^2+3\right). \end{aligned} \tag{58}$$

Substituting (57) and (58) into (56) gives

$$\frac{\left(2C+1\right)P(C)\left(4C-1\right)}{16\left(19C+117C^2-88C^3+32C^4+1\right)^2}=0, \tag{59}$$

where

$$\begin{aligned} P(C) \equiv\ & -3-28C+1208C^2+22\,264C^3+145\,428C^4+149\,024C^5 \\ & -797\,120C^6+420\,864C^7-53\,248C^8-131\,072C^9+65\,536C^{10}. \end{aligned}$$

One solution to (59) is $C=\frac{1}{4}$. To see that this is the only solution, note that by Lemma 3 and parts 2) and 3) of Lemma (4) we obtain

$$\begin{aligned} C &\leq \tfrac{3}{2}\tfrac{1-A}{2-A} && \text{(by part 3) of Lemma (4) and } D=\tfrac{1}{2}\text{)} \\ &< \tfrac{3}{2}\tfrac{1-\frac{1}{5}}{2-\frac{1}{5}} && \text{(by } A\in\left(\tfrac{1}{5},\tfrac{4}{5}\right)\text{ and monotonicity)} \\ &= \tfrac{2}{3} \end{aligned}$$

and

$$\begin{aligned} B &\geq \tfrac{1}{2}\tfrac{2-A}{1+A} && \text{(by part 2) of Lemma (4) and } D=\tfrac{1}{2}\text{)} \\ &\geq \tfrac{1}{2}\tfrac{2-\frac{4}{5}}{1+\frac{4}{5}} && \text{(by } A\in\left(\tfrac{1}{5},\tfrac{4}{5}\right)\text{ and monotonicity)} \\ &= \tfrac{1}{3} \end{aligned}$$

which by (57) implies

$$C\geq\frac{1}{19}.$$

One can show that $P(C)>0$ if $\frac{1}{19}\leq C\leq\frac{2}{3}$. Therefore, $C=\frac{1}{4}$ is the only solution to (59). ■

Proof of Proposition 6. Recall (NOT). Suppose by contradiction that at the optimal solution of (P) it just happens that $AB+A\left(1-B\right)D+\left(1-A\right)CD=1/2$, i.e., the ex-ante probability of victory for the two players is 1/2. By the same logic in the Proof of Claim 1, it then follows that

$$\tau^*\leq\left(\max_{(A,B,C,D)\in[0,1]^4}\theta\left(A,B,C,D\right)\quad \text{s.t.}\quad \left\{\begin{array}{l} B=C\frac{(1+D)(2-D)}{2C+D-D^2} \\ AB+A\left(1-B\right)D+\left(1-A\right)CD=\frac{1}{2} \end{array}\right.\right),$$

where $\theta(A,B,C,D)$ is defined in (54) in the Proof of Claim 1. Next, we show that

$$\left(\max_{(A,B,C,D)\in[0,1]^4}\theta(A,B,C,D) \quad \text{s.t.} \begin{cases} B = C\frac{(1+D)(2-D)}{2C+D-D^2} \\ AB + A(1-B)D + (1-A)CD = \frac{1}{2}\end{cases}\right) \le \frac{25}{76}, \tag{60}$$

thus obtaining $\tau^* \le \frac{25}{76}$, which is a contradiction to (17), under the extra constraint of identical ex-ante probabilities of victory across players. We prove (60) in the remaining of this proof. Note that we can focus on $A > 1/2$, as this simply renames who player X is. By Proposition 4 and Lemma 5, this implies $\alpha_2 < 1$, and so $C < \frac{D}{2}$.

We can solve the constraints for A and B in the maximization problem in (60) as

$$\begin{array}{l} A = \frac{(2CD-1)(2C+D(1-D))}{4C^2D-2C(2D^3-3D^2+D+2)-2(1-D)D^2} \\ B = C\frac{(1+D)(2-D)}{2C+D-D^2}, \end{array} \tag{61}$$

therefore $A > \frac{1}{2}$ is equivalent to

$$2(2CD-1)(2C+D(1-D)) < 4C^2D - 2C\left(2D^3-3D^2+D+2\right) - 2(1-D)D^2,$$

or

$$1 - D - 2C > 0.$$

From this, along with $C \le \frac{D}{2}$ and Lemma 3, we obtain the restrictions on C and D:

$$C < \min\left\{\frac{D}{2}, \frac{1-D}{2}\right\}, D \in \left(\frac{1}{5}, \frac{4}{5}\right). \tag{62}$$

Straightforward but tedious algebra shows that, using the values of A and B from (61),

$$\theta\left(\frac{(2CD-1)(2C+D(1-D))}{4C^2D-2C(2D^3-3D^2+D+2)-2(1-D)D^2}, C\frac{(1+D)(2-D)}{2C+D-D^2}, C, D\right) \le \frac{25}{76}$$

is equivalent to

$$\frac{C\cdot n_3(C,D) + l_1}{76\left(2C-D^2+D\right)^2 l_2} \ge 0, \tag{63}$$

where

$$\begin{aligned} l_1 \;\equiv\;& C^5\left(152D^2(1+(1-D)^2D)\right) \\ & +(1-2C)^2((1-8C)^2+8C^2)(1-D)^3D^4\left((5-6D)^2+2D^2+3D\right) \\ >\;& 0, \end{aligned}$$

$$
\begin{aligned}
l_2 & \equiv (1-2CD)\,C + C\,(1-D\,(1-D)\,(2D-1)) + (1-D)D^2 \\
& > \left(1-D^2\right)C + C\,(1-D\,(1-D)\,(2D-1)) + (1-D)D^2 \qquad \text{(by (62))} \\
& > 0
\end{aligned}
$$

where we used (62), and $n_3(C,D) \equiv a_3C^3 + b_3C^2 + c_3C + d_3$ with

$$
\begin{aligned}
a_3 &\equiv -4D\left(50 + D\,(1-D)\left(76+19D+1572D^2-7590D^3+12706D^4-9576D^5+2736D^6\right)\right) \\
b_3 &\equiv 2\left(24 + D\,(1-D)\left(50+85D-190D^2+4039D^3-18129D^4+30563D^5-23218D^6+6650D^7\right)\right) \\
c_3 &\equiv -D\,(1-D)\left(28+180D-581D^2+3478D^3-13517D^4+22918D^5-17670D^6+5092D^7\right) \\
d_3 &\equiv (1-D)^2D^2\left(50-179D+463D^2-1096D^3+1368D^4-570D^5\right).
\end{aligned}
$$

Thus, to establish (63), in the remaining of the proof, we prove that $n_3(C,D) > 0$. In doing so, similarly to the proof of **Claim 2**, we focus on the proof of the sign of bivariate polynomials (in both C and D). The more standard proofs of the sign of univariate polynomials (either in C, or in D) are omitted for brevity. In what follows, we first show that $n_3(C,D) > 0$ when $C \leq 1/8$ and then when $1/8 < C < 1/2$.

When $C \leq 1/8$, consider the polynomial $d_3\,(1-8C)\left(365C^2-54C+2\right)/2$ which is positive as $d_3 \geq 0$, $1-8C \geq 0$ and $365C^2-54C+2 > 0$ as its discriminant is $(54)^2 - 4\cdot 2\cdot 365 < 0$. We have

$$
n_3(C,D) - d_3\,(1-8C)\,\frac{\left(365C^2-54C+2\right)}{2} = \frac{1}{2}C \cdot n_2^{C\leq 1/8}(C,D),
$$

where $n_2^{C\leq 1/8}(C,D) \equiv a_2C^2 + b_2C + c_2$, with

$$
\begin{aligned}
a_2 &\equiv -8D(50 - D\,(1-D)\left(18174-83604D+232758D^2-561445D^3+886654D^4-697794D^5+205314D^6\right)), \\
b_2 &\equiv 96 + (D-1)D(-200+39510D-181753D^2+495518D^3-1170007D^4) \\
&\quad +(D-1)D^6(1841556-1451714D+427690D^2), \\
c_2 &\equiv 4D\,(1-D)\,(-14+785D-3717D^2+9496D^3-20524D^4+31661D^5-25080D^6+7429D^7).
\end{aligned}
$$

Therefore, if $n_2^{C\leq 1/8}(C,D) > 0$, then we have established $n_3(C,D) > 0$ and in turn (63) for $C \leq \frac{1}{8}$. The discriminant of $n_2^{C\leq 1/8}$ is $b_2^2 - 4a_2c_2$, which is a univariate polynomial in D, and one can show it is negative for $\frac{1}{5} \leq D \leq \frac{4}{5}$. Therefore, as $n_2^{C\leq 1/8}(0,D) > 0$, we proved that $n_2^{C\leq 1/8}(C,D) > 0$, thus concluding the proof for $C \leq 1/8$.

When $C > 1/8$, consider the polynomial $d_3\,(1-2C)\,(6C-1)^2$, which is positive as $d_3 \geq 0$ and $C \in [1/8, 1/2]$. We subtract such a polynomial from $n_3(C,D)$ and show that what is left is still positive. In particular, what is left is $C\left(\hat{a}_2C^2+\hat{b}_2C+\hat{c}_2\right)$ where

$$\begin{aligned}
\hat{a}_2 &\equiv -4D\left(50-(1-D)D\left(824-4141D+9984D^2-20472D^3+31646D^4-25308D^5+7524D^6\right)\right)\\
\hat{b}_2 &\equiv 2\left(24-(1-D)D\left(-50+1415D-6680D^2+15221D^3-28641D^4+43357D^5-34922D^6+10450D^7\right)\right)\\
\hat{c}_2 &\equiv D\left(1-D\right)\left(-28+520D-2625D^2+5510D^3-8309D^4+11578D^5-9462D^6+2888D^7\right).
\end{aligned}$$

As C is positive, we focus on the sign of $n_2^{C>1/8}\left(C,D\right)\equiv\hat{a}_2C^2+\hat{b}_2C+\hat{c}_2$. Note that $\hat{a}_2<0$ and hence $n_2^{C>1/8}\left(C,D\right)$ is concave in C. As $n_2^{C>1/8}\left(C,D\right)$ is also positive at all its bounds (namely, when $C=\min\left\{\frac{D}{2},\frac{1-D}{2}\right\}$ and $C=1/8$, see (62)), then $n_2^{C>1/8}\left(C,D\right)$ is positive everywhere in the parameter region of interest. Indeed, for $1/5\leq D\leq 1/2$,

$$\begin{aligned}
\frac{n_2^{C>1/8}\left(\frac{D}{2},D\right)}{D} &= -7524D^{10}+43\,282D^9-105\,214D^8+142\,747D^7-123\,494D^6\\
&\quad +77\,874D^5-40\,685D^4+17\,054D^3-4660D^2+598D-4\\
&> 0,
\end{aligned}$$

for $1/2\leq D\leq 4/5$,

$$\begin{aligned}
\frac{n_2^{C>1/8}\left(\frac{1-D}{2},D\right)}{1-D} &= 7524D^{10}-29906D^9+47302D^8-40255D^7+22154D^6\\
&\quad -9028D^5+2699D^4-319D^3-71D^2-28D+24\\
&> 0,
\end{aligned}$$

and for $1/5\leq D\leq 4/5$,

$$\begin{aligned}
16*n_2^{C>1/8}\left(\frac{1}{8},D\right) &= -11932D^9+48944D^8-80478D^7+82318D^6\\
&\quad -76112D^5+56681D^4-22905D^3+3732D^2-298D+96\\
&> 0.
\end{aligned}$$

This shows that $n_2^{C>1/8}\left(C,D\right)>0$, and thus it concludes the proof that $n_3\left(C,D\right)>0$ for any C such that $0\leq C\leq\min\left\{\frac{D}{2},\frac{1-D}{2}\right\}$ and therefore (63) holds, thus establishing that adding the constraint $AB+A\left(1-B\right)D+\left(1-A\right)CD=\frac{1}{2}$ strictly reduces the value of the objective function in (P). $\blacksquare$

B Appendix B: Proofs of results in Section 6

We begin by extending the preliminaries of Section 3 to $r \in (0,1)$ following Nti (1999). When $r \in (0,1)$, the equilibrium probabilities of victory in (5) and (6) are generalized by

$$p_X = \frac{\alpha\left(\Delta u_X\right)^r}{\alpha\left(\Delta u_X\right)^r + \left(\Delta u_Y\right)^r}, \tag{64}$$

$$p_Y = \frac{\left(\Delta u_Y\right)^r}{\alpha\left(\Delta u_X\right)^r + \left(\Delta u_Y\right)^r}, \tag{65}$$

the equilibrium payoffs (7) and (8) by

$$u_X = \Delta u_X p_X \left(1 - r p_Y\right) + u_X^L, \tag{66}$$

$$u_Y = \Delta u_Y p_Y \left(1 - r p_X\right) + u_Y^L, \tag{67}$$

and (9) by

$$x + y = r\left(\Delta u_X + \Delta u_Y\right) \cdot p_X p_Y. \tag{68}$$

With the above equations one can generalize Lemma 1 as follows:

Lemma 6 *Consider the best-of-three generalized Tullock contest between two ex-ante symmetric players described. The equilibrium probabilities of victory for X in each node are recursively determined as function of the vector of biases $\{\alpha_{(0,0)}, \alpha_{(1,0)}, \alpha_{(0,1)}, \alpha_{(1,1)}\}$ as follows:*

$$p_X^{(1,1)} = \left(1 + \frac{1}{\alpha_{(1,1)}}\right)^{-1}, \tag{69}$$

$$p_X^{(0,1)} = \left(1 + \frac{1}{\alpha_{(0,1)}}\left(\frac{1 + r p_Y^{(1,1)}}{1 - r p_Y^{(1,1)}}\right)^r\right)^{-1}, \tag{70}$$

$$p_X^{(1,0)} = \left(1 + \frac{1}{\alpha_{(1,0)}}\left(\frac{1 - r p_X^{(1,1)}}{1 + r p_X^{(1,1)}}\right)^r\right)^{-1}, \tag{71}$$

and $p_X^{(0,0)}$ solves

$$\frac{1}{p_X^{(0,0)}} = 1 + \frac{1}{\alpha_{(0,0)}}\left(\frac{p_Y^{(0,1)}\left(1 - r p_X^{(0,1)}\right) p_X^{(1,1)}\left(1 + r p_Y^{(1,1)}\right) + p_X^{(1,0)} p_Y^{(1,1)}\left(1 - r p_X^{(1,1)}\right)\left(1 + r p_Y^{(1,0)}\right)}{p_X^{(1,0)}\left(1 - r p_Y^{(1,0)}\right) p_Y^{(1,1)}\left(1 + r p_X^{(1,1)}\right) + p_X^{(1,1)} p_Y^{(0,1)}\left(1 + r p_X^{(0,1)}\right)\left(1 - r p_Y^{(1,1)}\right)}\right)^r, \tag{72}$$

where $p_Y^{(i,j)} = 1 - p_X^{(i,j)}$ for any $(i,j) \in \{0,1\}^2$.

Furthermore, in equilibrium TE in (2) satisfies

$$\begin{aligned}\frac{TE}{2rV} &= \left(p_X^{(1,0)}p_Y^{(1,1)} + p_X^{(1,1)}p_Y^{(0,1)} - r^2 p_X^{(1,1)}p_Y^{(1,1)}\left(p_X^{(0,1)}p_Y^{(0,1)} + p_X^{(1,0)}p_Y^{(1,0)}\right)\right)\cdot p_X^{(0,0)}p_Y^{(0,0)} \\ &\quad + p_X^{(0,0)}p_Y^{(1,1)}p_X^{(1,0)}p_Y^{(1,0)} + \left(1 - p_X^{(0,0)}\right)p_X^{(0,1)}p_Y^{(0,1)}p_X^{(1,1)} \\ &\quad + \left(p_X^{(0,0)}p_Y^{(1,0)} + p_Y^{(0,0)}p_X^{(0,1)}\right)p_X^{(1,1)}p_Y^{(1,1)}.\end{aligned} \tag{73}$$

Proof. Straightforward from the Proof of Lemma 1. ■

Proof of Proposition 7. This is the generalized proof of Proposition 1. For the general case of $r < 1$, the objective function must be modified with respect to (15). In particular, using (73) and letting $\tau_r \equiv \frac{TE}{2rV}$ we obtain that maximizing TE is equivalent to solving

$$\max_{A,B,C,D} \tau_r (A, B, C, D),$$

where

$$\begin{aligned}\tau_r (A, B, C, D) &= \left(B(1-D) + D(1-C) - r^2 D(1-D)(C(1-C) + B(1-B))\right) A(1-A) \\ &\quad + A(1-D) B(1-B) + (1-A) C(1-C) D + (A(1-B) + (1-A) C) D(1-D),\end{aligned} \tag{74}$$

and the choice variables are $\{A, B, C, D\} \in (0,1)^4$. In fact, as it can be seen in (64) and (65), the designer cannot induce, with her choice of biases, an equilibrium probability of victory in a match which is 0 or 1 because $\Delta u_X, \Delta u_Y > 0$ in every match. Since $\tau_r (A, B, C, D)$ is a polynomial, the "relaxed problem" in which $\{A, B, C, D\} \in [0,1]^4$ admits a solution by Weierstrass' theorem. We now show that the solution of the relaxed problem is interior, so the original and relaxed problems have the same solution.

Note also that

$$\tau_r \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) = \frac{3}{8} - \frac{r^2}{32}.$$

We show that the optimal $A \in (0,1)$. Indeed,

$$\tau_r (0, B, C, D) = DC(2 - D - C) \leq \frac{2}{3}\frac{2}{3}\left(2 - \frac{2}{3} - \frac{2}{3}\right) = \frac{8}{27} < \frac{3}{8} - \frac{r^2}{32},$$

and

$$\tau_r (1, B, C, D) = (1-B)(1-D)(B+D) \leq \left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{3}\right)\left(\frac{1}{3} + \frac{1}{3}\right) = \frac{8}{27} < \frac{3}{8} - \frac{r^2}{32}.$$

Similarly, we can establish that the optimal $D \in (0,1)$. Fix now any $(A, D) \in (0,1)^2$ and consider now τ_r only as a function of B and C, a function we denote with $\tau_r^{AD}(B,C)$. This function is strictly concave for $\{A,D\} \in (0,1)^2$. Indeed, the Hessian matrix of $\tau^{AD}(B,C)$ is the following negative definite matrix:

$$\left[\begin{array}{cc} -2A(1-D)(1-(1-A)Dr^2) & 0 \\ 0 & -2(1-A)D(1-A(1-D)r^2) \end{array}\right].$$

Therefore, setting the gradient of $\tau_r^{AD}(B,C)$ to zero, we see that $\tau_r^{AD}(B,C)$ is maximized at

$$\begin{cases} B = \beta_r\left(A,D\right) \equiv \frac{1}{2}\frac{(2-A)(1-D)+(1-A)D\left(1-r^2\right)}{1-D(1-A)r^2} \in \left(0,1\right), \\ C = \gamma_r\left(A,D\right) \equiv \frac{1}{2}\frac{(1-A)(2-D)+A(1-D)\left(1-r^2\right)}{1-A(1-D)r^2} \in \left(0,1\right). \end{cases} \tag{75}$$

Note that $\beta_r\left(A,D\right) > 0$. To see that $\beta_r\left(A,D\right) < 1$, note that this inequality reduces to $A + D - r^2 D\left(1-A\right) > 0$. As $r < 1$, this is true, as $A+D-r^2D\left(1-A\right) \geq A+D-D\left(1-A\right) = A+AD > 0$. Similar calculations apply to show that $\gamma_r\left(A,D\right) \in (0,1)$. Therefore, the optimal solution to the relaxed problem is interior.

We define

$$\tilde{\tau}_r\left(A,D\right) \equiv \tau_r^{AD}\left(\beta_r\left(A,D\right),\gamma_r\left(A,D\right)\right) = \tau_r\left(A,\beta_r\left(A,D\right),\gamma_r\left(A,D\right),D\right).$$

By concavity of τ_r^{AD}, we see that $\tau_r\left(A,B,C,D\right) \leq \tilde{\tau}_r\left(A,D\right)$. Simple algebra then shows

$$\tilde{\tau}_r\left(A,D\right) = \frac{1}{4}\left(T_{1r}\left(A,D\right) + T_{2r}\left(A,D\right)\right), \tag{76}$$

where

$$\begin{aligned} T_{1r}\left(A,D\right) &\equiv A\left(1-A\right) + D\left(1-D\right) + 1 - 2(1-A)A(1-D)Dr^2, \\ T_{2r}\left(A,D\right) &\equiv (A+D-1)^2\left(\frac{A(1-D)}{1-D\left(1-A\right)r^2} + \frac{D\left(1-A\right)}{1-A\left(1-D\right)r^2} - 1\right). \end{aligned} \tag{77}$$

Note now that $T_{2r}\left(A,D\right) \leq 0$ in any interior solution, as both $\frac{A(1-D)}{1-D(1-A)r^2}$ and $\frac{D(1-A)}{1-A(1-D)r^2}$ are increasing in r^2 and $r^2 \leq 1$, so we obtain

$$\begin{aligned} \frac{A(1-D)}{1-D\left(1-A\right)r^2} + \frac{D\left(1-A\right)}{1-A\left(1-D\right)r^2} - 1 &\leq \frac{A(1-D)}{1-D\left(1-A\right)} + \frac{D\left(1-A\right)}{1-A\left(1-D\right)} - 1 \\ &= \left(A\right)\frac{1-D}{1-D+AD} + \left(1-A\right)\frac{D}{1-A+AD} - 1 \\ &\leq A + \left(1-A\right) - 1 \\ &= 0. \end{aligned}$$

Using the change of variable $x \equiv A(1-A)$ and $y \equiv D(1-D)$, we can rewrite the RHS of (77) as $x + y - 2xyr^2 + 1$, which is strictly increasing in x and y, as both $x \leq \frac{1}{4}$ and $y \leq \frac{1}{4}$. Therefore, $A = D = \frac{1}{2}$ is a strict unique maximum for $T_{1r}(A, D)$ and, since $T_{2r}(A, D) \leq 0$ and $T_{2r}\left(\frac{1}{2}, \frac{1}{2}\right) = 0$, $A = D = \frac{1}{2}$ is a maximum for $T_{2r}(A, D)$. Therefore, $A = D = \frac{1}{2}$ is the unique maximum of $\tilde{\tau}_r(A, D)$. Since $\beta_r(1/2, 1/2) = \gamma_r(1/2, 1/2) = 1/2$, there is a unique global maximum for $\tau_r(A, B, C, D)$ at $\{A^*, B^*, C^*, D^*\} = \{1/2, 1/2, 1/2, 1/2\}$. Recalling (NOT), this implies $p_X^{(i,j)} = 1/2$ for any $i, j \in \{0, 1\}$. The values of α in the statement of this proposition then immediately obtain from $(69) - (72)$. ■

Proof of Proposition 8. We show that $\{1, 1, 1\}$ is not optimal by proving that, if we fix $\alpha_3 = 1$, and view TE as a function of α_1 and α_2, then $\alpha_1 = \alpha_2 = 1$ is a saddle point, where the definition of saddle point we adopt is the one in Simon and Blume (1994, p. 399): "A critical point x^* of F for which the Hessian $D^2F(x^*)$ is indefinite is called **saddle point** of F. A saddle point x^* is a min of F in some directions and a max of F in other directions." In particular, we do not require the directions to be orthogonal.

Recall (NOT). If we set $\alpha_3 = 1$, from (69) we obtain $D = \frac{1}{2}$, from (71) we obtain

$$B\left(\frac{2-r}{2+r}\right)^r = \alpha_2(1-B), \tag{78}$$

and from (70) we obtain

$$C\left(\frac{2+r}{2-r}\right)^r = \alpha_2(1-C). \tag{79}$$

Note that further assuming $\alpha_2 = 1$ implies $B + C = 1$. Indeed, if $\alpha_2 = 1$, then

$$B + C = \frac{1}{1 + \left(\frac{2-r}{2+r}\right)^r} + \frac{1}{1 + \left(\frac{2+r}{2-r}\right)^r} = \frac{2 + \left(\frac{2-r}{2+r}\right)^r + \left(\frac{2+r}{2-r}\right)^r}{\left(1 + \left(\frac{2-r}{2+r}\right)^r\right)\left(1 + \left(\frac{2+r}{2-r}\right)^r\right)} = 1$$

Furthermore, A then simplifies as

$$A = \left(1 + \frac{1}{\alpha_1}\right)^{-1}.$$

Note now that, by appropriately choosing α_1, any $A \in (0, 1]$ can be induced. Therefore, choosing α_1 and α_2 to maximize TE is equivalent to choosing A, B, C, and α_2 to maximize $\tau_r\left(A, B, C, \frac{1}{2}\right)$, defined in (74) subject to the constraints (78) and (79).The Lagrangean is

$$\begin{aligned} L &= \frac{1}{2}\left(B+(1-C)-r^2\frac{1}{2}(C(1-C)+B(1-B))\right)A(1-A) \\ &\quad +A\frac{1}{2}B(1-B)+(1-A)C(1-C)\frac{1}{2}+(A(1-B)+(1-A)C)\frac{1}{4} \\ &\quad -\mu_B\left[B\left(\frac{2-r}{2+r}\right)^r-\alpha_2(1-B)\right]-\mu_C\left[C\left(\frac{2+r}{2-r}\right)^r-\alpha_2(1-C)\right]. \end{aligned}$$

The FOC are

$$\begin{array}{lrl} (A) & (1-2A)\frac{1}{2}\left(B+(1-C)-r^2\frac{1}{2}(C(1-C)+B(1-B))\right) & \\ & +\frac{1}{2}B(1-B)-\frac{1}{2}C(1-C)+\frac{1}{4}(1-B-C) & =0 \\ (B) & \frac{1}{2}\left(1-r^2\frac{1}{2}(1-2B)\right)A(1-A)+\frac{A}{2}(1-2B)-\frac{A}{4}-\mu_B\left(\left(\frac{2-r}{2+r}\right)^r+\alpha_2\right) & =0 \\ (C) & \frac{1}{2}\left(-1-r^2\frac{1}{2}(1-2C)\right)A(1-A)+\frac{1-A}{2}(1-2C)+\frac{1-A}{4}-\mu_C\left(\left(\frac{2+r}{2-r}\right)^r+\alpha_2\right) & =0 \\ (\alpha_2) & \mu_B(1-B)+\mu_C(1-C) & =0. \end{array} \tag{80}$$

We now show that the point $A=\frac{1}{2}$, $\alpha_2=1$, and B and C that satisfy the constraints is a critical point of the Lagrangean. (This implies that $\alpha_1=1$ and $\alpha_2=1$ is a critical point if we view total effort as function of α_1 and α_2.)

Recall that $\alpha_2=1$ implies $B+C=1$. So $A=\frac{1}{2}$ solves the first FOC. The second becomes

$$\frac{(1-2B)\left(4-r^2\right)}{16}-\mu_B\left(\left(\frac{2-r}{2+r}\right)^r+1\right)=0,$$

which gives us μ_B, since B is determined by (78) as

$$B=\frac{1}{1+\left(\frac{2-r}{2+r}\right)^r},$$

so

$$\mu_B=\frac{(1-2B)B\left(4-r^2\right)}{16}.$$

Similarly, the third FOC becomes

$$\frac{(1-2C)\left(4-r^2\right)}{16}-\mu_C\left(\left(\frac{2+r}{2-r}\right)^r+1\right)=0$$

which gives us μ_C, since C is determined by (79) as

$$C = \frac{1}{1+\left(\frac{2+r}{2-r}\right)^r}.$$

Therefore,

$$\mu_C = \frac{(1-2C)\,C\left(4-r^2\right)}{16}$$

The last FOC is then automatically satisfied at $\alpha_2 = 1$ by $B + C = 1$; indeed we have

$$\begin{aligned}
\mu_B\,(1-B) + \mu_C\,(1-C) &= \frac{(1-2B)\left(4-r^2\right)}{16}B\,(1-B) + \frac{(1-2C)\left(4-r^2\right)}{16}C\,(1-C) \\
&= (1-2B+1-2C)\,\frac{\left(4-r^2\right)}{16}B\,(1-B) \\
&= 2\,(1-B-C)\,\frac{\left(4-r^2\right)}{16}B\,(1-B) \\
&= 0.
\end{aligned}$$

We now check the second-order condition by building the Hessian matrix of the Lagrangean. We have

$$\begin{aligned}
\frac{\partial^2 L}{\partial A^2} &= -\left(B+1-C-r^2\frac{1}{2}\left(C\,(1-C)+B\,(1-B)\right)\right), \\
\frac{\partial^2 L}{\partial A \partial B} &= (1-2A)\,\frac{1}{2}\left(1-r^2\frac{1}{2}\,(1-2B)\right)+\frac{1}{2}\,(1-2B)-\frac{1}{4}, \\
\frac{\partial^2 L}{\partial A \partial C} &= (1-2A)\,\frac{1}{2}\left(-1-r^2\frac{1}{2}\,(1-2C)\right)-\frac{1}{2}\,(1-2C)-\frac{1}{4}, \\
\frac{\partial^2 L}{\partial B^2} &= \frac{1}{2}r^2 A\,(1-A)-A, \\
\frac{\partial^2 L}{\partial C^2} &= \frac{1}{2}r^2\cdot A\,(1-A)-(1-A)\,, \\
\frac{\partial^2 L}{\partial B \partial \alpha_2} &= -\mu_B, \quad \frac{\partial^2 L}{\partial C \partial \alpha_2} = -\mu_C, \\
\frac{\partial^2 L}{\partial A \partial \alpha_2} &= \frac{\partial^2 L}{\partial B \partial C} = 0.
\end{aligned} \tag{81}$$

Furthermore, we need the Jacobian of the constraints, which is

$$\begin{bmatrix} 0 & \left(\frac{2-r}{2+r}\right)^r + \alpha_2 & 0 & -(1-B) \\ 0 & 0 & \left(\frac{2+r}{2-r}\right)^r + \alpha_2 & -(1-C) \end{bmatrix}.$$

At $\alpha_2 = 1$, $B + C = 1$, and $A = \frac{1}{2}$, we have that the Hessian becomes

$$\begin{bmatrix} -\left(2 - r^2(1-B)\right)B & \frac{1}{4} - B & \frac{1}{4} - B & 0 \\ \frac{1}{2}(1-2B) - \frac{1}{4} & \frac{r^2}{8} - \frac{1}{2} & 0 & -\frac{(1-2B)B\left(4-r^2\right)}{16} \\ \frac{1}{4} - B & 0 & \frac{r^2}{8} - \frac{1}{2} & -\frac{(1-B)(2B-1)\left(4-r^2\right)}{16} \\ 0 & -\frac{(1-2B)B\left(4-r^2\right)}{16} & -\frac{(1-B)(2B-1)\left(4-r^2\right)}{16} & 0 \end{bmatrix},$$

and the Jacobian, using the definition of B and C, simplifies as

$$\begin{bmatrix} 0 & \frac{1}{B} & 0 & -(1-B) \\ 0 & 0 & \frac{1}{C} & -(1-C) \end{bmatrix}.$$

We now use Theorem 16.4 in Simon and Blume (1994), with four variables and two constraints. The Bordered Hessian is

$$H = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{B} & 0 & -(1-B) \\ 0 & 0 & 0 & 0 & \frac{1}{C} & -(1-C) \\ 0 & 0 & -\left(2 - r^2(1-B)\right)B & \frac{1}{2}(1-2B) - \frac{1}{4} & -\frac{1}{2}(1-2C) - \frac{1}{4} & 0 \\ \frac{1}{B} & 0 & \frac{1}{2}(1-2B) - \frac{1}{4} & \frac{r^2}{8} - \frac{1}{2} & 0 & -\frac{(1-2B)B\left(4-r^2\right)}{16} \\ 0 & \frac{1}{C} & -\frac{1}{2}(1-2C) - \frac{1}{4} & 0 & \frac{r^2}{8} - \frac{1}{2} & -\frac{(1-2C)C\left(4-r^2\right)}{16} \\ -(1-B) & -(1-C) & 0 & -\frac{(1-2B)B\left(4-r^2\right)}{16} & -\frac{(1-2C)C\left(4-r^2\right)}{16} & 0 \end{bmatrix},$$

with determinant

$$-\frac{1}{8}\frac{n^D(B,r)}{1-B},$$

where

$$n^D(B,r) \equiv 96B^2 - 66B - 32B^3 - 6r^2 + r^4 + 40Br^2 - 7Br^4 - 60B^2r^2 + 24B^3r^2 + 12B^2r^4 - 6B^3r^4 + 10.$$

Finally, moving to next leading principal minor, i.e., In this case, let's continue with the next leading principal minor.

$$\begin{bmatrix} 0 & 0 & 0 & \frac{1}{B} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{1-B} \\ 0 & 0 & -\left(2 - r^2(1-B)\right)B & \frac{1}{4} - B & \frac{1}{4} - B \\ \frac{1}{B} & 0 & \frac{1}{2}(1-2B) - \frac{1}{4} & \frac{r^2}{8} - \frac{1}{2} & 0 \\ 0 & \frac{1}{1-B} & \frac{1}{4} - B & 0 & \frac{r^2}{8} - \frac{1}{2} \end{bmatrix},$$

we note that its determinant is

$$\frac{r^2(1-B) - 2}{B(1-B)^2} < 0.$$

Therefore, the pattern of signs of the determinant is such that if $n^D(B(r), r) > 0$ then we have that the Hessian Matrix is indefinite in the constraints (recall Theorem 16.4 in Simon and Blume; 1994). Note that, with $r = 1$, we have $B = \frac{3}{4}$ and $n^D(B, r) = \frac{43}{32} > 0$ and we have an indefinite Hessian Matrix, and thus $\alpha_1 = \alpha_2 = \alpha_3 = 1$ does not maximize TE.

We next show that n^D starts negative for low values of r, then it becomes positive right after $\hat{r} = 0.826581$, and it stays positive past $r = 1$.

The first step is to note that B is a strictly increasing function of $r \in (0, 1]$, which can be see with routine algebra. The second step is to show that, for any given $r \in (0, 1]$, $n^D(B, r)$ is an increasing function of B when $B \in [1/2, 1]$. In fact,

$$\frac{\partial n^D(B, r)}{\partial B} = 192B - 120Br^2 + 24Br^4 + 72B^2r^2 - 18B^2r^4 - 96B^2 + 40r^2 - 7r^4 - 66,$$

which is always positive as it starts positive by

$$\left.\frac{\partial n^D(B, r)}{\partial B}\right|_{B=1/2} = 6 - 2r^2 + \frac{r^4}{2} > 0,$$

and it is strictly increasing in B

$$\begin{aligned}\frac{\partial^2 n^D(B, r)}{\partial B^2} &= 24(8 - 5r^2 + r^4) - 12B(16 - 12r^2 + 3r^4) \\ &> 24(8 - 5r^2 + r^4) - 12(16 - 12r^2 + 3r^4) \\ &= 12r^2\left(2 - r^2\right) \\ &> 0.\end{aligned}$$

Therefore, $n^D(B(r), r)$ is an increasing function of r. Furthermore, with $r = 1$, we have $n^D(B(1), 1) = \frac{43}{32} > 0$ and with $r \to 0$ we have $\lim_{r \to 0} n^D(B(r), r) = -3$. Therefore, all in all, $n^D(B(r), r)$ is an increasing function of r which starts negative, takes value 0 for $\hat{r} \approx 0.826581$ only, and is then positive till $r = 1$.

Therefore, Theorem 16.4 in Simon and Blume (1994) part (c) applies, and we have an indefinite Hessian on the constraint set for $r > \hat{r}$. Therefore, $\alpha_1 = \alpha_2 = \alpha_3 = 1$ fails the necessary second order condition for a constrained maximum for $r > \hat{r}$ (see Theorem 19.6 in Simon and Blume (1994), and the note on necessary condition on page 468). ■

Proof of Proposition 9. Recall (NOT) and (74). Let $B(\alpha_2)$ be the solution of (78) and $C(\alpha_2)$ that of (79). Let $A(\alpha_2) = A\left(\frac{1}{\alpha_2}, B(\alpha_2), C(\alpha_2)\right)$ be the solution of (72) where $\alpha_{(0,0)} = \frac{1}{\alpha_2}$, $p_X^{(1,0)} = B(\alpha_2)$ and $p_X^{(0,1)} = C(\alpha_2)$. As we are focusing on alternating contest, we fix $D = \frac{1}{2}$.

First, note that

$$\frac{d\tau_r\left(A\left(\alpha_2\right),B\left(\alpha_2\right),C\left(\alpha_2\right),\frac{1}{2}\right)}{d\alpha_2}=\frac{\partial\tau_r}{\partial A}\cdot A'\left(\alpha_2\right)+\frac{\partial\tau_r}{\partial B}B'\left(\alpha_2\right)+\frac{\partial\tau_r}{\partial C}C'\left(\alpha_2\right).$$

This will equal zero at $\alpha_2=1$. Indeed,

$$\frac{\partial\tau_r}{\partial A}=(1-2A)\frac{1}{2}\left(B+(1-C)-r^2\frac{1}{2}\left(C\left(1-C\right)+B\left(1-B\right)\right)\right)+\frac{1}{2}B\left(1-B\right)-\frac{1}{2}C\left(1-C\right)+\frac{1}{4}\left(1-B-C\right)=0$$

as calculated in (80) in the Proof of Proposition 8. Moreover,

$$\begin{aligned}\frac{\partial\tau_r}{\partial B}&=\frac{1}{2}\left(1-r^2\frac{1}{2}\left(1-2B\right)\right)A\left(1-A\right)+\frac{A}{2}\left(1-2B\right)-\frac{A}{4}=\frac{1}{16}\left(r-2\right)\left(r+2\right)\left(2B-1\right),\\ \frac{\partial\tau_r}{\partial C}&=\frac{1}{2}\left(-1-r^2\frac{1}{2}\left(1-2C\right)\right)A\left(1-A\right)+\frac{1-A}{2}\left(1-2C\right)+\frac{1-A}{4}=\frac{1}{16}\left(r-2\right)\left(r+2\right)\left(2C-1\right),\end{aligned}$$

and by $B+C=1$,

$$\frac{\partial\tau_r}{\partial B}=-\frac{\partial\tau_r}{\partial C}.$$

Finally, note that from the constraints, $B=\left[1+\frac{1}{\alpha_2}\left(\frac{2-r}{2+r}\right)^r\right]^{-1}$, so $B'\left(\alpha_2\right)=\frac{B(1-B)}{\alpha_2}$. Similarly, $C'\left(\alpha_2\right)=\frac{C(1-C)}{\alpha_2}$. Therefore, with $\alpha_2=1$, $B'\left(\alpha_2\right)=C'\left(\alpha_2\right)=B\left(1-B\right)$ as $B+C=1$. And

$$\frac{d\tau_r\left(A\left(\alpha_2\right),B\left(\alpha_2\right),C\left(\alpha_2\right),\frac{1}{2}\right)}{d\alpha_2}=0$$

then follows.

Moving on to the second derivative, we have

$$\begin{aligned}\frac{d^2\tau_r\left(A(\alpha_2),B(\alpha_2),C(\alpha_2),\frac{1}{2}\right)}{\left(d\alpha_2\right)^2}&=\frac{d\left(\frac{\partial\tau_r}{\partial A}\right)}{d\alpha_2}\cdot A'\left(\alpha_2\right)+\frac{d\left(\frac{\partial\tau_r}{\partial B}\right)}{d\alpha_2}B'\left(\alpha_2\right)+\frac{d\left(\frac{\partial\tau_r}{\partial C}\right)}{d\alpha_2}C'\left(\alpha_2\right)\\ &+\frac{\partial\tau_r}{\partial A}\cdot A''\left(\alpha_2\right)+\frac{\partial\tau_r}{\partial B}B''\left(\alpha_2\right)+\frac{\partial\tau_r}{\partial C}C''\left(\alpha_2\right).\end{aligned}\tag{82}$$

Several of the terms in (82) can be found in (80) and (81) in the Proof of Proposition 8. Moreover, $B'\left(\alpha_2\right)$ and $C'\left(\alpha_2\right)$ are calculated above. The missing terms are $A''\left(\alpha_2\right),B''\left(\alpha_2\right),C''\left(\alpha_2\right)$, and $A'\left(\alpha_2\right)$.

Since $\frac{\partial\tau_r}{\partial A}=0$ as calculated above when $\alpha_2=1$, we do not need an expression for $A''\left(\alpha_2\right)$. Proceeding with the calculations of (82), from (78) and (79),

$$B''\left(\alpha_2\right)=-\frac{2B^2\left(1-B\right)}{\left(\alpha_2\right)^2},\ C''\left(\alpha_2\right)=-\frac{2C^2\left(1-C\right)}{\left(\alpha_2\right)^2},$$

which, evaluated at $\alpha_2 = 1$, yields

$$B''(\alpha_2) = -2B^2(1-B), \ C''(\alpha_2) = -2(1-B)B(1-B).$$

The last missing term is $A'(\alpha_2)$. With $\alpha_1 = \frac{1}{\alpha_2}$ and $D = \frac{1}{2}$, we can write from (72),

$$A(\alpha_2) = \left(1 + \alpha_2 \left(\frac{n^A(B(\alpha_2), C(\alpha_2))}{d^A(B(\alpha_2), C(\alpha_2))}\right)^r\right)^{-1},$$

where

$$\begin{aligned} n^A(B,C) &\equiv \left(1+\frac{r}{2}\right)(1-C)(1-rC) + \left(1-\frac{r}{2}\right)B(1+r(1-B)), \\ d^A(B,C) &\equiv \left(1+\frac{r}{2}\right)B(1-r(1-B)) + \left(1-\frac{r}{2}\right)(1-C)(1+rC). \end{aligned}$$

Therefore, suppressing for convenience the dependence of B and C from α_2 when clear, we have

$$\begin{aligned} A'(\alpha_2) &= -\frac{\left(\left(\frac{n^A}{d^A}\right)^r + \alpha_2 r \left(\frac{n^A}{d^A}\right)^{r-1} \frac{\left(\frac{\partial n^A}{\partial B}B'(\alpha_2) + \frac{\partial n^A}{\partial C}C'(\alpha_2)\right)d^A - \left(\frac{\partial d^A}{\partial B}B'(\alpha_2) + \frac{\partial d^A}{\partial C}C'(\alpha_2)\right)n^A}{(d^A)^2}\right)}{\left(1+\alpha_2\left(\frac{n^A}{d^A}\right)^r\right)^2} \\ &= -A^2\left(\frac{1}{\alpha_2}\left(\frac{1-A}{A}\right)\right) \\ &\quad -A^2\alpha_2 r\left(\frac{1}{\alpha_2}\left(\frac{1-A}{A}\right)\right)\frac{d^A}{n^A}\frac{\left(\frac{\partial n^A}{\partial B}B'(\alpha_2) + \frac{\partial n^A}{\partial C}C'(\alpha_2)\right)d^A - \left(\frac{\partial d^A}{\partial B}B'(\alpha_2) + \frac{\partial d^A}{\partial C}C'(\alpha_2)\right)n^A}{(d^A)^2}. \end{aligned}$$

Note now that

$$\begin{aligned} \frac{\partial n^A}{\partial B} &= \left(1-\frac{r}{2}\right)(1+r(1-2B)), \ \frac{\partial n^A}{\partial C} = -\left(1+\frac{r}{2}\right)(1+r(1-2C)), \\ \frac{\partial d^A}{\partial B} &= \left(1+\frac{r}{2}\right)(1-r(1-2B)), \ \frac{\partial d^A}{\partial C} = \left(1-\frac{r}{2}\right)(-1+r(1-2C)). \end{aligned}$$

Therefore, when $\alpha_2 = 1$, we have $n^A(B,C) = d^A(B,C) = B\left(2 - r^2(1-B)\right), \ B'(\alpha_2) = C'(\alpha_2)$, and we obtain

$$
\begin{aligned}
\left.A'(\alpha_2)\right|_{\alpha_2=1} &= -A^2 + A^2 r \frac{B(1-B)}{\alpha_2} \frac{2r(2B-2C+1)}{B(2-r^2(1-B))} \\
&= -\frac{1}{4} + \frac{1}{2} r (1-B) \frac{r(2B-2C+1)}{2-r^2(1-B)} \\
&= -\frac{1}{4} + \frac{1}{2} r C \frac{r(3-4C)}{2-r^2 C} \\
&= \frac{8C^2 r^2 - 7Cr^2 + 2}{4Cr^2 - 8},
\end{aligned}
$$

where we used $A = 1/2$ and $B + C = 1$.

We are now ready to compute the second derivative of τ_r. From (82), straightforward but tedious algebra, using $C = 1 - B$, shows:

$$
\left.\frac{d^2\tau_r\left(A(\alpha_2), B(\alpha_2), C(\alpha_2), \frac{1}{2}\right)}{(d\alpha_2)^2}\right|_{\alpha_2=1} = B \frac{n^S(B,r)}{32 + 16(B-1)r^2}, \tag{83}
$$

where

$$
n^S(B,r) = 4 - 8B(9+4B(-5+3B)) - 4(-1+B)(-5+B(41-74B+44B^2))r^2 - (-1+B)^2(-1-4B+52B^2)r^4.
$$

As the denominator of (83) is positive, we only analyze the sign of $n^S(B,r)$ in what follows. Using the definition of B, for $r \in [0,1]$, we have that $B \in [1/2, 3/4]$.

We conclude the proof by showing that $n^S(B,r)$ is increasing in r when $B \in [1/2, 3/4]$, so that there is a unique $\tilde{r} \in (0,1]$ such that if $r > \tilde{r}$ then the fully unbiased contest is a minimum, and if $r < \tilde{r}$, it is not.

We write

$$
\frac{\partial n^S(B(r), r)}{\partial r} = \frac{\partial n^S}{\partial B} \frac{\partial B}{\partial r} + \frac{\partial n^S}{\partial r},
$$

where we obtain

$$
\begin{aligned}
\tfrac{\partial n^S}{\partial B} = \;& 8(40B - 36B^2 - 9) + 8(23 - 115B + 177B^2 - 88B^3)r^2 \\
& -2(B-1)(104B^2 - 58B + 1)r^4,
\end{aligned} \tag{84}
$$

$$
\frac{\partial n^S}{\partial r} = -8(B-1)(44B^3 - 74B^2 + 41B - 5)r - 4(B-1)^2(52B^2 - 4B - 1)r^3. \tag{85}
$$

In what follows we show that $\partial n^S/\partial B > 0$ in Step 1 and $\partial B(r)/\partial r > 2r/5$ in Step 2, while Step 3 uses these two results to conclude the proof.

Step 1. The term $\partial n^S/\partial B$ is a polynomial of degree two in r^2. One can show that the two roots solving $\partial n^S/\partial B = 0$ are *not* in the range $r \in [0,1]$, and for $r = 0$ we have that $\partial n^S/\partial B > 0$

(recall that $B \in [1/2, 3/4]$). Thus, $\partial n^S / \partial B > 0$ $\forall r \in [0,1]$ and $\forall B \in [1/2, 3/4]$.

Step 2. We show that

$$\begin{aligned}
\frac{\partial B(r)}{\partial r} &> \frac{2r}{5} \iff -\frac{\left(-1+\frac{4}{2+r}\right)^r \left[4r + (-4+r^2) Log\left(-1+\frac{4}{2+r}\right)\right]}{(-4+r^2)\left[1+(-1+\frac{4}{2+r})^r\right]^2} > \frac{2r}{5} \\
&\iff 5\left(\frac{2-r}{2+r}\right)^r \left[\frac{2}{(2-r)(2+r)} - \frac{1}{2r} Log\left(\frac{2-r}{2+r}\right)\right] > \left[1+\left(\frac{2-r}{2+r}\right)^r\right]^2 \qquad (86)
\end{aligned}$$

We use the following logarithmic inequality from the second row of Table 1 of Topsøe (2006):

$$\log(1+x) \leq \frac{x(6+x)}{6+4x} \;\forall x > -1,$$

so that we obtain

$$-\frac{1}{2r} Log\left(\frac{2-r}{2+r}\right) = -\frac{1}{2r} Log\left(1-\frac{2r}{2+r}\right) > -\frac{1}{2r}\frac{-\frac{2r}{2+r}\left(6-\frac{2r}{2+r}\right)}{6+4\left(-\frac{2r}{2+r}\right)} = \frac{2r+6}{(2+r)(6-r)}. \qquad (87)$$

Now, using (87) into (86) we obtain the following sufficient condition for $\partial B(r)/\partial r > 2r/5$:

$$\begin{aligned}
5\left(\frac{2-r}{2+r}\right)^r \left[\frac{2}{(2-r)(2+r)} + \frac{2r+6}{(2+r)(6-r)}\right] &> \left[1+\left(\frac{2-r}{2+r}\right)^r\right]^2 \\
10\left[\frac{1}{(2-r)(2+r)} + \frac{r+3}{(2+r)(6-r)}\right] &> \left(\frac{2+r}{2-r}\right)^r + 2 + \left(\frac{2-r}{2+r}\right)^r,
\end{aligned}$$

and as

$$\left(\frac{2+r}{2-r}\right)^r + 2 + \left(\frac{2-r}{2+r}\right)^r < \left(\frac{2+1}{2-1}\right)^r + 2 + \left(\frac{2-0}{2+0}\right)^r = 3^r + 3,$$

it suffices to show that

$$10\left[\frac{1}{(2-r)(2+r)} + \frac{r+3}{(2+r)(6-r)}\right] > 3^r + 3. \qquad (88)$$

Finally, the LHS of (88) is greater than $5 - r + 2r^2$, which is in turn greater than the RHS of (88). In fact, the former condition is equivalent to

$$(r-6)(r-2)(r-1)r(r+2)(24-8r-11r^2+2r^3) < 0,$$

which holds true as the polynomial of degree three of the last factor is positive in $r \in (0,1]$. The latter condition is equivalent to

$$\begin{aligned} 5 - r + 2r^2 &> 3^r + 3 \\ &\iff 2 - r > 3^r - 2r^2, \end{aligned}$$

where at $r = 0$ the LHS equals 2 and the RHS equals 0, at $r = 1$ both the LHS and RHS equal 0, and the first derivative of the LHS is always smaller than that of the RHS. It follows that $\partial B(r)/\partial r > 2r/5$.

Step 3. Hence, from $\partial n^S/\partial B > 0$ (Step 1) and $\partial B(r)/\partial r > 2r/5$ (Step 2), using the expressions (84) and (85), we obtain

$$\begin{aligned} \frac{\partial n^S(B(r), r)}{\partial r} &> \frac{\partial n^S}{\partial B} \cdot \frac{2r}{5} + \frac{\partial n^S}{\partial r} \\ &= \frac{r}{5} \left[\begin{array}{c} -344 + 2480B - 5176B^2 + 4720B^3 - 1760B^4 \\ + \left(388 - 1800B + 1652B^2 + 752B^3 - 1040B^4\right) r^2 \\ -4(B-1)(1 - 58B + 104B^2) r^4 \end{array} \right]. \end{aligned} \tag{89}$$

Hence, in order to show that there is a unique $\tilde{r} \in (0, 1]$ such that if $r > \tilde{r}$ then the fully unbiased contest is a minimum, and if $r < \tilde{r}$, it is not, we will show that the term in the square bracket of (89) is positive.

First, we show that the term in square bracket of (89) is decreasing in $r^2 \in (0, 1]$. Its derivative with respect to r^2 equals

$$\left(388 - 1800B + 1652B^2 + 752B^3 - 1040B^4\right) + 8(B-1)(1 - 58B + 104B^2) r^2. \tag{90}$$

Now, the first bracket is always negative as $\forall B \in [1/2, 3/4]$ we have

$$\underbrace{388 - 1800B + 1652B^2}_{\leq 30} + \underbrace{752B^3 - 1040B^4}_{\leq -30}. \tag{91}$$

The coefficient of r^2 in (90) could be positive or negative. If it is negative, we are done proving that the term in square bracket of (89) is decreasing in r^2. If it is positive, (90) is convex in r, and it takes negative values both at $r = 0$ and $r = 1$. At $r = 0$ it follows by (91). At $r = 1$ its negativity is equivalent to

$$-95 + 332B - 89B^2 - 396B^3 + 260B^4 > 0, \tag{92}$$

where the LHS is concave in B for $B \in [1/2, 3/4]$ and takes value $31/2$ and $1225/64$ at $B = 1/2$ and $B = 3/4$, respectively. Hence, (92) follows, and thus the term in square bracket of (89) is decreasing in r^2.

As the term in the square bracket of (89) is decreasing in r^2, it is greater than its value at $r = 0$, which is

$$\begin{aligned} -344 + 2480B - 5176B^2 + 4720B^3 - 1760B^4 &= -344 + 2480B - 5176B^2 + 3400B^3 + 1320B^3 - 1760B^4 \\ &> -344 + 2480B - 5176B^2 + 3400B^3, \end{aligned}$$

which, in the range $B \in [1/2, 3/4]$ has a unique minimum at $(647 + \sqrt{23359})/1275$, where it takes positive value. ■

Proof of propositions 10 and 11. The proofs of propositions 10 and 11 are very similar, and hence we merge them in what follows. Both rely on the same equilibrium characterization in Proposition 1 of Ewerhart (2017). We begin by adapting Proposition 1 in Ewerhart (2017) to our setting with bias α. Consider a generic match where the payoffs of player X and Y, respectively, are

$$\frac{\alpha x^r}{\alpha x^r + y^r}\Delta u_X - x + u_X^L \text{ and } \frac{y^r}{\alpha x^r + y^r}\Delta u_Y - y + u_Y^L,$$

where $\Delta u_k = u_k^W - u_k^L$, for $k \in \{X, Y\}$ is the effective prize of the match. To match Proposition 1 in Ewerhart (2017), we work with the notation $\tilde{x} \equiv \alpha^{\frac{1}{r}} x$, and y. Hence, the payoff of player X can be rewritten as

$$\frac{\tilde{x}^r}{\tilde{x}^r + y^r}\Delta u_X - \frac{\tilde{x}}{\alpha^{\frac{1}{r}}} + u_X^L = \frac{1}{\alpha^{\frac{1}{r}}}\left(\frac{\tilde{x}^r}{\tilde{x}^r + y^r}\alpha^{\frac{1}{r}}\Delta u_X - \tilde{x} + \alpha^{\frac{1}{r}} u_X^L\right).$$

Similarly, for player Y the payoff is

$$\frac{y^r}{\tilde{x}^r + y^r}\Delta u_Y - y + u_Y^L.$$

To apply Proposition 1 in Ewerhart, we need to define two cases that distinguish who is the "strong" player (i.e., the one with the largest prize valuation). This depends on whether $\alpha^{\frac{1}{r}}\Delta u_X$ is larger or smaller than Δu_Y.[36] In particular, we have two cases:

Case I: $\alpha^{\frac{1}{r}}\Delta u_X \geq \Delta u_Y$, so player X is player 1 in Ewerhart's notation. Then we know there are no pure-strategy equilibria and Proposition 1 in Ewerhart (2017) implies that, in any mixed-strategy equilibrium with $r \in (2, \infty)$, player X bids in expectation $E[\tilde{x}] = \Delta u_Y/2$, so that, using

[36] If $\alpha^{\frac{1}{r}}\Delta u_X = \Delta u_Y$ the two cases merge. For technical convenience, this possibility is included in both cases below.

the same notation of Section 3, where we omit the expectation operator for simplicity, we obtain

$$\begin{aligned}
x &= \frac{\Delta u_Y}{2\alpha^{\frac{1}{r}}},\ y = \frac{\Delta u_Y^2}{2\alpha^{\frac{1}{r}}\Delta u_X}, \\
p_X &= 1 - \frac{\Delta u_Y}{2\alpha^{\frac{1}{r}}\Delta u_X},\ p_Y = \frac{\Delta u_Y}{2\alpha^{\frac{1}{r}}\Delta u_X}, \\
u_X &= \frac{1}{\alpha^{\frac{1}{r}}}\left[\alpha^{\frac{1}{r}}\Delta u_X - \Delta u_Y + \alpha^{\frac{1}{r}} u_X^L\right] = u_X^W - \frac{\Delta u_Y}{\alpha^{\frac{1}{r}}},\ u_Y = u_Y^L.
\end{aligned}$$

Case II: $\alpha^{\frac{1}{r}}\Delta u_X \leq \Delta u_Y$, so player Y is player 1 in Ewerhart's notation. As above, there are no pure-strategy equilibria. Proposition 1 in Ewerhart (2017) implies that, in any mixed strategy equilibrium with $r \in (2, \infty)$, we obtain

$$\begin{aligned}
x &= \frac{\alpha^{\frac{1}{r}}\Delta u_X^2}{2\Delta u_Y},\ y = \frac{\alpha^{\frac{1}{r}}\Delta u_X}{2}, \\
p_X &= \frac{\alpha^{\frac{1}{r}}\Delta u_X}{2\Delta u_Y},\ p_Y = 1 - \frac{\alpha^{\frac{1}{r}}\Delta u_X}{2\Delta u_Y}, \\
u_X &= u_X^L,\ u_Y = u_Y^W - \alpha^{\frac{1}{r}}\Delta u_X.
\end{aligned}$$

We can now use the above results to find the TE-maximizing vector of biases. For convenience of exposition, we start the analysis proving Proposition 11, with victory-independent biases. We start with the equilibrium analysis node-by-node for a given vector $\{\alpha_1, \alpha_2, \alpha_3\}$. We normalize the prize to 1. We begin with $\alpha_3 \in (0, 1)$; that is, player Y has the advantage in the tie-break. (The case $\alpha_3 > 1$ is symmetric, and the case $\alpha_3 = 1$ is discussed separately.)

In node $(1, 1)$ we are in Case II above, as both players have a winning payoff of 1 and a losing payoff of 0. Hence, in equilibrium player X expected bid is $\left(\alpha_3^{1/r}\right)/2$, X wins with probability $\left(\alpha_3^{1/r}\right)/2$, and has a payoff of 0, while player Y bids in expectation $\left(\alpha_3^{1/r}\right)/2$, wins with probability $1 - \left(\alpha_3^{1/r}\right)/2$, and has a payoff of $1 - \alpha_3^{1/r}$.

In node $(0, 1)$, as X has both a winning and losing payoff of 0 whereas Y's effective prize is strictly positive, by assumption (A) both players bid 0, player Y wins for sure and has a payoff of 1, and player X loses for sure and has a payoff of 0.

In node $(1, 0)$ X is fighting for 1 and Y is fighting for $1 - \alpha_3^{1/r}$. Thus, the analysis distinguishes two possibilities: we are in Case I if $\alpha_2^{1/r} \cdot 1 \geq 1 - \alpha_3^{1/r}$, or $\alpha_2^{1/r} + \alpha_3^{1/r} \geq 1$ and Case II if $\alpha_2^{1/r} + \alpha_3^{1/r} \leq 1$.

We now show that $\alpha_2^{1/r} + \alpha_3^{1/r} < 1$ always implies $TE = 0$, and hence the optimum has to comply with Case I. In node $(1, 0)$, under Case II, the payoff of Y is $1 - \alpha_3^{1/r} - \alpha_2^{1/r}$ and that of X is 0. Then, in node $(0, 0)$ X is fighting for 0 and Y is fighting for $1 - \left(1 - \alpha_3^{1/r} - \alpha_2^{1/r}\right) = \alpha_3^{1/r} + \alpha_2^{1/r}$. Hence, Y wins for sure the first match with no effort by assumption (A). All in all, Y wins for sure also the first two matches with no effort, and thus $TE = 0$. This result is intuitive. As we assumed

$\alpha_3 < 1$, the case of $\alpha_2^{1/r} + \alpha_3^{1/r} < 1$ gives too large an advantage to player Y, who wins the entire contest with no effort; player X anticipates in node $(0,0)$ that even a victory would give her an expected payoff of 0 because of the future advantages given to player Y in second and third match.

Having ruled out $\alpha_2^{1/r} + \alpha_3^{1/r} < 1$, we henceforth consider $\alpha_2^{1/r} + \alpha_3^{1/r} \geq 1$. Then, in node $(1,0)$, we obtain

$$\begin{aligned}
x^{(1,0)} &= \frac{1-\alpha_3^{1/r}}{2\alpha_2^{1/r}},\ y^{(1,0)} = \frac{\left(1-\alpha_3^{1/r}\right)^2}{2\alpha_2^{1/r}}, \\
p_X^{(1,0)} &= 1 - \frac{1-\alpha_3^{1/r}}{2\alpha_2^{1/r}},\ p_Y^{(1,0)} = \frac{1-\alpha_3^{1/r}}{2\alpha_2^{1/r}}, \\
u_X^{(1,0)} &= 1 - \frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}},\ u_Y^{(1,0)} = 0.
\end{aligned}$$

Finally, in node $(0,0)$, X is fighting for $\Delta u_X = 1 - \frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}$ and Y is fighting for $\Delta u_Y = 1$. As $\alpha_2^{1/r} + \alpha_3^{1/r} = 1$ would imply as before that Y wins for sure with no effort and $TE = 0$, we can now focus on $\alpha_2^{1/r} + \alpha_3^{1/r} > 1$. To know whether node $(0,0)$ complies with Case I or Case II, we need to check whether

$$\begin{aligned}
\alpha_1^{1/r}\Delta u_x &\geq \Delta u_Y \\
&\iff \alpha_1^{1/r}\left(1 - \frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}\right) \geq 1 \\
&\iff \frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}} \leq 1 - \frac{1}{\alpha_1^{1/r}}. \qquad (93)
\end{aligned}$$

Under the previously imposed $\alpha_2^{1/r} + \alpha_3^{1/r} > 1$, inequality (93) remains undecided, so we study both cases separately in the remainder of the proof.

Case I: $\frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}} \leq 1 - \frac{1}{\alpha_1^{1/r}}$. Then node $(0,0)$ follows Case I. So, we obtain

$$\begin{aligned}
x^{(0,0)} &= \frac{1}{2\alpha_1^{1/r}},\ y^{(0,0)} = \frac{1}{2\alpha_1^{1/r}\left(1 - \frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}\right)}, \\
p_X^{(0,0)} &= 1 - \frac{1}{2\alpha_1^{1/r}\left(1 - \frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}\right)},\ p_Y^{(0,0)} = \frac{1}{2\alpha_1^{1/r}\left(1 - \frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}\right)}, \\
u_X^{(0,0)} &= \left(1 - \frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}\right) - \frac{1}{\alpha_1^{1/r}},\ u_Y^{(0,0)} = 0.
\end{aligned}$$

Using (2), we can now characterize the optimal TE if $\alpha_3 \in (0,1)$. As efforts in node $(0,1)$ are zero, we obtain

$$\begin{aligned} TE &= \frac{1}{2\alpha_1^{1/r}} + \frac{1}{2\alpha_1^{1/r}\left(1 - \frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}\right)} \\ &\quad + \left(1 - \frac{1}{2\alpha_1^{1/r}\left(1 - \frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}\right)}\right)\left(\frac{1-\alpha_3^{1/r}}{2\alpha_2^{1/r}} + \frac{\left(1-\alpha_3^{1/r}\right)^2}{2\alpha_2^{1/r}} + \frac{1-\alpha_3^{1/r}}{2\alpha_2^{1/r}}\alpha_3^{1/r}\right) \\ &= \frac{1}{\alpha_1^{1/r}} + \frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}. \end{aligned} \tag{94}$$

Now, the problem is to maximize TE choosing the vector of α's given these five constraints

$$\begin{aligned} \alpha_1 &> 0,\ \alpha_2 > 0,\ 0 < \alpha_3 < 1, \\ \alpha_2^{1/r} + \alpha_3^{1/r} &> 1,\ \frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}} \leq 1 - \frac{1}{\alpha_1^{1/r}}. \end{aligned}$$

The constraint $\alpha_2^{1/r} + \alpha_3^{1/r} > 1$ can be written as $\frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}} < 1$, and it is thus implied by the last constraint. Hence, we are left with the following four constraints

$$\begin{aligned} \alpha_1 &> 0,\ \alpha_2 > 0,\ 0 < \alpha_3 < 1, \\ \frac{1}{\alpha_1^{1/r}} &\leq 1 - \frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}. \end{aligned}$$

Hence, given (94), it is clear that the last constraint has to be binding in the optimum. And, if so, by substituting into (94) the binding constraint, we obtain full-rent extraction $TE = 1$. Rearranging the last constraint, which derives from (93), and imposing equality, we obtain that there is a continuum of maximizers satisfying (22) with $\alpha_3^* \in (0,1)$ and $\alpha_1^* \geq 1$.

Case II: $\frac{1-\alpha_3^{\frac{1}{r}}}{\alpha_2^{\frac{1}{r}}} \geq 1 - \frac{1}{\alpha_1^{\frac{1}{r}}}$, then node $(0,0)$ follows Case II, with $\Delta u_X = 1 - \frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}$ and $\Delta u_Y = 1$.

We obtain

$$\begin{aligned}
x^{(0,0)} &= \frac{\alpha_1^{1/r}}{2}\left(1-\frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}\right)^2,\ y^{(0,0)} = \frac{\alpha_1^{1/r}}{2}\left(1-\frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}\right), \\
p_X^{(0,0)} &= \frac{\alpha_1^{1/r}}{2}\left(1-\frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}\right),\ p_Y^{(0,0)} = 1-\frac{\alpha_1^{1/r}}{2}\left(1-\frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}\right), \\
u_X^{(0,0)} &= 0,\ u_Y^{(0,0)} = 1-\alpha_1^{1/r}\left(1-\frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}\right).
\end{aligned}$$

We can finally analyze TE in this case. As efforts in node $(0,1)$ are zero, we can write

$$\begin{aligned}
TE &= \frac{\alpha_1^{1/r}}{2}\left(1-\frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}\right)+\frac{\alpha_1^{1/r}}{2}\left(1-\frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}\right)^2 \\
&\quad +\frac{\alpha_1^{1/r}}{2}\left(1-\frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}\right)\left(\frac{1-\alpha_3^{1/r}}{2\alpha_2^{1/r}}+\frac{\left(1-\alpha_3^{1/r}\right)^2}{2\alpha_2^{1/r}}+\frac{1-\alpha_3^{1/r}}{2\alpha_2^{1/r}}\alpha_3^{1/r}\right) \\
&= \alpha_1^{1/r}-\alpha_1^{1/r}\frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}.
\end{aligned}$$

Now, the constraints under which to maximize TE are

$$\begin{aligned}
\alpha_1 &> 0,\ \alpha_2>0,\ 0<\alpha_3<1, \\
\alpha_2^{1/r}+\alpha_3^{1/r} &> 1,\ \frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}} \geq 1-\frac{1}{\alpha_1^{1/r}},
\end{aligned}$$

and merging the last two constraints, the problem boils down to

$$\max_{\alpha_1>0,\alpha_2>0,\alpha_3\in(0,1]}\left(\alpha_1^{1/r}-\alpha_1^{1/r}\frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}}\right) \text{ such that } 1-\frac{1}{\alpha_1^{1/r}} \leq \frac{1-\alpha_3^{1/r}}{\alpha_2^{1/r}} < 1,$$

and thus, once again, inequality (93) must be binding in the optimum. And, if so, we obtain full-rent extraction $TE=1$. Therefore, we obtain the same continuum of maximizers as in (22) with $\alpha_3^* \in (0,1)$ and $\alpha_1^* \geq 1$.

The next step of the proof is to note that the case $\alpha_3 > 1$ is symmetric to that of $\alpha_3 \in (0,1)$, and thus we can drop the condition $\alpha_3^* \in (0,1)$ and $\alpha_1^* \geq 1$ in the optimum (22). To finish the proof, consider now $\alpha_3 = 1$. In node $(1,1)$ both players have a winning payoff of 1 and a losing payoff of 0. Hence, in equilibrium both player's expected bid is 1/2, X wins with probability 1/2, and both players have payoff 0. In node $(0,1)$, as X has both a winning and losing payoff of 0 whereas Y's effective prize is strictly positive, by assumption (A) both players bid 0, player Y wins for sure and has a payoff of 1, and player X loses for sure and has a payoff of 0. In node $(1,0)$ X is fighting for 1 and Y is fighting for 0. Therefore, by assumption (A) both players bid 0, player X wins for sure and has a payoff of 1, and player Y loses for sure and has a payoff of 0. In node $(0,0)$ both players have a winning payoff of 1 and a losing payoff of 0. Therefore, the unique maximum for TE is obtained for $\alpha_1 = 1$. This concludes the proof of Proposition 11.

We now move to the proof of Proposition 10. We consider only $\alpha_{(1,1)} \in (0,1)$ as the other cases are similar. Note that in the proof of Proposition 11, α_2 enters TE only through efforts and probabilities in node $(1,0)$, as efforts in node $(0,1)$ are zero for any $\alpha_2 \in (0,\infty)$. Hence, the fact that efforts in node $(0,1)$ are zero carries over to the setup of victory-dependent biases with $\alpha_{(1,1)} \in (0,1)$, and the proof of Proposition 11 carries over to the case of victory-dependent biases by just replacing α_1 with $\alpha_{(0,0)}$, α_2 with $\alpha_{(1,0)}$, and α_3 with $\alpha_{(1,1)}$. The continuum of TE-maximizing vectors in (22) remains valid $\forall \alpha^*_{(0,1)} > 0$. And this concludes the proof of Proposition 10. ■

Proof of Proposition 12. First, we prove that $p_X^{(i,j)} = 1/2$ for any $i,j \in \{0,1\}$ is optimal if $r < \check{r}$. Note that setting $p_X^{(i,j)} = 1/2$ for any $i,j \in \{0,1\}$ ensures that a pure-strategy equilibrium exists for any $r \in (1,2]$. Therefore, the level of total effort that is achieved by setting $p_X^{(i,j)} = 1/2$ for any $i,j \in \{0,1\}$ remains feasible even if $r \in (1,2]$. But, when $r \in (1,2]$, the strategy of proof of Proposition 1 has an additional layer of complexity to consider: the values of $p_X^{(1,0)}$ and $p_X^{(0,1)}$ in (75) do not necessarily lie in the interval $\left[1-\frac{1}{r}, \frac{1}{r}\right]$. This happens for choices of $p_X^{(0,0)}$ or $p_X^{(1,1)}$ that are sufficiently far away from $\frac{1}{2}$. It is intuitive that it is nodes $(1,0)$ and $(0,1)$ that fail first to deliver the pure-strategy equilibrium, rather than nodes $(0,0)$ and $(1,1)$, as in nodes $(1,0)$ and $(0,1)$ one player is one-match ahead. For instance, one can show that if $r \leq 1.1$, then $\beta_r(A,D)$ and $\gamma_r(A,D)$ do not lie in the interval $\left[1-\frac{1}{r}, \frac{1}{r}\right]$ only if A and D do not lie in $[0.2, 0.8]$.

Recall from Proposition 7 that $p_X^{(i,j)} = 1/2$ for any $i,j \in \{0,1\}$ is the unique global maximum for $r \leq 1$ and that expected effort changes continuously as we switch from a pure-strategy equilibrium to a semi-mixed one, as it can be seen by comparing (68) and (24) with the threshold value of $p_X = 1/r$. Therefore, by continuity, there exists some $\check{r} > 1$ such that, if $r \in (1,\check{r})$, then $p_X^{(i,j)} = 1/2$ for any $i,j \in \{0,1\}$ remains a strict global maximum for TE. This is because the semi-mixed candidates not considered in the Proof of Proposition 1 are relevant only for $p_X^{(0,0)}$ and $p_X^{(1,1)}$ that at a discrete distance from $\frac{1}{2}$ (in the example $r \leq 1.1$, this requires $p_X^{(0,0)} \in [0,0.2] \cup [0.8,1]$ and $p_X^{(1,1)} \in [0,0.2] \cup [0.8,1]$), and hence these candidates cannot be optimal because efforts are continuous in r and $p_X^{(i,j)} = 1/2$ for any $i,j \in \{0,1\}$ is the unique global maximum if $r < \hat{r}$.

Second, we prove that $p_X^{(i,j)} = 1/2$ for any $i,j \in \{0,1\}$ is *not* optimal if $r > 2\left(\sqrt{3}-1\right)$. In fact, we show that the point $p_X^{(0,0)} = p_X^{(1,0)} = p_X^{(0,1)} = p_X^{(1,1)} = \frac{1}{2}$ is outperformed by the point $p_X^{(0,0)} = \frac{1}{r}, p_X^{(1,0)} = 1/2, p_X^{(0,1)} = 0$, and $p_X^{(1,1)} = 1 - \frac{1}{r}$. Consider this last point, and starting in node

$(1,1)$. We obtain

$$\Delta u_X^{(1,1)} = \Delta u_Y^{(1,1)} = V, \Delta u_X^{(1,1)} + \Delta u_Y^{(1,1)} = 2V.$$

If we set $\alpha^{(1,1)} = r-1$ we are in Case II, and hence using (25) and (27), we obtain

$$p_X^{(1,1)} = 1 - \frac{1}{r},\ u_X^{(1,1)} = 0,\ u_Y^{(1,1)} = V\left(\frac{2}{r} - 1\right).$$

In node $(0,1)$, $\Delta u_X^{(0,1)} = u_X^{(1,1)} - 0 = 0$, but $\Delta u_Y^{(0,1)} > 0$. So, by assumption A, it does not matter what is assumed for the kind of equilibrium: Y wins for sure with zero effort. Therefore, $p_X^{(0,1)} = 0$. In node $(1,0)$, we have

$$\Delta u_X^{(1,0)} = V - u_X^{(1,1)} = V, \Delta u_Y^{(1,0)} = u_Y^{(1,1)} = V\left(\frac{2}{r} - 1\right),\ \Delta u_X^{(1,0)} + \Delta u_Y^{(1,0)} = \frac{2}{r}V.$$

If we set $\alpha^{(1,0)} = \left(\frac{2-r}{r}\right)^r$ we are neither in Case I nor in Case II, so players play the pure-strategy equilibrium which generates formulas (64), (66) , and (67). We then have

$$p_X^{(1,0)} = \frac{1}{2},\ u_X^{(1,0)} = V\frac{1}{2}\left(1 - \frac{r}{2}\right),\ u_Y^{(1,0)} = V\left(\frac{2}{r} - 1\right)\frac{1}{2}\left(1 - \frac{r}{2}\right).$$

In node $(0,0)$, the above implies

$$\begin{aligned}\Delta u_X^{(0,0)} &= V\frac{1}{2}\left(1 - \frac{r}{2}\right), \Delta u_Y^{(0,0)} = V - V\left(\frac{2}{r} - 1\right)\frac{1}{2}\left(1 - \frac{r}{2}\right),\\ \Delta u_X^{(0,0)} + \Delta u_Y^{(0,0)} &= V\frac{5r - 2 - r^2}{2r}.\end{aligned}$$

If we set

$$\alpha^{(0,0)} = \left(\frac{1 - \left(\frac{2}{r} - 1\right)\frac{1}{2}\left(1 - \frac{r}{2}\right)}{\frac{1}{2}\left(1 - \frac{r}{2}\right)}\right)^r \frac{1}{(r-1)}$$

we are in Case I, hence using (23), we obtain

$$p_X^{(0,0)} = \frac{1}{r}.$$

Thus, using (24) , (26), and (68), we obtain

$$
\begin{aligned}
TE &= \left[x^{(0,0)}+y^{(0,0)}\right]+p_X^{(0,0)}\left[x^{(1,0)}+y^{(1,0)}\right]+\left(1-p_X^{(0,0)}\right)\left[x^{(0,1)}+y^{(0,1)}\right] \\
&\quad +\left(p_X^{(0,0)}\cdot\left(1-p_X^{(1,0)}\right)+\left(1-p_X^{(0,0)}\right)p_X^{(0,1)}\right)\left[x^{(1,1)}+y^{(1,1)}\right] \\
&= \left[2V\frac{5r-2-r^2}{4r}\left(1-\frac{1}{r}\right)\right]+\frac{1}{r}\left[\frac{2}{r}Vr\frac{1}{2}\left(1-\frac{1}{2}\right)\right]+0 \\
&\quad +\left(\frac{1}{r}\cdot\frac{1}{2}+0\right)\left[2V\left(1-\frac{1}{r}\right)\right] \\
&= \frac{V}{2r}\left(6r-r^2-4\right).
\end{aligned}
$$

If we compare this value with what we obtain by setting $p_X^{(0,0)}=p_X^{(1,0)}=p_X^{(0,1)}=p_X^{(1,1)}=\frac{1}{2}$, which is $V\frac{r\left(12-r^2\right)}{16}$, we see that

$$
\begin{aligned}
V\frac{1}{2r}\left(6r-r^2-4\right) &\geq V\frac{r\left(12-r^2\right)}{16} \\
&\iff 3-\frac{2}{r}-\frac{5r}{4}+\frac{r^3}{16}\geq 0 \\
&\iff r\geq 2\left(\sqrt{3}-1\right)\approx 1.4641.
\end{aligned}
$$

■

Proof of Proposition 13. We show that if $r\leq r'$, then, after fixing $\alpha_3=1$, $\alpha_1=\alpha_2=1$ is a saddle point for TE in $\mathbb{R}^2_{>0}$. Fix $D=\frac{1}{2}$, $A=\frac{1}{2}$, and let C,B be what comes out of setting $\alpha_2=1$ in (71) and (70). We obtain

$$
B=\frac{1}{1+\left(\frac{2-r}{2+r}\right)^r},\quad C=\frac{1}{1+\left(\frac{2+r}{2-r}\right)^r}.
$$

Note that r' is defined to ensure that, if $r\leq r'$, then the above values of B and C are in the interval $[1-1/r,1/r]$, hence implying a pure-strategy equilibrium, so that the proof in Proposition 8 carries over and the final statement of the proposition is proved. ■

Proof of Proposition 14. Recall (NOT). Using (28) and recalling that in an alternating contest $\alpha_{(0,0)}=\alpha_1=\alpha$, $\alpha_{(1,0)}=\alpha_{(0,1)}=\alpha_2=1/\alpha$, and $\alpha_{(1,1)}=\alpha_3=1$, we can calculate the analogous expressions to (32) , (33), (34) , and (35) as

$$\begin{aligned}
D(\beta) &= \left(1+\frac{1}{\beta}\right)^{-1}, \\
C(\alpha,\beta) &= \left(1+\frac{\alpha(2-D(\beta))}{\beta D(\beta)}\right)^{-1}, \\
B(\alpha,\beta) &= \left(1+\frac{\alpha(1-D(\beta))}{\beta(1+D(\beta))}\right)^{-1}, \\
A(\alpha,\beta) &= \left(1+\frac{1}{\alpha\beta}\frac{((1-C(\alpha,\beta))^2 D(\beta)(2-D(\beta))+B(\alpha,\beta)(1-D(\beta))^2(2-B(\alpha,\beta)))}{((B(\alpha,\beta))^2(1-D(\beta))(1+D(\beta))+(D(\beta))^2(1-C(\alpha,\beta))(1+C(\alpha,\beta)))}\right)^{-1}.
\end{aligned}$$

Then, using τ in (15), we can define total effort as a function of (α,β) as

$$TE(\alpha,\beta) \equiv 2V\tau(A(\alpha,\beta), B(\alpha,\beta), C(\alpha,\beta), D(\beta)).$$

We obtain

$$\left.\frac{\partial TE(\alpha,\beta)}{\partial\beta}\right|_{(\alpha,1)} = \frac{4V(1-\alpha)\alpha^2 P_\beta(\alpha)}{(\alpha+1)^3(\alpha+3)^3(3\alpha+1)^3\left(3\alpha^4+26\alpha^3+166\alpha^2+26\alpha+3\right)^3},$$

where

$$\begin{aligned}
P_\beta(\alpha) \equiv\ & 729\alpha^{16}-15066\alpha^{15}-474228\alpha^{14}-7696998\alpha^{13}-53893404\alpha^{12}-196576690\alpha^{11} \\
&-139964940\alpha^{10}+453592818\alpha^9+866279814\alpha^8+453592818\alpha^7-139964940\alpha^6 \\
&-196576690\alpha^5-53893404\alpha^4-7696998\alpha^3-474228\alpha^2-15066\alpha+729.
\end{aligned}$$

Simple algebra shows that

$$P_\beta(\alpha) = \alpha^{16} P_\beta\left(\frac{1}{\alpha}\right). \tag{95}$$

Hence, using (95) we conclude that

$$\begin{aligned}
\left.\frac{\partial TE(\alpha,\beta)}{\partial\beta}\right|_{(\alpha,1)} &= \frac{4V(1-\alpha)\alpha^{18} P_\beta\left(\frac{1}{\alpha}\right)}{(\alpha+1)^3(\alpha+3)^3(3\alpha+1)^3\left(3\alpha^4+26\alpha^3+166\alpha^2+26\alpha+3\right)^3} \\
&= -\frac{4V(1-\frac{1}{\alpha})\frac{1}{\alpha^2}P_\beta\left(\frac{1}{\alpha}\right)}{(1+\frac{1}{\alpha})^3(1+\frac{3}{\alpha})^3(3+\frac{1}{\alpha})^3\left(3\left(\frac{1}{\alpha}\right)^4+26\left(\frac{1}{\alpha}\right)^3+166\left(\frac{1}{\alpha}\right)^2+26\left(\frac{1}{\alpha}\right)+3\right)^3} \\
&= -\left.\frac{\partial TE(\alpha,\beta)}{\partial\beta}\right|_{(1/\alpha,1)}.
\end{aligned}$$

Furthermore, note that, by Descartes' rule of signs, $P_\beta(\alpha)=0$ has at most five positive solutions.

As

$$
\begin{aligned}
P_\beta(0) &> 0,\ P_\beta\left(\frac{1}{4.3}\right) < 0, \\
P_\beta\left(\frac{1}{4.2}\right) &< 0, P_\beta(1) > 0, \\
P_\beta(1) &> 0, P_\beta(2) < 0, \\
P_\beta(42) &< 0, P_\beta(43) > 0,
\end{aligned}
$$

by continuity we derive four such solutions: one between 0 and $\frac{1}{4.3}$ (i.e., smaller than $1/\alpha^*$), one between $\frac{1}{4.2}$ and 1 (i.e., larger than $1/\alpha^*$), one between 1 and 2 (i.e., smaller than α^*), and one between 42 and 43 (i.e., larger than α^*). It then follows that $P_\beta(\alpha) < 0$ for any $\alpha \in (4.2, 4.3)$, otherwise we would have at least six positive solutions to $P_\beta(\alpha) = 0$ by (95). Therefore, at $\alpha^* \in (4.2, 4.3)$, $P_\beta(\alpha^*) < 0$ and

$$
\left.\frac{\partial TE(\alpha,\beta)}{\partial \beta}\right|_{(\alpha^*,1)} > 0,
$$

thus completing the proof of (29). To see the rest of the proof, consider $\beta \in \left(1, \bar{\beta}\right)$. (The case $\beta \in \left(\hat{\beta}, 1\right)$ is specular.) Recall that, by Proposition 3, α^* and $1/\alpha^*$ are the unique global maximizers of $TE(\alpha, 1)$. By the envelope theorem and (29), the objective function increases in β for $\alpha = \alpha^*$, and it decreases in β for $\alpha = 1/\alpha^*$. Therefore, $\arg\max_\alpha TE(\alpha,\beta)$ belongs to a neighborhood of α^*, by the maximum theorem. To see uniqueness of $\arg\max_\alpha TE(\alpha,\beta)$ for $\beta \in \left(1, \bar{\beta}\right)$, recall that, for $\beta = 1$ there is a unique maximizer of $TE(\alpha, 1)$ in a neighborhood of α^*, by Proposition 3. To see that uniqueness of $\arg\max_\alpha TE(\alpha,\beta)$ is maintained for $\beta \in \left(1, \bar{\beta}\right)$, note that

$$
\left.\frac{\partial^2 TE(\alpha,\beta)}{\partial \alpha^2}\right|_{(\alpha^*,1)} < 0,
$$

as one can calculate from 40; therefore $\arg\max_\alpha TE(\alpha,\beta)$ does not bifurcate from α^* as we increase β from 1, by the implicit function theorem. ■

Proof of Proposition 15. A simple extensions of the Proof of Lemma 1 shows that

$$
\begin{aligned}
D &= \left(1+\frac{1}{\alpha_{(1,1)}}\frac{V_Y}{V_X}\frac{c_X}{c_Y}\right)^{-1}, \\
C &= \left(1+\frac{1}{\alpha_{(0,1)}}\frac{V_Y}{V_X}\frac{c_X}{c_Y}\frac{2-D}{D}\right)^{-1}, \\
B &= \left(1+\frac{1}{\alpha_{(1,0)}}\frac{V_Y}{V_X}\frac{c_X}{c_Y}\frac{1-D}{1+D}\right)^{-1}, \\
A &= \left(1+\frac{1}{\alpha_{(0,0)}}\frac{V_Y}{V_X}\frac{c_X}{c_Y}\frac{(1-C)^2 D(2-D)+B(1-D)^2(2-B)}{B^2(1-D)(1+D)+D^2(1-C)(1+C)}\right)^{-1}.
\end{aligned}
$$

Hence, TE reads

$$
\begin{aligned}
TE = \; & \tfrac{V_X}{c_X}\left(B^2\left(1-D^2\right)+D^2\left(1-C^2\right)\right)A(1-A) \\
& +\tfrac{V_Y}{c_Y}\left((1-C)^2 D(2-D)+B(1-D)^2(2-B)\right)A(1-A) \\
& +A\left(\tfrac{V_X}{c_X}\left(1-D^2\right)+\tfrac{V_Y}{c_Y}(1-D)^2\right)B(1-B) \\
& +(1-A)\left(\tfrac{V_X}{c_X}D^2+\tfrac{V_Y}{c_Y}D(2-D)\right)C(1-C) \\
& +(A(1-B)+(1-A)C)\left(\tfrac{V_X}{c_X}+\tfrac{V_Y}{c_Y}\right)D(1-D).
\end{aligned}
\tag{96}
$$

Now consider TE where $A=B=C=D=w\in[0,1]$. Then, from (96), we obtain

$$
TE=\phi(w)\equiv w^2(1-w)^2\left(\left(3+4w+2w^2\right)\frac{V_X}{c_X}+\left(9-8w+2w^2\right)\frac{V_Y}{c_Y}\right).
$$

Since

$$
\phi(w)'\big|_{w=1/2}=\frac{3}{8}\left(\frac{V_X}{c_X}-\frac{V_Y}{c_Y}\right)\neq 0,
$$

then it is not optimal to set $A=B=C=D=1/2$. ■

Proof of Proposition 16. Set $\alpha_j=1$ for any $j\geq 3$. Let $TE^{(i,k)}$ be the expected total effort conditional on reaching node (i,k) of all matches from the third onwards. We then have

$$
\begin{aligned}
u_X^{(2,0)}+u_Y^{(2,0)} &\equiv VH,\, TE^{(2,0)}=V(1-H), \\
u_X^{(1,1)}+u_Y^{(1,1)} &\equiv VL,\, TE^{(1,1)}=V(1-L), \\
u_X^{(0,2)}+u_Y^{(0,2)} &= VH,\; TE^{(0,2)}=V(1-H).
\end{aligned}
$$

We also know that

$$\begin{aligned} u_X^{(1,1)} &= u_Y^{(1,1)} = \frac{1}{2}VL, \\ u_X^{(0,2)} &= u_Y^{(2,0)} = \varepsilon VH, \\ u_X^{(2,0)} &= u_Y^{(0,2)} = (1-\varepsilon)\,VH, \end{aligned}$$

with $\varepsilon < \frac{1}{2}$, and $\frac{1}{2}L > \varepsilon H$ as established in (i) and (ii) in the Proof of Proposition 3 of Klumpp and Polborn (2006). Applying (3)—(9) we obtain:

$$\begin{aligned} \Delta u_X^{(1,0)} + \Delta u_Y^{(1,0)} &= \left(u_X^{(2,0)} - u_X^{(1,1)}\right) + \left(u_Y^{(1,1)} - u_Y^{(2,0)}\right) = (1-2\varepsilon)\,VH \\ u_X^{(1,0)} &= \left(u_X^{(2,0)} - u_X^{(1,1)}\right)B^2 + u_X^{(1,1)} = V\left(\left((1-\varepsilon)\,H - \frac{1}{2}L\right)B^2 + \frac{1}{2}L\right) \\ u_Y^{(1,0)} &= \left(u_Y^{(1,1)} - u_Y^{(2,0)}\right)(1-B)^2 + u_Y^{(2,0)} = V\left(\left(\frac{1}{2}L - \varepsilon H\right)(1-B)^2 + \varepsilon H\right) \end{aligned}$$

$$\begin{aligned} \Delta u_X^{(0,1)} + \Delta u_Y^{(0,1)} &= \left(u_X^{(1,1)} - u_X^{(0,2)}\right) + \left(u_Y^{(0,2)} - u_Y^{(1,1)}\right) = (1-2\varepsilon)\,VH \\ u_X^{(0,1)} &= \left(u_X^{(1,1)} - u_X^{(0,2)}\right)C^2 + u_X^{(0,2)} = V\left(\left(\frac{1}{2}L - \varepsilon H\right)C^2 + \varepsilon H\right) \\ u_Y^{(0,1)} &= \left(u_Y^{(0,2)} - u_Y^{(1,1)}\right)(1-C)^2 + u_Y^{(1,1)} = V\left(\left((1-\varepsilon)\,H - \frac{1}{2}L\right)(1-C)^2 + \frac{1}{2}L\right) \end{aligned}$$

$$\begin{aligned} \Delta u_X^{(0,0)} &= \left(u_X^{(2,0)} - u_X^{(1,1)}\right)B^2 + u_X^{(1,1)} - \left(\left(u_X^{(1,1)} - u_X^{(0,2)}\right)C^2 + u_X^{(0,2)}\right) \\ &= \left(u_X^{(2,0)} - u_X^{(1,1)}\right)B^2 + \left(u_X^{(1,1)} - u_X^{(0,2)}\right)\left(1-C^2\right) \\ &= V\left(\left((1-\varepsilon)\,H - \frac{1}{2}L\right)B^2 + \left(\frac{1}{2}L - \varepsilon H\right)\left(1-C^2\right)\right) \\ \Delta u_Y^{(0,0)} &= \left(u_Y^{(0,2)} - u_Y^{(1,1)}\right)(1-C)^2 + u_Y^{(1,1)} - \left(\left(u_Y^{(1,1)} - u_Y^{(2,0)}\right)(1-B)^2 + u_Y^{(2,0)}\right) \\ &= \left(u_Y^{(0,2)} - u_Y^{(1,1)}\right)(1-C)^2 + \left(u_Y^{(1,1)} - u_Y^{(2,0)}\right)\left(1-(1-B)^2\right) \\ &= V\left(\left((1-\varepsilon)\,H - \frac{1}{2}L\right)(1-C)^2 + \left(\frac{1}{2}L - \varepsilon H\right)B\,(2-B)\right) \end{aligned}$$

Using the above, we have

$$
\begin{aligned}
TE &= \left(\Delta u_X^{(0,0)} + \Delta u_Y^{(0,0)}\right) A(1-A) \\
&+A\left(\Delta u_X^{(1,0)} + \Delta u_Y^{(1,0)}\right) B(1-B) + (1-A)\left(\Delta u_X^{(0,1)} + \Delta u_Y^{(0,1)}\right) C(1-C) \\
&+ABV(1-H) + (A(1-B) + (1-A)C)V(1-L) + (1-A)(1-C)V(1-H)
\end{aligned}
$$

In equilibrium, we also have

$$
\begin{aligned}
A &= \frac{1}{1+\frac{1}{\alpha_1}\frac{\Delta u_Y^{(0,0)}}{\Delta u_X^{(0,0)}}}, \\
B &= \frac{\alpha_2}{\alpha_2 + \frac{u_Y^{(1,1)} - u_Y^{(2,0)}}{u_X^{(2,0)} - u_X^{(1,1)}}} = \frac{\alpha_2}{\alpha_2 + \frac{\frac{1}{2}L - \varepsilon H}{(1-\varepsilon)H - \frac{1}{2}L}}, \\
C &= \frac{\alpha_2}{\alpha_2 + \frac{u_Y^{(0,2)} - u_Y^{(1,1)}}{u_X^{(1,1)} - u_X^{(0,2)}}} = \frac{\alpha_2}{\alpha_2 + \frac{(1-\varepsilon)H - \frac{1}{2}L}{\frac{1}{2}L - \varepsilon H}}.
\end{aligned}
$$

As $\Delta u_Y^{(0,0)} \neq 0$, by appropriately choosing α_1, any $A \in (0,1]$ can be induced. Therefore, choosing α_1 and α_2 to maximize TE is equivalent to choosing A, B, and C, and α_2, to maximize TE. Using the above results, we obtain:

$$
\begin{aligned}
\frac{TE}{V} &= \left(\left((1-\varepsilon)H - \frac{1}{2}L\right)\left(B^2 + (1-C)^2\right) + \left(\frac{1}{2}L - \varepsilon H\right)\left(1 - C^2 + B(2-B)\right)\right) A(1-A) \\
&+ (1-2\varepsilon)H(AB(1-B) + (1-A)C(1-C)) \\
&+AB(1-H) + (A(1-B) + (1-A)C)(1-L) + (1-A)(1-C)(1-H)
\end{aligned}
$$

subject to the constraints,

$$
B\left(\alpha_2 + \frac{\frac{1}{2}L - \varepsilon H}{(1-\varepsilon)H - \frac{1}{2}L}\right) = \alpha_2, \text{ and } C\left(\alpha_2 + \frac{(1-\varepsilon)H - \frac{1}{2}L}{\frac{1}{2}L - \varepsilon H}\right) = \alpha_2.
$$

The Lagrangean is

$$
L(A,B,C,\alpha_2) = \frac{TE}{V} - \mu_B\left(B\left(\alpha_2 + \frac{\frac{1}{2}L - \varepsilon H}{(1-\varepsilon)H - \frac{1}{2}L}\right) - \alpha_2\right) - \mu_C\left(C\left(\alpha_2 + \frac{(1-\varepsilon)H - \frac{1}{2}L}{\frac{1}{2}L - \varepsilon H}\right) - \alpha_2\right).
$$

The FOC are:

$$\begin{aligned}
(A):&\quad \left(\left((1-\varepsilon)H-\tfrac{1}{2}L\right)\left(B^2+(1-C)^2\right)+\left(\tfrac{1}{2}L-\varepsilon H\right)\left(1-C^2+B(2-B)\right)\right)(1-2A)\\
&\quad +(1-2\varepsilon)H\left(B(1-B)-C(1-C)\right)+(L-H)(B+C-1) &= 0\\
(B):&\quad \left(\left((1-\varepsilon)H-\tfrac{1}{2}L\right)2B+\left(\tfrac{1}{2}L-\varepsilon H\right)2(1-B)\right)A(1-A)\\
&\quad +(1-2\varepsilon)HA(1-2B)+A(L-H)-\mu_B\left(\alpha_2+\frac{\frac{1}{2}L-\varepsilon H}{(1-\varepsilon)H-\frac{1}{2}L}\right) &= 0\\
(C):&\quad -\left(\left((1-\varepsilon)H-\tfrac{1}{2}L\right)2(1-C)+\left(\tfrac{1}{2}L-\varepsilon H\right)2C\right)A(1-A)\\
&\quad +(1-2\varepsilon)H(1-A)(1-2C)+(L-H)(A-1)-\mu_C\left(\alpha_2+\frac{(1-\varepsilon)H-\frac{1}{2}L}{\frac{1}{2}L-\varepsilon H}\right) &= 0\\
(\alpha_2):&\quad \mu_B(1-B)+\mu_C(1-C) &= 0
\end{aligned}$$

We now show that the point $A=1/2$, $\alpha_2=1$, and B and C that satisfy the constraints is a critical point of the Lagrangean. (This implies that $\alpha_1=1$ and $\alpha_2=1$ is a critical point if we view total effort as function of α_1 and α_2.)

Setting $\alpha_2=1$ implies

$$\begin{aligned}
B &= \frac{1}{1+\frac{\frac{1}{2}L-\varepsilon H}{(1-\varepsilon)H-\frac{1}{2}L}}=\frac{(1-\varepsilon)H-\frac{1}{2}L}{H(1-2\varepsilon)},\\
C &= \frac{1}{1+\frac{(1-\varepsilon)H-\frac{1}{2}L}{\frac{1}{2}L-\varepsilon H}}=\frac{\frac{1}{2}L-\varepsilon H}{H(1-2\varepsilon)}
\end{aligned}$$

and $B+C=1$. Therefore

$$\begin{aligned}
\Delta u_X^{(0,0)} &= V\left(\left((1-\varepsilon)H-\frac{1}{2}L\right)B^2+\left(\frac{1}{2}L-\varepsilon H\right)\left(1-C^2\right)\right)\\
&= V\left(\left((1-\varepsilon)H-\frac{1}{2}L\right)(1-C)^2+\left(\frac{1}{2}L-\varepsilon H\right)\left(1-(1-B)^2\right)\right)\\
&= \Delta u_Y^{(0,0)},
\end{aligned}$$

so $\alpha_1=1$ is equivalent to $A=\frac{1}{2}$. Thus, the FOC for A is satisfied. The FOC for B becomes

$$\mu_B=\left((1-\varepsilon)H-\frac{1}{2}L\right)\frac{\left((1-\varepsilon)H-\frac{1}{2}L\right)^2+\left(\frac{1}{2}L-\varepsilon H\right)^2}{\left(H(1-2\varepsilon)\right)^2}\frac{1}{2}+(L-H)\frac{(1-\varepsilon)H-\frac{1}{2}L}{(1-2\varepsilon)H}.$$

The FOC for C becomes

$$\mu_C=-\left(\frac{1}{2}L-\varepsilon H\right)\frac{\left((1-\varepsilon)H-\frac{1}{2}L\right)^2+\left(\frac{1}{2}L-\varepsilon H\right)^2}{\left(H(1-2\varepsilon)\right)^2}\frac{1}{2}+(H-L)\frac{\frac{1}{2}L-\varepsilon H}{(1-2\varepsilon)H}.$$

And the FOC for α_2 is satisfied, as we have

$$\begin{aligned}
\mu_B (1-B) + \mu_C (1-C) &= \frac{\left((1-\varepsilon) H - \frac{1}{2}L\right)^2 + \left(\frac{1}{2}L - \varepsilon H\right)^2}{H(1-2\varepsilon)} \frac{1}{2} B (1-B) + (L-H) B (1-B) \\
&\quad - \frac{\left((1-\varepsilon) H - \frac{1}{2}L\right)^2 + \left(\frac{1}{2}L - \varepsilon H\right)^2}{H(1-2\varepsilon)} \frac{1}{2} C (1-C) + (H-L) C (1-C) \\
&= 0.
\end{aligned}$$

We now check the second-order condition by building the bordered Hessian matrix of the Lagrangean, evaluated at the critical point above. We have the following second derivatives, all evaluated at $A = 1/2$ for the sake of space,

$$\begin{aligned}
\frac{\partial^2 L}{\partial A^2} &= -2\left(\left((1-\varepsilon) H - \frac{1}{2}L\right)\left(B^2 + (1-C)^2\right) + \left(\frac{1}{2}L - \varepsilon H\right)\left(1 - C^2 + B(2-B)\right)\right), \\
\frac{\partial^2 L}{\partial A \partial B} &= (1-2\varepsilon) H (1-2B) + (L-H), \\
\frac{\partial^2 L}{\partial A \partial C} &= -(1-2\varepsilon) H (1-2C) + (L-H), \\
\frac{\partial^2 L}{\partial B^2} &= (H-L)\frac{1}{2} - (1-2\varepsilon) H, \\
\frac{\partial^2 L}{\partial C^2} &= (H-L)\frac{1}{2} - (1-2\varepsilon) H, \\
\frac{\partial^2 L}{\partial B \partial \alpha_2} &= -\mu_B, \ \frac{\partial^2 L}{\partial C \partial \alpha_2} = -\mu_C, \\
\frac{\partial^2 L}{\partial A \partial \alpha_2} &= \frac{\partial^2 L}{\partial B \partial C} = 0.
\end{aligned}$$

The Jacobian of the constraints is

$$\left[\begin{array}{cccc} 0 & \frac{1}{B} & 0 & -(1-B) \\ 0 & 0 & \frac{1}{C} & -(1-C) \end{array}\right].$$

Substituting the values for B, C, μ_B, and μ_C, one can show that the sign of the determinant of the bordered Hessian is negative if and only if

$$8x^5 - 4y^5 - 20xy^4 + 32x^4 y - 5x^2 y^3 + 32x^3 y^2 > 0, \tag{97}$$

where we defined $x \equiv H - L$ and $y \equiv L - 2H\varepsilon$. Assuming $x \geq y > 0$, the left-hand side of (97) takes value $43y^5 > 0$ at $x = y$, and is increasing in x as its derivative has the same sign as

$$20x^3 + 24x^2 y - 5y^3 \geq 20x^3 + 24x^2 y - 5x^3 > 0.$$

Hence, we showed that $x \geq y > 0$ is a sufficient condition for (97). Condition $x \geq y$ is equivalent to

$$H\left(1+2\varepsilon\right) \geq 2L,$$

which is implied by $H \geq 2L$.

The next leading principal minor evaluates to

$$-\left(1-2\varepsilon\right)\frac{4\left(1-B\left(1-B\right)+\left(1-B\right)^{2}\right)H}{\left(1-B\right)^{2}B} < 0.$$

Therefore, applying part (c) of Theorem 16.4 and Theorem 19.6 in Simon and Blume (1994), we conclude that we have a saddle point, and therefore that $\alpha_1 = \alpha_2 = 1$ does not maximize TE. ■

Proof of Proposition 17. We follow a similar structure to the proof of Proposition 3. Recall (NOT). From (31) with (32)-(35), and using the definitions of $\alpha' s$ in an alternating contest (i.e., $\alpha_{(0,0)} = \alpha$, $\alpha_{(1,0)} = \alpha_{(0,1)} = 1/\alpha$, and $\alpha_{(1,1)} = 1$), then we can characterize

$$\frac{WE\left(\alpha\right)}{V} = \frac{3\alpha P_1^{WE}\left(\alpha\right)}{(\alpha+1)^2(\alpha+3)^3(3\alpha+1)^3\left(3\alpha^4+26\alpha^3+166\alpha^2+26\alpha+3\right)^3}+\frac{1}{4},$$

where

$$\begin{aligned}P_1^{WE}\left(\alpha\right) \equiv\ & 729+22842\alpha+379485\alpha^2+4034448\alpha^3+29814840\alpha^4+156934872\alpha^5+588302512\alpha^6\\ &+1522678000\alpha^7+2610776546\alpha^8+3121927900\alpha^9+2610776546\alpha^{10}+1522678000\alpha^{11}\\ &+588302512\alpha^{12}+156934872\alpha^{13}+29814840\alpha^{14}+4034448\alpha^{15}+379485\alpha^{16}+22842\alpha^{17}\\ &+729\alpha^{18}.\end{aligned}$$

As $WE\left(1\right) = 91/256$ and $\lim_{\alpha\to\infty} WE\left(\alpha\right) = \lim_{\alpha\to 0} WE\left(\alpha\right) = 1/4 < 91/256$, then $WE\left(\alpha\right)$ admits a global maximum in the interval $\alpha \in (0,\infty)$. The critical points of $WE\left(\alpha\right)$ are characterized by $\frac{\partial WE(\alpha)}{\partial\alpha} = 0$, with

$$\frac{\partial WE\left(\alpha\right)}{\partial\alpha} = \frac{3(1-\alpha)P^{WE}\left(\alpha\right)}{2(\alpha+1)^3(\alpha+3)^4(3\alpha+1)^4\left(3\alpha^4+26\alpha^3+166\alpha^2+26\alpha+3\right)^4},$$

where

$$\begin{aligned} P^{WE}(\alpha) \equiv\ & 6561\alpha^{24} + 253692\alpha^{23} + 4922208\alpha^{22} + 61592724\alpha^{21} + 541561626\alpha^{20} + 3489394788\alpha^{19} \\ & +16794128160\alpha^{18} + 59042042604\alpha^{17} + 136263542639\alpha^{16} + 123266598488\alpha^{15} \\ & -351013150336\alpha^{14} - 1294426619192\alpha^{13} - 1807569895508\alpha^{12} - 1294426619192\alpha^{11} \\ & -351013150336\alpha^{10} + 123266598488\alpha^{9} + 136263542639\alpha^{8} + 59042042604\alpha^{7} + 16794128160\alpha^{6} \\ & +3489394788\alpha^{5} + 541561626\alpha^{4} + 61592724\alpha^{3} + 4922208\alpha^{2} + 253692\alpha + 6561. \end{aligned}$$

As P^{WE} is continuous and $P^{WE}(1) < 0$, then $\alpha = 1$ is a local minimum for $WE(\alpha)$, as $\frac{\partial WE(\alpha)}{\partial \alpha} < 0$ in a left neighborhood of $\alpha = 1$, and $\frac{\partial WE(\alpha)}{\partial \alpha} > 0$ in a right neighborhood of $\alpha = 1$.

Furthermore, besides $\alpha = 1$, $\frac{\partial WE(\alpha)}{\partial \alpha} = 0$ has at most two positive roots by the Descartes' rule of signs applied to $P^{WE}(\alpha)$. One can further show that $P^{WE}(1.9) < 0$ and $P^{WE}(2) > 0$, and hence α^*_{WE}, which is the solution of $P^{WE}(\alpha) = 0$ between 1.9 and 2, is a maximum since $\frac{\partial WE(\alpha)}{\partial \alpha} > 0$ in a left neighborhood of α^*_{WE} and $\frac{\partial WE(\alpha)}{\partial \alpha} < 0$ in a right neighborhood of α^*_{WE} (see (40)). Similar considerations yield that $1/\alpha^*_{WE}$ is also a maximum. As $WE(\alpha) = WE(1/\alpha)$, both α^*_{WE} and $1/\alpha^*_{WE}$ are global maxima. The proof that the probability of victory is identical across players if and only if $\alpha = 1$ is unchanged with respect to that in Proposition 3. ■

C Appendix C: Comparing $\alpha = 1$ and $\alpha = 2$ in alternating contests

Here, we compare alternating contests with $\alpha = 1$ in Figure 4 and $\alpha = 2$ in Figure 5, where we normalize $V = 1$. The formulae of Section 3 help the reader follow the algebra. The values of $\Delta u_X^{(1,0)}$ and $\Delta u_Y^{(1,0)}$ are the differences in utilities at nodes (2,0) and (1,1), while the values of $p_X^{(1,0)}$, $u_X^{(1,0)}$ and $u_Y^{(1,0)}$ are computed using (5), (7) and (8). The values at all the other nodes are similarly computed. In Figure 5, we approximated the values in the first and second matches, without losing the qualitative features of any comparisons. At the bottom of each figure, the total effort at node (0,0) is calculated using (9), and labeled "first effect." The label "third effect" describes the total effort at node (1,1) weighted by the probability of reaching node (1,1).[37]

[37] This probability is denoted by $\Pr\{\exists (1,1)\}$.

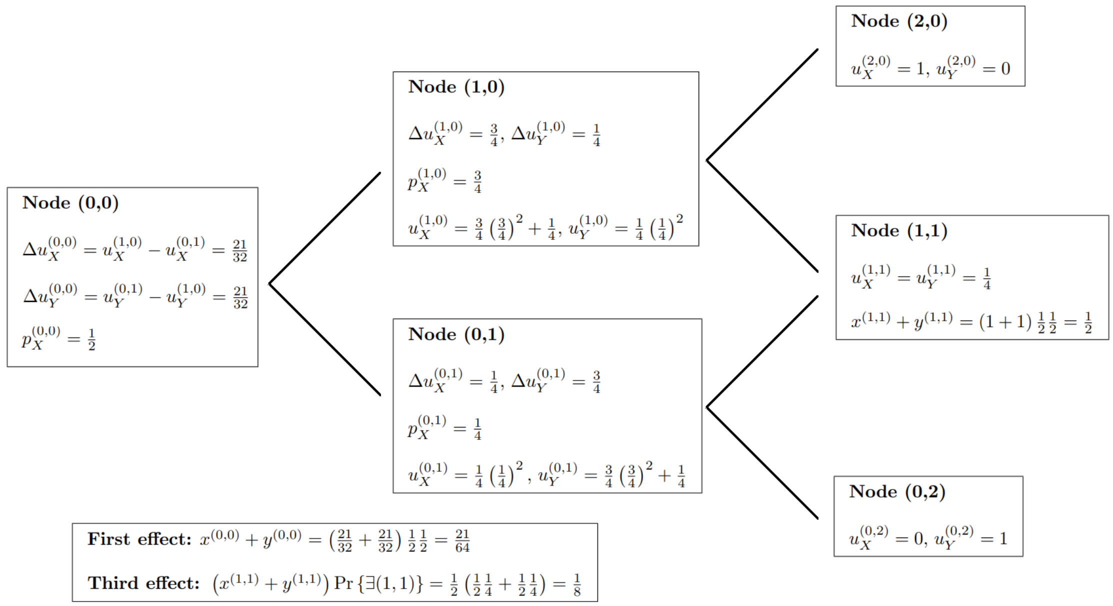

Figure 4: Fully unbiased contest with $\alpha = 1$ and $V = 1$.

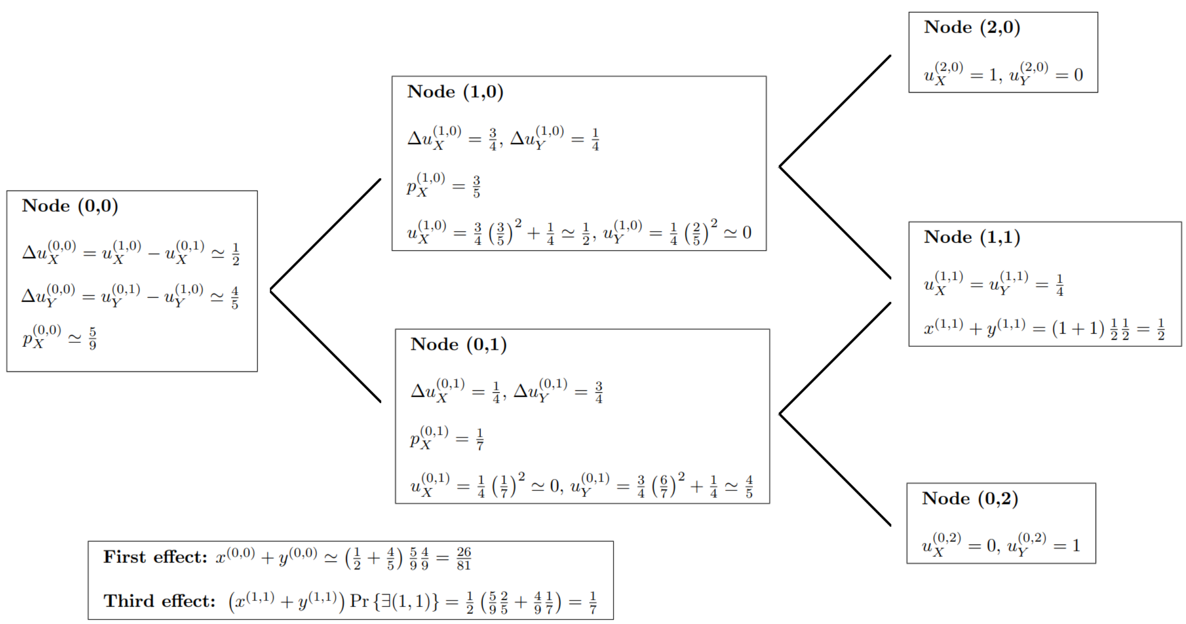

Figure 5: Alternating contest with $\alpha = 2$ and $V = 1$.

References

[1] Aoyagi, M., 2010. Information feedback in a dynamic tournament. Games and Economic Behavior 70 (2), 242–260.

[2] Badminton World Federation, 2019. Last access 19th June 2020: https://extranet.bwfbadminton.com/docs/document-system/81/1466/1470/Section%204.1%20-%20Laws%20of%20Badminton.pdf

[3] Baik, K.H., 1994. Effort levels in contests with two asymmetric players. Southern Economic Journal, pp.367-378.

[4] Baye, M.R., Kovenock, D. and De Vries, C.G., 1993. Rigging the lobbying process: an application of the all-pay auction. The american economic review, 83(1), pp.289-294.

[5] Baye, M.R., Kovenock, D. and De Vries, C.G., 1996. The all-pay auction with complete information. Economic Theory, 8(2), pp.291-305.

[6] Beviá, C., and L. Corchón, 2013. Endogenous Strength in Conflicts. International Journal of Industrial Organization 31 (3), 195–306.

[7] Bialik, C. (2016). Serving is a Disadvantage in Some Olympic Sports. FiveThirtyEight. Available at: https://fivethirtyeight.com/features/serving-is-a-disadvantage-in-some-olympic-sports/

[8] Budd, C., Harris, C. and Vickers, J., 1993. A model of the evolution of duopoly: Does the asymmetry between firms tend to increase or decrease?. The Review of Economic Studies, 60(3), pp.543-573.

[9] Chan, W., Courty, P. and Hao, L., 2009. Suspense: Dynamic incentives in sports contests. The Economic Journal, 119(534), pp.24-46.

[10] Che, Y.-K. and T. Hendershott, 2008, How to divide the possession of a football? Economics Letters, 99, pp. 561-565.

[11] Clark, D. J. and Riis, C., 1998. Contest success functions: an extension. Economic Theory, 11(1), pp.201-204.

[12] Clark, D. J. and Riis, C., 2000. Allocation efficiency in a competitive bribery game. Journal of Economic Behavior & Organization 42, 109–124.

[13] Clark, D. J., Nilssen, T. and J. Y. Sand, 2012. Motivating over time: Dynamic win effects in sequential contests. Memorandum, Department of Economics, University of Oslo, No. 28/2012, Dep. of Economics, Univ. of Oslo, Oslo

[14] Dechenaux, E., Kovenock, D., & Sheremeta, R. M., 2015. A survey of experimental research on contests, all-pay auctions and tournaments. Experimental Economics, 18(4), 609-669.

[15] Drugov, M. and D. Ryvkin, 2017. Biased Contests for Symmetric Players. Games and Economic Behavior 103, 116-144.

[16] Dukerich, J. M., M. L. Nichols, D. R. Elm, and D.A.Vollrath, 1990. Moral Reasoning in Groups: Leaders Make a Difference. Human Relations, 43, 473-493.

[17] Ederer, F., 2010. Feedback and motivation in dynamic tournaments. Journal of Econonomics & Management Strategy 19 (3), 733–769.

[18] Ely, J., Frankel, A. and Kamenica, E., 2015. Suspense and surprise. Journal of Political Economy, 123(1), pp.215-260.

[19] Epstein, G. S., Y. Mealem, and S. Nitzan, 2011. Political culture and discrimination in contests. Journal of Public Economics 95, 88-93.

[20] Esteve-González, P., 2016. Moral hazard in repeated procurement of services. International Journal of Industrial Organization 48, 244-269.

[21] Ewerhart, C., 2017. Contests with small noise and the robustness of the all-pay auction. Games and Economic Behavior, 105, 195-211.

[22] Fain, J. R., 2009. Affirmative Action Can Increase Effort. Journal of Labor Research, 30, 168–175.

[23] Feng, X., and Lu, J. 2017. Uniqueness of equilibrium in two-player asymmetric Tullock contests with intermediate discriminatory power. Economics Letters, 159, 61-64.

[24] Feng, X., and Lu, J., 2018. How to split the pie: Optimal rewards in dynamic multi-battle competitions. Journal of Public Economics 160, 82-95.

[25] Ferrall, C., and A., Smith, 1999. A sequential game model of sports championship series: theory and estimation. Review of Economics and Statistics 81, 704-719.

[26] FIVB, Fédération Internationale de Volleyball, 2020. Official Volleyball Rules 2017-2020: Approved by the 35th FIVB Congress 2016.

[27] Franke, J., 2012. Affirmative Action in Contest Games. European Journal of Political Economy 28 (1), 105-118.

[28] Franke, J., C. Kanzow, W. Leininger, and A. Schwartz, 2013. Effort Maximization in Asymmetric Contest Games with Heterogeneous Contestants, Economic Theory, 52(2), 589-630.

[29] Fu, Q., 2006. A theory of affirmative action in college admissions. Econ. Inquiry 44 (3), 420–428.

[30] Gradstein, M., 1995. Intensity of competition, entry and entry deterrence in rent seeking contests. Economics & Politics, 7(1), pp.79-91.

[31] Gradstein, M. and Konrad, K.A., 1999. Orchestrating rent seeking contests. The Economic Journal, 109(458), pp.536-545.

[32] Harris, C., and Vickers, J., 1985. Perfect equilibrium in a model of a race. The Review of Economic Studies, 52(2), 193-209.

[33] Harris, C. and Vickers, J., 1987. Racing with uncertainty. The Review of Economic Studies, 54(1), pp.1-21.

[34] Höffler, F. and D., Sliwka, 2003. Do new brooms sweep clean? When and why dismissing a manager increases the subordinates' performance. European Econonomic Review 47 (5), 877–890.

[35] Jamieson, J. P., 2010. The Home Field Advantage in Athletics: A Meta-Analysis. Journal of Applied Social Psychology 40 (7), 1819–1848.

[36] Jia, H., 2008. A stochastic derivation of the ratio form of contest success functions. Public Choice, 135(3-4), pp.125-130.

[37] Judd, K.L., 1998. Numerical methods in economics. MIT press, Cambridge: Massachusetts.

[38] Kawamura, Kohei, Moreno de Barreda, Inés, 2014. Biasing selection contests with ex-ante identical agents. Economic Letters 123 (2), 240–243.

[39] Klein, A. H., and A. Schmutzler, 2017. Optimal effort incentives in dynamic tournaments. Games and Economic Behavior, 103, 199-224.

[40] Klumpp, T. and M. K. Polborn, 2006. Primaries and the New Hampshire effect. Journal of Public Economics 90 (6-7), 1073-1114.

[41] Konrad, K.A., 2009. Strategy and Dynamics in Contests. Oxford: Oxford University Press.

[42] Konrad, K.A., and D. Kovenock, 2009. Multi-battle contests. Games and Economic Behavior 66, 256-274.

[43] Lazear, E. P. and S. Rosen, 1981. Rank-Order Tournaments as Optimum Labor Contracts. Journal of Political Economy 89 (5), 841–864.

[44] Lee, S.-H., 2013. The incentive effect of a handicap. Economics Letters, 118, 42-45.

[45] Leininger, W., 1993. More efficient rent-seeking—a Münchhausen solution. Public Choice, 75(1), pp.43-62.

[46] Lien, D.H.D., 1986. A note on competitive bribery games. Economics Letters, 22(4), pp.337-341.

[47] Lien, D.H.D., 1990. Corruption and allocation efficiency. Journal of development economics, 33(1), pp.153-164.

[48] Lizzeri, A., Meyer, M., and N. Persico, 1999. Interim evaluations in dynamic tournaments: the effects of midterm exams. Mimeo.

[49] Lizzeri, A., Meyer, M., and N. Persico, 2002. The incentive effects of interim performance evaluations. CARESS Working Paper #02-09.

[50] Mago, S. D., Sheremeta, S., and A. Yates, 2013. Best-of-three contest experiments: Strategic vs. Psychological Momentum. International Journal of Industrial Organization, 31, 287-296.

[51] Malueg, D. A., and A. J. Yates, 2006. Best-of-three contests between equally-skilled players. Working Paper

[52] 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.

[53] Mayer, W., 2004. The Basic Dynamics of Contemporary Nomination Process. In: Mayer, W. (Eds.), The making of the presidential candidate. Rowman and Littlefied Publishers, Inc, USA, 83-132.

[54] Meyer, M. A., 1991. Learning from coarse information: biased contests and career profiles. Review of Econonomic Studies 58 (1), 15–41.

[55] Meyer, M. A., 1992. Biased contests and moral hazard: implications for career profiles. Annales d'Économie et de Statistique 25/26, 165–187.

[56] Moldovanu, B. and Sela, A., 2001. The optimal allocation of prizes in contests. American Economic Review, 91(3), pp.542-558.

[57] Möller, M., 2012. Incentives versus Competitive Balance. Economics Letters 117, 505-508.

[58] Nevill, A. M., and R. L. Holder, 1999. Home advantage in sport: An overview of studies on the advantage of playing at home. Sports Medicine 28, 221–236.

[59] Nti, K.O., 1999. Rent-seeking with asymmetric valuations. Public Choice, 98(3-4), pp.415-430.

[60] Nti, K.O., 2004. Maximum efforts in contests with asymmetric valuations. European Journal of Political Economy 20, 1059–1066.

[61] Pérez-Castrillo, J. D. and Wettstein, D., 2016. Discrimination in a model of contests with incomplete information about ability. International Econonomic Review, 57, 3.

[62] Ridlon, R., and Shin, J., 2013. Favoring the winner or loser in repeated contests. Marketing Science, 32(5), 768-785.

[63] Schotter, A., and K. Weigelt, 1992. Asymmetric Tournaments, Equal Opportunity Laws, and Affirmative Action: Some Experimental Results. Quarterly Journal of Economics, 107, 511–539.

[64] Serena, M. 2017. Quality Contests. European Journal of Political Economy, 46, 15-25.

[65] Simon, C.P. and Blume, L., 1994. Mathematics for economists. New York: Norton.

[66] Skaperdas, S., 1996. Contest success functions. Economic Theory, 7, 283–290.

[67] Stein, W.E., 2002. Asymmetric rent-seeking with more than two contestants. Public Choice, 113(3-4), pp.325-336.

[68] Topsøe, F., 2006. Some bounds for the logarithmic function. Inequality theory and applications, 4, p.137.

[69] Tullock, G., 1980. Efficient Rent-seeking. In J.M. Buchanan, R.D. Tollison and G. Tullock (eds), Towards a Theory of a Rent-Seeking Society, College Station, TX: Texas A&M University Press, 97–112.

[70] Vergin, R. C. and J. J. Sosik, 1999. No place like home: an examination of the home field advantage in gambling strategies in NFL football. Journal of Economics and Business 51 (1), 21-31.

[71] Wang, Z., 2010. The optimal accuracy level in asymmetric contests. B. E. Journal of Theoretical Economics, 10 (1).