Econ Theory (2012) 51:351–382
DOI 10.1007/s00199-009-0463-z

SYMPOSIUM

The optimal multi-stage contest

Qiang Fu · Jingfeng Lu

Received: 17 April 2006 / Accepted: 7 April 2009 / Published online: 28 April 2009

Abstract This paper investigates the optimal (effort-maximizing) structure of multi-stage sequential-elimination contests. We allow the contest organizer to design the contest structure using two instruments: contest sequence (the number of stages, and the number of contestants remaining after each stage), and prize allocation. When the contest technology is sufficiently noisy, we find that multi-stage contests elicit more effort than single-stage contests. For concave and moderately convex impact functions, the contest organizer should allocate the entire prize purse to a single final prize, regardless of the contest sequence. Additional stages always increase total effort. Therefore, the optimal contest eliminates one contestant at each stage until the finale when a single winner obtains the entire prize purse. Our results thus rationalize various forms of multi-stage contests that are conducted in the real world.

We owe special thanks to Kyung Hwan Baik for the very inspiring discussion. The paper has tremendously benefited from the constructive comments and suggestions provided by an anonymous referee and coeditor Dan Kovenock. We are grateful to Atsu Amegashie, Michael Baye, Indranil Chakaraborty, Jimmy Chan, Jiahua Che, Yeon-Koo Che, John Conlon, Sudipto Dasgupta, Jeff Ely, Hanming Fang, Yuk-Fai Fong, Qiang Gong, Tanjim Hossain, Xinyu Hua, Kai Konrad, Tilman Klumpp, Lawrance Martin, Johannes Münster, Benny Moldovanu, Marco Ottaviani, Ivan Png, Aner Sela, Curtis Taylor, Lixin Ye, Hussein Yildrim, Hongjun Zhong and Wen Zhou. We would like to acknowledge the helpful comments and suggestions made by participants of 2006 Midwest Economic Theory Meeting, the 2006 PET conference, the 2007 SET conference, the 2008 Young Research Workshop on Contests (WZB, Berlin), the 2008 CESifo Venice Workshop on Contests, and the 2008 HKUST IO Workshop. The authors gratefully acknowledge the financial support from National University of Singapore [R-313-000-068-112 (Q. Fu) and R-122-000-108-112 (J. Lu)].

Q. Fu (✉)
Department of Business Policy, National University of Singapore,
1 Business Link, Singapore 117592, Singapore
e-mail: bizfq@nus.edu.sg

J. Lu
Department of Economics, National University of Singapore,
10 Kent Ridge Crescent, Singapore 119260, Singapore
e-mail: ecsljf@nus.edu.sg

Keywords Multi-stage contest · Effort maximization · Prize allocation · Winner-take-all · Sequential elimination

JEL Classification C7 · D7 · J3

1 Introduction

A wide range of competitive events can be viewed as contests. In such contests, economic agents expend scarce resources in order to win a limited number of prizes. These contests appear in a diverse array of areas including influence politics, sports, R&D races, college admissions and even labor market competition within firms. Due to the ubiquity of contests, a huge body of economic literature has been developed to uncover the various strategic aspects of these activities.[1] A large proportion of the existing literature views a contest as a static battle, under the conventional assumption that a contestant can accomplish success in a single stroke (one-shot effort). In reality however, many contests last several stages, and require contestants to endure a long line of shots before they make the win.

Numerous examples are available to illustrate the multi-stage nature of most contests. One of them is the "election of London" to host the 2012 Summer Olympic Games: While nine cities initially submitted applications, only five (London, Madrid, Moscow, New York and Paris) were shortlisted for the final election. In research tournaments, the procurement firms select the most attractive ideas from a larger pool of proposals, and only the selected are eligible for further development.[2] When recruiting new faculty members, economics departments similarly interview a large number of candidates, but extend on-campus visit invitations only to a small subset. In the early stages of all the aforementioned scenarios, contestants strive mainly to avoid elimination.

Central to the contest literature is the question of how the structure or the rule of the contest impacts the total effort supplied by contestants. As argued by Gradstein and Konrad (1999), "…contest structures result from the careful consideration of a variety of objectives, one of which is to maximize the effort of contenders". This study follows in this strand of the literature, and explores effort-maximizing contest designs in a setting that involves sequential elimination. We recognize that the equilibrium level of effort supplied in this particular context depends on two major structural elements: *the contest sequence* (which indicates the number of stages, and the number of remaining contestants in each stage), and *the prize allocation rule* (which spreads the prize money among a set of prizes of differing ranks). A number of questions naturally arise out of such an exploration: Firstly, given the sequence of a multi-stage contest and a fixed sum of prize money, does a "winner-take-all" contest (that concentrates the entire budget on the top prize awarded to the final winner) necessarily dominate a contest that awards intermediate prizes or a second prize? Secondly, does a multi-stage contest (that successively eliminates contestants) generate more effort than a

[1] See Konrad (2007) for a thorough survey of theoretical work on contests.

[2] See Fullerton and McAfee (1999).

single-stage contest that demands only one-shot effort from each contestant? Thirdly, how are the sequence of the contest and the allocation of prize money intertwined in the way they impact the output of a contest?

These questions are particularly intriguing in the context of imperfectly discriminatory contests with complete information. The standard Tullock model, as well as its numerous variations, has been widely employed in the literature to model competitive events where factors other than a contestant's autonomous effort contribute to one's success. Unlike all-pay auctions with complete-information that fully dissipates the rent (when contestants are identical), contestants in a Tullock contest could end up with a positive surplus in one-shot competition—as long as the contest success function is sufficiently "noisy" in converting the effort into winning probabilities. Because a one-shot contest does not fully dissipate the rent, there is room in the context outlined here for sophisticated contest design that will provide additional incentives for effort.

We attempt to answer these questions in a unified framework by considering N identical contestants, who are eligible for a fixed total of prize money Γ_0. The contestants are successively eliminated from the race through L stages,[3] and survivors in each stage compete against all others to advance further in the competition. We allow the contest organizer to maximize the total effort by choosing the optimal contest sequence and optimally allocating the prize money: Non-negative prizes could be available to contestants in any stage prior to the finale, as well as to all finalists.

As one could imagine, outcomes in various competitive events (such as sports and art performance competition) could be influenced by not only autonomous efforts, but also many other factors. In the process of determining winners, in order to mimic a setting of multiple-prize contests with sufficient "noise" (imperfectly discriminatory contests), we model the competition in each stage as a multi-winner nested contest as suggested by Clark and Riis (1996, 1998a). This framework extends the basic Tullock contest model and allows a block of prizes to be awarded to winners. Following Clark and Riis (1998a), we measure the impact of a contestant on the winning probabilities through a power function of his/her effort. We focus mainly on settings with a concave impact function that exhibits decreasing return on effort.[4] The results are summarized as follows:

1. *"Hierarchical Winner-Take-All Property": it is optimal to award the entire prize purse (only) to the contestant who wins the first prize in the finale, regardless of the sequence of the contest.*
2. *Given that the entire prize purse is concentrated on the top of the hierarchical ladder, inserting an additional stage of elimination always increases the total effort supplied by the contestants, regardless of the existing contest sequence.*

As a consequence, the optimal contest that maximizes the total effort is organized as a $(N-1)$-stage "Pyramid" contest that eliminates one contestant at each stage, and allows a single final winner to earn the entire prize purse. Thus, our study provides rationales for (i) the multi-stage contests widely observed in reality; and (ii) the

[3] The contest reduces to a single-stage one when $L = 1$.

[4] The concave impact functions are then modeled as power functions with exponential terms less than or equal to one.

winner-take-all principle commonly assumed in modeling rent-seeking competition in a more general setting. To better understand the underlying logic of our results, and to further explore the robustness and the limit of the generality of our main results, we extend the basic model to allow for increasing return to effort.[5] We find that multi-stage contests could still prevail for moderately convex impact function. In addition, we investigate whether stochastically eliminating one or more contestants within any stage could elicit more effort. We show that in our setting, a stochastic elimination rule does not increase the total supply of effort.

1.1 The relation to the literature

Baye et al. (1993) raise the question of "why politicians frequently 'announce' that they have narrowed down a set of potential recipients of a 'prize' to a slate of finalists?" While the process of shortlisting modeled by Baye et al. (1993) does not involve any activity on the part of contestants, an increasing number of papers do in fact allow contestants to compete actively for advancement toward the finale through a number of stages.

The usual way of modeling the elimination/shortlisting process in multi-stage contests is to divide remaining contestants into groups such that a single winner stands out from each group. This strand of literature can be traced back to the seminal paper of Rosen (1986). He considers a 2^N-contestant N-stage sequential contest: In each stage, two of the remaining contestants are matched for head-to-head confrontation, and the winner of each pairwise battle survives into the next stage. Assuming an imperfectly discriminatory contest technology,[6] Rosen (1986) searches for a reward scheme that maintains incentives along the ladder of the given hierarchical structure, and shows that a disproportionate share of prize money should be allocated to the top prize. Our paper is more closely related to studies that explore the optimal (effort-maximizing) structure of multi-stage contests with complete information. While Rosen (1986) treats the hierarchical contest as exogenously given, Gradstein and Konrad (1999) consider the contest structure to be the endogenous choice of the contest organizer. The contest organizer is allowed to decide how many stages the contest includes, and how remaining contestants are matched at each stage. They show that the optimality of a contest design crucially depends on how discriminatory the (Tullock) contest is: an N-stage contest emerges as the optimum when the contest is sufficiently noisy, i.e., the impact function $f(e) = e^r$ in the ratio-form contest success function is concave with an exponent $r \leq 1$; while a single-stage contest prevails when $r \geq 1$.[7,8,9]

[5] In this paper, "increasing return" refers to impact functions that are modeled as power functions with exponents greater than 1.

[6] Rosen (1986) allows for a general impact function.

[7] In Gradstein and Konrad (1999), multi-stage contests and static contests elicit the same amount of total effort when $r = 1$.

[8] Other papers contributing to the research agenda on multi-stage contests include Amegashie (1999, 2000), Harbaugh and Klumpp (2005), Matros (2005), and Konrad and Kovenock (2006).

[9] Moldovanu and Sela (2006) compare two-stage contests to static contests in an all-pay auction setting with incomplete-information. They find that a two-stage contest elicits more effort if the effort-cost function is convex.

A distinct feature of our analysis is that we consider an elimination process that "pools" rather than "matches" competitors in each preliminary stage. In contrast to the aforementioned papers in which the contest consists of a series of single-winner battles (in each battle a contestant meets a subset of remaining competitors), we let remaining contestants confront "all others" in each stage to determine the set of survivors for the next round.[10] The competition in each stage can be modeled as a multiple-winner contest. A handful of papers have examined contests that involve more than one prize recipient. We adopt the multiple-winner multiple-loser nested Tullock contest as suggested by Clark and Riis (1996, 1998a). This framework is also adopted by Yates and Heckelman (2001), Amegashie (2000) and Fu and Lu (2008a) to model imperfectly discriminatory contests. By contrast, Fu and Lu (2008b) and Barut and Kovenock (1998) have thoroughly analyzed multiple-winner perfectly discriminatory contests (all-pay auctions).[11]

Amegashie (2000) compares the two elimination procedures ("pooling" contestants and "matching" contestants) in a two-stage Tullock contest with linear contest technology, and concludes that the procedure of "pooling competition" generates more effort.[12] Our paper allows for a more general contest technology. We endogenize the contest sequence and search for the optimal sequence of pooling competitions such that total effort is maximized. We show that a "Pyramid" contest could emerge as the optimum design, even if the impact function is moderately convex.

This paper is linked also to the literature on optimal prize allocation in which the reward structure of a contest is treated as an endogenous choice. Most of these papers assume that the contest organizer attempts to maximize overall effort. For this purpose, Krishna and Morgan (1998) justify the winner-take-all principle in small two-stage tournaments that concentrate the entire prize purse on the first prize in the top rank. Similar results are obtained by Matros (2005) in a two-stage Tullock contest setting. In addition, Moldovanu and Sela (2001) examine the optimal allocation of a prize budget in a one-stage incomplete information all-pay auction. They find that a winner-take-all contest maximizes the total effort for concave or linear cost functions. Our paper departs from these studies in that the contest organizer is endowed with the flexibility to design the contest by simultaneously choosing both the elimination (the contest sequence) and prize allocation rules.

Our paper proceeds as follows. Section 2 sets up the model. In Sect. 3, we present our results on optimal contest structures with concave impact function ($r \leq 1$ for impaction function $f(e) = e^r$). In Sect. 4, we extend the basic setting by allowing for a convex impact function ($r > 1$) and for stochastic elimination in order to check for the robustness and generality of the results established in Sect. 3. We further discuss possible future extensions to conclude the paper.

[10] The aforementioned competition to host the Olympic games may serve as one example of a contest that involves the pooling of contestants in the early stages.

[11] In contrast to Clark and Riis (1998b), Barut and Kovenock (1998), Moldovanu and Sela (2001) and Moldovanu et al. (2007) investigate multiple-winner all-pay auctions with incomplete information.

[12] Fu and Lu (2008a) also provide theoretical evidence supporting the fact that a "pooling" competition elicits more effort than any split contests.

2 The model

2.1 Setup

There are $N(\geq 3)$ risk-neutral contestants involved in a multi-stage contest. The contest organizer has a total budget of Γ_0 for prize allocation. Let L denote the number of stages in the contest, and N_l denote the number of contestants in a stage $l \in \{1, 2, \ldots, L\}$, with $N_1 \equiv N$. In each "preliminary" stage $l \in \{1, 2, \ldots, L-1\}$, N_l contestants participate, and N_{l+1} of them survive and proceed to the next stage. Besides the "tickets" to the next round, N_l nonnegative intermediate prizes $W_l^m, m \in \{1, \ldots, N_l\}$ are awarded in each stage $l \in \{1, 2, \ldots, L-1\}$. In the "final" stage L, N_L surviving contestants compete for N_L nonnegative final prizes $W_L^m, m \in \{1, \ldots, N_L\}$.[13] The sequence of a contest is therefore represented by a L-term non-increasing sequence $\{N_1, N_2, \ldots, N_L\}$ with $N_1 = N \geq N_2 \geq \cdots \geq N_L \geq 1$, which indicates the number of stages and the number of remaining contestants in each stage.

In this study, we consider the competition in each stage to be a multiple-winner contest where each contestant competes against all others. Appropriately then, we adopt the multiple-winner nested contest model suggested by Clark and Riis (1996, 1998a). A number of prizes are given away in each stage $l \in \{1, 2, \ldots, L\}$, and each remaining contestant is eligible for only one prize. The contestants simultaneously exert their one-shot efforts $e_l^i, i = 1, 2, \ldots, N_l$ to increase their probabilities of winning. Once a contestant is selected as a winner for a prize, he/she is immediately removed from the pool, while the rest of the contestants are readied for the next draw. The effort levels of the contestants remain fixed at their chosen levels while the draws are taking place. The same set of effort entries (although excluding those of previous winners) thus continue to be utilized to select the subsequent winner. This procedure is repeated until the last prize is given away.[14]

We assume that in each stage of the contest, the "tickets" to the next stage and the stage prizes are allocated in a sequential lottery process. We define Ω_l^m to be the set of remaining contestants up for the mth draw in stage l, with $m \in \{1, 2, \ldots, N_l\}$. Denote $(e_l^1, \ldots, e_l^{i-1}, e_l^{i+1}, \ldots, e_l^{N_l})$ by $\mathbf{e}_l^{-i}$. We assume that the conditional probability that a contestant $i \in \Omega_l^m$ is selected in the mth draw is given by the ratio-form contest success function

$$p\left(e_l^i, \mathbf{e}_l^{-i}; \Omega_l^m\right) = \left(e_l^i\right)^r \Big/ \sum_{j \in \Omega_l^m} \left(e_l^j\right)^r, \quad r \in \left(0, \frac{N}{N-1}\right], \tag{1}$$

where the *impact function* $f(e_i) = (e_l^i)^r$ represents the contestants' output in the contest, with r measuring the marginal impact of an increase in the contestants' effort. This particular form of success function, which projects effort entries into winning odds, was axiomatized by Skaperdas (1996). It has been widely employed in the literature

[13] Some final prizes are allowed to be zero.

[14] Although this multiple-winner nested contest can be conveniently viewed as a sequential-lottery contest, Fu and Lu (2008b) find that a unique (simultaneous) noisy ranking process exists that underpins this model.

to model imperfectly discriminatory contests, in which the contestant who puts in the greatest amount of effort may not end up winning, although such an effort would no doubt increase his or her chances. The higher the r, the less noisy the selection process. When $r \geq \frac{N}{N-1}$, a single stage winner-take-all contest fully dissipates the rent, and is thus seen as the optimal contest. We thus consider the more interesting case of $r \in (0, \frac{N}{N-1})$.

If all contestants who go up for a draw make zero effort, we assume that the winner would be randomly picked from the pool. Moreover, we assume that if Ω_l^m is reduced to a singleton, i.e., only one contestant is up for the mth draw, then the only contestant automatically wins. At stage $l \in \{1, 2, \ldots, L\}$, the contestant selected in the mth draw is awarded the prize W_l^m. In addition, in a "preliminary" stage $l \in \{1, 2, \ldots, L-1\}$, the contestants who are selected in the first N_{l+1} draws proceed to the $(l+1)$th stage, while the other $N_l - N_{l+1}$ contestants are eliminated. We define $\Gamma_l \equiv \sum_{m=1}^{N_l} W_l^m$ to be the sum of prizes awarded in stage l, and $\Gamma \equiv \sum_{l=1}^{L} \Gamma_l$ to be the entire prize purse for the contest.

2.2 Symmetric equilibrium

For the sake of tractability, we focus on the symmetric equilibria in which all remaining contestants follow the same strategy in each stage of the contest. We solve this symmetric equilibrium through backward induction. Denoted by V_l is the conditional (symmetric) equilibrium expected payoff of a representative contestant at stage l. For the sake of descriptive convenience, we define $V_{L+1} \equiv 0$, $N_{L+1} \equiv 1$. At stage $l \in \{1, 2, \ldots, L\}$, a remaining contestant i rationally chooses his effort e_l^i to maximize his expected payoff

$$V_l^i = \sum_{m=1}^{N_{l+1}} \left[P_m\left(e_l^i, \mathbf{e}_l^{-i}\right)\left(V_{l+1} + W_l^m\right) \right] + \sum_{m=N_{l+1}+1}^{N_l} \left[P_m\left(e_l^i, \mathbf{e}_l^{-i}\right) W_l^m \right] - e_l^i, \quad (2)$$

where $P_m(e_l^i, \mathbf{e}_l^{-i})$ is the probability that contestant i is selected in the mth draw.

In a symmetric Nash equilibrium, all remaining contestants choose the same effort outlay e_l (nonnegative). For ease of notation, we define Φ_l as $\Phi_l \equiv \sum_{m=1}^{N_{l+1}}[(1 - \sum_{g=0}^{m-1} \frac{1}{N_l - g})(V_{l+1} + W_l^m)] + \sum_{m=N_{l+1}+1}^{N_l}[(1 - \sum_{g=0}^{m-1} \frac{1}{N_l - g}) W_l^m]$. The Kuhn–Tucker conditions state that if $e_l > 0$, the following must be satisfied

$$\frac{r\Phi_l}{N_l e_l} - 1 = 0. \quad (3)$$

A unique symmetric interior equilibrium $e_l = \frac{r\Phi_l}{N_l}$ exists if $\Phi_l \geq 0$; while a symmetric corner solution equilibrium with $e_l = 0$ would emerge otherwise. In a symmetric equilibrium, every contestant has the same chance of winning each component of the total stage-award $N_{l+1}V_{l+1} + \Gamma_l$ (including N_l stage prizes and N_{l+1} tickets to the next stage). Therefore, the conditional equilibrium expected payoff of a representative contestant at stage l is $V_l = (N_{l+1}V_{l+1} + \Gamma_l)/N_l - e_l$, where e_l is his/her equilibrium effort at stage l. We summarize these results in the following proposition.[15]

Proposition 1 *In the symmetric subgame perfect Nash equilibrium of the contest, every remaining contestant in stage $l \in \{1, 2, \ldots, L\}$ exerts an effort*

$$e_l = \begin{cases} \frac{r\Phi_l}{N_l} & \text{if } \Phi_l \geq 0, \\ 0 & \text{if } \Phi_l < 0, \end{cases} \tag{4}$$

and the payoff V_l of a representative contestant at stage l is

$$V_l = \frac{N_{l+1}V_{l+1} + \Gamma_l}{N_l} - e_l. \tag{5}$$

Proof See Appendix. □

From (4) and (5), e_l and V_l can be solved recursively from the last stage of the game as we have $V_{L+1} \equiv 0$, $N_{L+1} \equiv 1$. Note that when $r \in (0, \frac{N}{N-1}]$, it is easy to verify that every V_l is nonnegative.[16] Thus, the sign of Φ_l depends on the prize allocation rule $\{W_l^m\}_{m=1}^{N_l}$ in stage l. The terms $1 - \sum_{g=0}^{m-1} \frac{1}{N_l - g}$ strictly decreases with m, i.e., the order of the draws. In addition, $\sum_{m=1}^{N_l}(1 - \sum_{g=0}^{m-1} \frac{1}{N_l-g})$ always equals zero.[17] It implies that the coefficient $1 - \sum_{g=0}^{m-1} \frac{1}{N_l-g}$, which is assigned to a prize, is positive for "earlier" draws, but turns negative when m is sufficiently large. Thus, when sufficiently large prizes are awarded for the latest draws, a corner equilibrium solution may emerge, as the contestants would prefer not to make positive effort (which increase the odds of winning earlier prizes), but to wait for the more generous prizes awarded in last draws.

However, we now argue that for the purpose of eliciting effort, there is no loss of generality to restrict our attention to only prize allocations that render interior equilibria, i.e., $\Phi_l \geq 0$ holds and the equilibrium solution $e_l = \frac{r\Phi_l}{N_l}$ applies. To illustrate, suppose the current prize allocation $\{W_l^m\}_{m=1}^{N_l}$ makes $\Phi_l < 0$. Recall that $1 - \sum_{g=0}^{m-1} \frac{1}{N_l-g}$ decreases with m and $\sum_{m=1}^{N_l}(1 - \sum_{g=0}^{m-1} \frac{1}{N_l-g})$ is always equals to zero. Thus, Φ_l would strictly increase if the prize money is hypothetically shifted from later prizes toward the first prize W_l^1. Because Φ_l is affine in $\{W_l^m\}_{m=1}^{N_l}$, by the continuity argument, an alternative prize allocation $\{\tilde{W}_l^m\}_{m=1}^{N_l}$ always exists, which delivers exactly $\Phi_l = 0$. It reinstates the formula for an interior equilibrium of $e_l = \frac{r\Phi_l}{N_l} = 0$, while the equilibrium outcome is equivalent to that which arises from the initial prize allocation. In particular, the alternative prize allocation $\{\tilde{W}_l^m\}_{m=1}^{N_l}$ does not alter the contestants' expected equilibrium payoff V_l in stage l, which represents the value of a "ticket" to

[15] Please refer to the Appendix for detailed proof of these results. These findings are consistent with those of Clark and Riis (1998a).

[16] Detailed proof is available from the authors upon request.

[17] To see this, rewrite $\sum_{m=1}^{N_l}(1 - \sum_{g=0}^{m-1} \frac{1}{N_l-g})$ as $N_l - \sum_{g=0}^{N_l-1} \frac{N_l-g}{N_l-g} = 0$.

contestants in stage $l-1$. This means that the alternative prize allocation $\{\tilde{W}_l^m\}_{m=1}^{N_l}$ does not provide contestants in previous stages with a different incentive from that provided by the initial allocation rule $\{W_l^m\}_{m=1}^{N_l}$. Hence, we can ignore all prize allocations that yield $\Phi_l < 0$, and rely instead on the interior equilibrium of $e_l = \frac{r\Phi_l}{N_l}$ for the rest of this paper.

We assume that the effort accrues to the benefit of the contest organizer. Thus, the contest organizer chooses the optimal sequence $\{N_l\}_{l=1}^{L}$ and prize allocation $\{W_l^m | m = 1, \ldots, N_l; l = 1, \ldots, L\}$ to maximize the total effort $E = \sum_{l=1}^{L} N_l e_l$, subject to budget constraints

$$\Gamma \leq \Gamma_0, \tag{6}$$

and

$$\Phi_l \geq 0, \quad l = 1, \ldots, L. \tag{7}$$

As we have shown above, in a symmetric equilibrium, each remaining contestant at stage l expects to receive a payoff $V_l = (N_{l+1}V_{l+1} + \Gamma_l)/N_l - e_l$. Thus, we have

$$N_l e_l = N_{l+1}V_{l+1} + \Gamma_l - N_l V_l. \tag{8}$$

Lemma 1 $E = \Gamma - NV_1$.

Proof Summing up (8) over the L stages gives

$$E \equiv \sum_{l=1}^{L} N_l e_l = \sum_{l=1}^{L} (N_{l+1}V_{l+1} - N_l V_l) + \sum_{l=1}^{L} \Gamma_l = \Gamma - NV_1.$$

□

Lemma 1 is fairly intuitive, but is of essential importance in this paper. V_1 not only represents the payoff a representative contestant expects to receive in the first stage, it also represents the payoff every contestant expects in the very beginning from the entire series of competitions. Thus, NV_1 represents the total surplus available to the N contestants, which, by the risk neutrality of the contestants, is equivalent to the difference between the prize purse Γ and the total effort E. Obviously, the optimal contest structure that maximizes the total effort must minimize the net rent received by the contestants. Thus, Lemma 1 allows us to focus on the minimization of V_1 in the subsequent analysis. This analysis is equivalent to the original effort maximization problem that was detailed in the introduction.

Using (4) and (5), we are able to recursively solve for V_1, viewing it as a function of the prizes in the current and all future stages. The proof is omitted here for the sake of brevity, but will be made available upon request.

Lemma 2 $V_1 = \sum_{l=1}^{L}\{(\prod_{j=1}^{l}\frac{1}{N_j})(\prod_{j=1}^{l-1}[N_{j+1}(1-r)+r\sum_{g=0}^{N_{j+1}-1}\frac{N_{j+1}-g}{N_j-g}])$ $(\sum_{m=1}^{N_l}W_l^m[(1-r)+r\sum_{k=1}^{m}\frac{1}{N_l-k+1}])\}$.[18]

Lemma 2 writes V_1 as a function affine in the set of prizes $\{W_l^m\}$. Combining Lemmas 1 and 2, we obtain the following result. Again, the proof is omitted for brevity.

Proposition 2 *A N-person L-stage sequential-elimination contest, with a sequence $\{N_l\}_{l=1}^{L}$, a prize allocation $\{W_l^m|m=1,\ldots,N_l;l=1,\ldots,L\}$ and a total prize purse Γ, generates a total effort*

$$E=\Gamma-N\sum_{l=1}^{L}\left\{\left(\prod_{j=1}^{l}\frac{1}{N_j}\right)\left(\prod_{j=1}^{l-1}\left[N_{j+1}(1-r)+r\sum_{g=0}^{N_{j+1}-1}\frac{N_{j+1}-g}{N_j-g}\right]\right)\right.$$
$$\left.\times\left(\sum_{m=1}^{N_l}W_l^m\left[(1-r)+r\sum_{k=1}^{m}\frac{1}{N_l-k+1}\right]\right)\right\}.$$

3 Concave impact functions ($r \leq 1$)

In this section, we focus on the concave impact function $f(e)=e^r$ with $r\in(0,1]$. We first establish the optimality of the "hierarchical winner-take-all" principle, which requires that all the prize money be allocated to the winner of the finale of each contest sequence. We then establish that a full-sequence elimination process is optimal among all possibilities, and that this requires exactly one contestant to be eliminated in each stage until the last survivor receives the top prize.

3.1 The optimal prize allocation

Given the sequence of the contest, a contestant's equilibrium surplus V_1 depends purely on the allocation of the prize purse. Lemma 2 shows that V_1 can be written as a weighted sum of the prizes awarded in the contest. Based on this result, we investigate the optimal prize allocation rule that minimizes V_1 and we first obtain the following proposition.

Proposition 3 *When $r\in(0,1]$, the optimal prize allocation requires $W_l^1=\Gamma_l$ and $W_l^m=0$, $\forall m>1$, $l\in\{1,2,\ldots,L\}$.*

Proof We denote the weight on W_l^m by D_l^m, i.e.,
$D_l^m \equiv (\prod_{j=1}^{l}\frac{1}{N_j})\ (\prod_{j=1}^{l-1}[N_{j+1}(1-r)+r\sum_{g=0}^{N_{j+1}-1}\frac{N_{j+1}-g}{N_j-g}])\ [(1-r)+r\sum_{k=1}^{m}$ $\frac{1}{N_l-k+1}]$. Within a stage l, D_l^m contains the common component

[18] We define $\prod_{j=1}^{0}[N_{j+1}(1-r)+r\sum_{g=0}^{N_{j+1}-1}\frac{N_{j+1}-g}{N_j-g}]\equiv 1$.

$(\prod_{j=1}^{l}\frac{1}{N_j})(\prod_{j=1}^{l-1}[N_{j+1}(1-r)+r\sum_{g=0}^{N_{j+1}-1}\frac{N_{j+1}-g}{N_j-g}])$, but differs in the other component $[(1-r)+r\sum_{k=1}^{m}\frac{1}{N_l-k+1}]$. It is straightforward to see that $\sum_{k=1}^{m}\frac{1}{N_l-k+1}$ strictly increases with m, which leads to that D_l^m strictly decreases with m, the index for the order of the draw. Thus, to minimize V_1, all the prize money available at stage l should be allocated only to the first prize W_l^1. □

Proposition 3 states that in each stage l, zero prize money should be assigned to W_l^j for $j \geq 2$. This result is, in essence, not different from the standard "winner-take-all" result in static contests. Obviously, in each stage of the contest, winning a higher-ranked prize demands more effort from contestants. A more generous second prize would weaken contestants' incentive to compete for the first prize, and encourage them to wait for the next draw. Assume that a prize purse Γ_l is allocated to a stage l. *Within* stage l, the contest organizer can encourage contestants to exert more effort by concentrating the entire prize purse for this stage on the first prize, i.e., $W_l^1 = \Gamma_l$.[19]

How to allocate the prize purse Γ_0 across stages remains to be investigated. Define $\theta_l \equiv (\prod_{j=1}^{l}\frac{1}{N_j})(\prod_{j=1}^{l-1}[N_{j+1}(1-r)+r\sum_{g=0}^{N_{j+1}-1}\frac{N_{j+1}-g}{N_j-g}])[(1-r)+\frac{r}{N_l}]$, which is the coefficient of W_l^1 in the expression of V_1. According to Proposition 3, we can write V_1 as the weighted sum of $\{\Gamma_l\}$

$$V_1 = \sum_{l=1}^{L}\theta_l W_l^1, \tag{9}$$

where $W_l^1 = \Gamma_l$ for all $l \in \{1, 2, \ldots, L\}$.

The weight θ_l represents the net rent one unit of prize W_l^1 contributes to a contestant. To minimize V_1 in a given contest sequence, the prize purse must be concentrated on the prize with the smallest weight. We show in the Appendix that θ_l decreases with l when $r \in (0, 1]$.[20] We thus establish the following "Hierarchical Winner-Take-All" principle for the optimal prize allocation.

Theorem 1 *"Hierarchical Winner-Take-All" principle: When* $r \in (0, 1]$, *for any given contest sequence* $\{N_l\}_{l=1}^{L}$, *it is optimal to allocate the entire prize purse to the first prize in the final stage, i.e.,* $W_L^1 = \Gamma_0$.

Proof See Appendix. □

From Theorem 1, it is clear that a multi-stage contest that maximizes the total effort must combine all the resource into a single final prize and reward it to a single final winner. The rule applies regardless of the sequence of the contest. The weight θ_l decreases with l throughout the sequence, which implies that a prize at a higher rung along the hierarchical ladder contributes less surplus to a contestant. Intuitively, a prize at a higher rank in the hierarchy could demand more effort from a contestant,

[19] This effect is consistent with the "winner-take-all" principle established by Clark and Riis (1998a). It is easy to verify that Proposition 3 still holds when $r \in (1, \frac{N}{N-1})$.

[20] θ_l is strictly decreasing with l when $r \in (0, 1]$ if N_l is strictly decreasing.

because he has to repeatedly exert his effort to climb toward this higher rank. Thus, all resources should be allocated to the top prize to induce the greatest level of total effort. As argued by Rosen (1986), a greater top prize "effectively extends the players' horizon".[21] Our results provide a rationale for the commonly assumed "winner-take-all" principle in a multi-stage contest setting.

Theoretically however, moving prize money from a preliminary stage to the finale generates two effects. On one hand, it increases the effort contestants would put into the competition after this stage. On the other hand, it reduces the payoff to the contestants in that stage as well as all stages prior to it, and therefore reduces the effort that is put into those stages. Nevertheless, when the impact function $f(e) = e^r$ exhibits decreasing returns to effort ($r \leq 1$), the increase in effort generated by the former effect dominates the drop in effort due to the latter.

3.2 The optimality of multi-stage contests

Having established the "hierarchical winner-take-all" principle as the unique optimal prize allocation rule, the optimal sequence of the contest (the number of stages and the number of remaining contestants in each stage) that maximizes the total effort remains to be established. Theorem 1 indicates that the allocation of the prize purse is independent of the contest sequence when $r \in (0, 1]$. We therefore restrict our attention to the contest structure with a single prize $W_L^1 = \Gamma_0$. Rewriting V_1 by setting all prizes other than W_L^1 at zero yields

$$V_1 = \frac{\Gamma_0}{\prod_{j=1}^{L} \frac{1}{N_j}} \left\{ \prod_{j=1}^{L-1} \left[N_{j+1}(1-r) + r \sum_{g=0}^{N_{j+1}-1} \frac{N_{j+1}-g}{N_j - g} \right] \right\} \left[(1-r) + r\frac{1}{N_L} \right]. \tag{10}$$

For description convenience, we assume that all candidates for the optimal contest sequence end with $N_L = 1$ without loss of generality. By the optimality of the "hierarchical winner-take-all" principle in prize allocation, a single winner stands out in the last stage of competition and takes over the entire prize purse. Thus, a L-stage contest with $N_L > 1$ is equivalent to a hypothetical $(L + 1)$-stage contest represented by the sequence $\{\{N_l\}_{l=1}^{L}, 1\}$, i.e., $N_{L+1} = 1$. In the hypothetical $(L + 1)$-stage contest, one contestant is selected at stage L to enter stage $L + 1$, but does not receive a tangible prize in that stage. Thus, the "last man standing" in stage $L + 1$ would automatically win Γ_0 without exerting any additional effort. Consequently, we have $e_{L+1} = 0$ and $V_L = \Gamma_0$, and Eq. (11) continues to apply. Thus, in the following analysis, we consider only contest sequences with $N_L = 1$.[22] We further assume that the

[21] In the setting of Rosen (1986), the contest organizer has a different objective from that in our paper. However, the intuition underlying his results applies in our scenario.

[22] We construct this hypothetical sequence for technical convenience and generality. In the subsequent analysis, we ask the question whether inserting a stage between two consecutive stages would increase total effort. Without constructing this (equivalent) sequence with $N_L = 1$, the analysis does not immediately apply when we have more than two contestants remain in the last stage of competition and one more stage is added to further narrow the set of finalists.

contest sequence $\{N_l\}_{l=1}^{L}$ is strictly decreasing. Obviously, when $N_l = N_{l+1}$, stage l would not elicit positive effort without rewarding tangible prizes. Hence, eliminating stage l does not affect V_1. Based on these arguments, we can search for the optimal contest sequence by considering only strictly decreasing sequences $\{N_l\}_{l=1}^{L}$, with $N_1 = N > N_2 > \cdots > N_L = 1$, without loss of generality.

We proceed with the following thought experiment. Suppose there exists an integer $J(< L)$ such that $N_J - N_{J+1} > 1$, i.e., more than one contestant is eliminated in stage J. It is therefore feasible to insert an additional stage between stage J and stage $J+1$. Let $M \in \{N_{J+1}+1, \ldots, N_J - 1\}$ contestants be selected from the N_J contestants in the Jth stage. Let the M survivors compete for the N_{J+1} "tickets" to move on to the next stage. Does adding this additional stage necessarily elicit more effort?

The effect of this additional stage on the level of effort exerted is mixed. Although the additional stage M creates an additional input of effort, it alters the contestants' incentive to supply effort in the previous stages. Firstly, in stage J, the total effort $N_J e_J$ could either decrease or increase. On one hand, the additional stage reduces the value of each "ticket" to the next stage ($V_M = (N_{J+1}V_{J+1} - Me_M)/M < V_{J+1}$), which tends to weaken contestants' incentives to exert their effort; On the other hand, the number of survivors in stage J increases from N_{J+1} to M. Its impact on $N_J e_J$ is indefinite, because $N_J e_J$ is not a monotonic function of the number of survivors.[23] Thus, the overall effect of the additional stage on $N_J e_J$ is ambiguous, as is its impact on V_J, a contestant' expected payoff in stage J. Secondly, as the direction of the change in V_J remains obscure, the effort supplied in all stages prior to stage J could either decrease or increase as well, because the effort in prior stages strictly increases with V_J. Consequently, the ramifications of the additional stage cannot be intuitively inferred. Nevertheless, the following theorem provides an unambiguous answer to the question of whether the positive effects always outweigh the negative effects in the context that is of interest in this paper.

Let $\mathcal{E}(\{N_l\})$ denote the set composed of all the elements in the sequence $\{N_l\}$.

Definition 1 Let $\{N_l\}$ and $\{\widetilde{N}_l\}$ be two contest sequences with $N_1 = \widetilde{N}_1 = N$, where N is the number of contestants available for the contest. The sequence $\{\widetilde{N}_l\}$ is more *complete* than $\{N_l\}$ if and only if $\mathcal{E}(\{N_l\}) \subset \mathcal{E}(\{\widetilde{N}_l\})$.

We show in the following theorem that any additional stage always increases the total effort, regardless of the existing contest structure.

Theorem 2 *When* $r \in (0, 1]$, *the more complete the contest sequence, the higher the total effort induced.*

Proof See Appendix. □

Theorem 2 is important as it establishes that an additional stage of competition always increases the total effort. A contest sequence is not optimal, as long as it leaves

[23] The following example shows that the component $N_{l+1} - \sum_{g=0}^{N_{l+1}-1} \frac{N_{l+1}-g}{N_l-g}$ is not a monotonic function of N_{l+1}. Assume $N_l = 5$. Then $N_{l+1} - \sum_{g=0}^{N_{l+1}-1} \frac{N_{l+1}-g}{N_l-g} = 1.35$ when $N_{l+1} = 2$; $N_{l+1} - \sum_{g=0}^{N_{l+1}-1} \frac{N_{l+1}-g}{N_l-g} = 1.57$ when $N_{l+1} = 3$; $N_{l+1} - \sum_{g=0}^{N_{l+1}-1} \frac{N_{l+1}-g}{N_l-g} = 1.28$ when $N_{l+1} = 4$.

room for an additional stage, i.e., the difference between any two successive terms in the sequence $\{N_l\}_{l=1}^{L}$ exceeds one. Thus, the optimal contest sequence is represented by a N-term strictly decreasing arithmetic sequence $\{N_l | N_l = N - l + 1, l = 1, 2, \ldots, N.\}$. The last term $N_N = 1$ represents the unique final winner. In other words, the contest asts for $N - 1$ stages, and one contestant is eliminated in each stage.[24] We call this a *full-sequence "Pyramid"* contest.

Theorem 3 *Suppose* $r \in (0, 1]$. *In a setting with N contestants and a prize budget* Γ_0, *the effort-maximizing sequential contest with pooling competition in each stage lasts for* $N - 1$ *stages. By eliminating one contestant in each stage, a single final winner takes over the entire prize purse of* Γ_0.

Theorem 3 naturally stems from Theorems 1 and 2. We conclude that the optimal contest must be organized as a "winner-take-all" full-sequence "Pyramid" contest.

Theorem 4 *The optimally designed N-person contest with the total prize purse of* Γ_0, *i.e., the "winner-take-all" "Pyramid" contest, elicits an equilibrium total effort*

$$E = \Gamma_0 \left\{ 1 - \left(\prod_{l=1}^{N} \frac{1}{N - l + 1} \right) \times \left(\prod_{l=1}^{N-1} \left[(N - l)(1 - r) + r \sum_{g=0}^{N-l-1} \frac{(N - l) - g}{(N - l + 1) - g} \right] \right) \right\}. \tag{11}$$

Theorem 4 explicitly derives the equilibrium total effort in the optimally designed N-person contest. The result directly arises from Lemma 1 and (11), as well as from the fact that the optimal contest structure is represented by a full-sequence of integers from N to 1, with a single winner taking over the entire prize purse.

4 Discussion and extensions

So far, we have shown that when the impact function $f(e) = e^r$ exhibits decreasing returns to effort ($r \leq 1$), the contest would elicit more effort if the contestants have to survive a longer line of shots before they win the final prize. To further understand the logic underlying our theoretical results and inspect the robustness of our main results, we extend our basic model to two alternative settings. Each of them illuminates one particular aspect of the theory on multi-stage contests. We further discuss possible future extensions to conclude this paper.

4.1 Extension 1: convex impact functions ($r > 1$)

We first extend our basic model to allow the contest technology to exhibit increasing returns, i.e., we allow the parameter r to exceed 1. Gradstein and Konrad (1999)

[24] In stage $N - 1$, two remaining contestants compete for one final prize.

establish $r = 1$ as the "watershed" for the optimality of contest designs: multi-stage contests would no longer dominate a static contest once the exponent reaches the threshold. We consider a selection procedure that is different from theirs,[25] and have shown that with "pooling competition" in each stage, a multi-stage contest would strictly dominate static contests even if $r = 1$. This raises a number of interesting questions: In our setting, (1) can multi-stage contests still dominate one-stage contests for strictly convex contest impact functions ($r > 1$)? and (2) to what extent would a multi-stage contest continue to prevail?

In the subsequent analysis, we investigate these issues for the case of $r \in (1, \frac{N}{N-1})$. Intuitively, a multi-stage contest may not increase effort if the impact function is excessively convex. When $r \geq \frac{N}{N-1}$, the optimal contest design is self-evident. It is well known in the literature that when $r \geq \frac{N}{N-1}$, a single-stage winner-take-all contest suffices to fully dissipate the rent (see Baye et al. 1994, 1999).[26] In this case, a static contest would be the optimal choice to maximize effort and no other mechanisms would be in demand. Is it possible, therefore, to identify a visible upper bound $\bar{r} \in (1, \frac{N}{N-1})$ such that as long as $r < \bar{r}$, a multi-stage contest would emerge as the optimum? In the subsequent analysis, we first establish a bound $\bar{r}_1 \in (1, \frac{N}{N-1})$ such that as long as $r \leq \bar{r}_1$, the "*hierarchical winner-take-all*" prize allocation principle remains valid. We argue that all multi-stage contests elicit more effort than single-stage contests for any r that is lower than that of the upper bound. We then further establish a smaller bound $\bar{r}_2 \in (1, \frac{N}{N-1})$ such that as long as $r \leq \bar{r}_2 (< \bar{r}_1)$, the result of Theorem 2 would hold and a full-sequence contest remains optimal.

4.1.1 "Hierarchical winner-take-all" principle

The proof of Proposition 3 indicates that the result does not lose its bite when $r \in (1, \frac{N}{N-1})$: the entire prize purse available to each stage must be concentrated on the first prize of that stage. Thus, V_1 can be written as $V_1 = \sum_{l=1}^{L} \theta_l W_l^1$ as in Sect. 3. In fact, it has indeed been implied that a "winner-take-all" (single-prize) structure can always emerge as the optimum. Because L is finite, a complete ranking among all θ_ls exists. To minimize V_1, the contest organizer only needs to allocate the entire prize purse to a prize W_l^1 with the smallest weight θ_l.

However, the "*hierarchical winner-take-all* " principle established in Theorem 1 requires the entire prize purse to be concentrated on the first prize in the top rank regardless of the contest sequence. It requires $\theta_L = \min_{l=1}^{L}\{\theta_l\}$, which implies that the final prize W_L^1 contributes the least surplus to contestants along the hierarchical ladder. However, in the case that there exists a minimum weight $\theta_{l'}$ such that $l' < L$, the

[25] Gradstein and Konrad (1999) assume contestants are divided into groups in each stage and that a single winner survives each group. In this paper, however, we allow remaining contestants to be pooled and to compete against all others.

[26] In this contest, when $r = \frac{N}{N-1}$, a symmetric pure-strategy equilibrium exists and induces individual effort $e = \frac{r(N-1)}{N^2}\Gamma_0 = \frac{\Gamma_0}{N}$, which fully dissipates the rent. When r exceeds $\frac{N}{N-1}$, a mixed-strategy equilibrium exists. In such an equilibrium, contestants can break-even on average and full rent-dissipation emerges.

prize money should be concentrated on the prize $W_{l'}^1$ at the optimum. As a result, the contest ends after stage l' and the rest of the sequence would be redundant. Obviously, this could happen only if the size of r is sufficiently large. Consider the following numerical example.

Example 1 Consider $N = 20$ and the contest sequence {20,19,18,17,16,2}. When $r = 1.0248$, the values of θ_ls are given by ($\theta_1 = 0.001318, \theta_2 = 0.001249, \theta_3 = 0.001178, \theta_4 = 0.001105, \theta_5 = 0.001032, \theta_6 = 0.000942$). $\theta_6 = 0.000942$ turns out to be the minimum. When r =1.04, the values of θ_ls are given by ($\theta_1 = 0.00060, \theta_2 = 0.00063, \theta_3 = 0.00065, \theta_4 = 0.000656, \theta_5 = 0.000652, \theta_6 = 0.000746$). $\theta_1 = 0.00060$ turns out to be the minimum.

This example, among many others, shows that θ_L can be the minimum when r is relatively small, but θ_1 will be the minimum when r approaches $\frac{N}{N-1}$. Thus, when r is excessively large, all subsequent prizes elicit less effort than W_1^1, i.e., the first prize awarded in the first stage. In this particular case, the "*hierarchical winner-take-all*" principle no longer holds and a single-stage contest dominates all other organizing rules for the given contest sequence. Nevertheless, we can establish the following upper bound $\bar{r}_1(> 1)$ such that the "*hierarchical winner-take-all*" principle holds if $r < \bar{r}_1$.

Proposition 4 *The "hierarchical winner-take-all" principle holds for any $r \in (1, \bar{r}_1)$, where $\bar{r}_1 = \frac{2N+1}{2N}$.*

Proof See Appendix. □

Proposition 4 states that it is optimal to concentrate the entire prize purse on the first prize in the top rank, regardless of the contest sequence, for any $r \in (1, \bar{r}_1)$.

As Rosen (1986) argues, moving the prize money toward the top of the hierarchy would effectively extend the horizon of the competition. The *"hierarchical winner-take-all"* principle established by Theorem 1 confirms that such an allocation rule elicits more effort. However, as indicated by Example 1, this principle may not hold when r is excessively large, i.e., when the contest technology is sufficiently discriminatory. Thus, if the contest organizer is able to control the course of the competition, a part of the given contest sequence would be redundant and the contest does not have to last that long. However, when the sequence of a contest is fixed such that the contest organizer is unable to flexibly adjust the course of the competition, spreading the prize money toward intermediate prizes could increase the output of the contest.

It is worth pointing out, however, that $\bar{r}_1$ is a rather conservative upper bound for the principle to hold. For Proposition 4 to hold, a sufficient (but not necessary) condition is to have the weights on W_l^1 decrease with l. This property would require a much less stringent condition:

$$\sum_{g=0}^{N_{l+1}-1} \frac{1}{N_l - g} \geq \frac{N_{l+1}}{N_l \left[N_{l+1} - (N_{l+1} - 1)r\right]}. \tag{12}$$

Unfortunately, the sign of $\sum_{g=0}^{N_{l+1}-1} \frac{1}{N_l-g} - \frac{N_{l+1}}{N_l[N_{l+1}-(N_{l+1}-1)r]}$ depends on subtle interaction among N_l, N_{l+1} and r. Identifying an upper bound of r for (13) to hold is difficult to do without losing efficiency, as the property of the sum $\sum_{g=0}^{N_{l+1}-1} \frac{1}{N_l-g}$ remains obscure. Few restrictions can be imposed effectively on N_l and N_{l+1}.

4.1.2 The optimality of multi-stage contests

In the case of $r \in (0, 1]$ that corresponds to a concave impact function, all the prize money should be allocated to the top prize regardless of the sequence of the contest. We have shown, however, that this principle may not hold and that the two main structural elements of a contest, i.e., contest sequence and the prize allocation rule, could interact subtly when r is sufficiently large. In Example 1, when $r = 1.04$, to maximize the total effort, all prize money should be allocated to the first prize awarded in the first stage. To increase the effort, the contest would be reduced to a static contest where 20 contestants exert one-shot effort to vie for the single prize. Thus, when r is sufficiently large, the legitimacy of a multi-stage contest could be called into question if the prize money can be flexibly distributed among all possible prizes. The contest sequence and the prize allocation rule represent the two sides of the same coin, which is exemplified by the following theorem.

Theorem 5 *All winner-take-all multi-stage contests elicit more effort than the single-stage winner-take-all contest for any $r \in (1, \bar{r}_1)$.*

Proof Consider a single-stage contest with N persons competing for a single prize Γ_0. From the literature, one would expect that each contestant in a single-stage winner-take-all contest expects to receive a payoff of

$$\overline{V} = \theta_1 \Gamma_0, \quad \text{where } \theta_1 = \frac{1}{N}\left((1-r) + \frac{r}{N}\right). \tag{13}$$

In any L-stage contest where $L \geq 2$ and N_l strictly decreases with l, a contestant's expected surplus V_1 is

$$V_1 = \sum_{l=1}^{L} \theta_l \Gamma_l, \quad L \geq 2. \tag{14}$$

As N_l strictly decreases with l, θ_l too strictly decreases with l in accordance with the proof of Proposition 4. Thus, $V_1 \leq \overline{V}$. For this reason, all hierarchical winner-take-all multi-stage contests must strictly dominate their single-stage counterparts in terms of effort induction. □

Theorem 5 states a sufficient condition for the optimality of multi-stage contests. It shows that as long as the "*hierarchical winner-take-all*" principle holds, any multi-stage contest would dominate a single-stage one, regardless of its contest sequence. Thus, all multi-stage contests could increase the output of the contest—even when the impact function is moderately convex ($r \in (1, \bar{r}_1)$). In fact, this result directly stems from Proposition 4. Note that the weight assigned to the first prize in the first stage W_l^1 is constant regardless of the sequence of the contest. It also represents the weight on the unique prize of a one-stage "winner-take-all" contest. When $r \in (1, \bar{r}_1)$, θ_l strictly decreases as the hierarchical ladder ascends. It naturally follows that a prize awarded in a later stage contributes less to a contestant's surplus, and a multi-stage contest that requires contestants to win the prize through more stages would induce more effort. This result is in contrast to the findings of Gradstein and Konrad (1999) and the difference lies mainly in the ways that elimination procedures are modeled.

It should be noted again that the bound $\bar{r}_1$ is rather conservative. The following example demonstrates that a single stage contest may not be optimal even when $r > \bar{r}_1$.

Example 2 Consider again the case of $N = 20$ and the contest sequence {20,19,18, 17,16,2}. In this case, $\bar{r}_1 = 1.025$. When $r = 1.0366$, the values of θ_ls are given by ($\theta_1 = 0.0007625, \theta_2 = 0.0007715, \theta_3 = 0.000769, \theta_4 = 0.000757, \theta_5 = 0.000737, \theta_6 = 0.00079$). $\theta_5 = 0.000737$ turns out to be the minimum.

Next, we further establish an upper bound for r such that the "Pyramid" contest can be optimal.

Theorem 6 *A full-sequence winner-take-all Pyramid contest maximizes the total effort for* $r \in (1, \bar{r}_2]$, *where* $\bar{r}_2 = 1 + \frac{1}{8N^2}$. *It lasts for* $N - 1$ *stages while eliminating one contestant at each stage, and a single winner takes over the entire prize purse of* Γ_0.

Proof See Appendix. □

Bound $\bar{r}_2$ is also very conservative. Consider a contest with $N = 10$, and $\bar{r}_2 = 801/800$. When $r = 1.088 > \bar{r}_2$, our numerical results show that a full-sequence winner-take-all Pyramid contest continues to be optimal.

Unfortunately, it is analytically difficult to characterize the exact form of the optimal contest sequence when $r > \bar{r}_2$. The analysis of the model heavily involves the harmonic series $\{\frac{1}{N_l - g}\}_{g=1}^{N_{l+1}-1}$, and it is mathematically difficult to derive a tractable form of the sum of such a series. Gradstein and Konrad (1999) in their setting establish "$r = 1$" as the "the watershed" level for the optimality of multi-stage contests. It remains an open question whether such a visible "watershed" $\bar{r} \in (1, \frac{N}{N-1})$ also exists in our setting, such that a full sequence maximizes the total effort when $r \in (0, \bar{r})$, while a single-stage one does otherwise. Nevertheless, ample numerical evidence exists to show that such a threshold $\bar{r}$ exists at least for $N \leq 15$. In these cases, $\bar{r}$ varies to the total number of contestants N, and it equalizes the first weight θ_1 and the last weight θ_N in a full sequence, i.e., $\theta_1(\bar{r}) = \theta_N(\bar{r})$.

Recall that when $r \geq \frac{N}{N-1}$, the optimal contest can take the form of a single-stage winner-take-all contest. Since $\frac{N}{N-1}$ converges monotonically to one as N approaches to infinity, a single-stage winner-take-all contest would be optimal for any $r > 1$ when N is sufficiently large.

4.2 Extension 2: stochastic elimination

Our results imply that the effort supplied in a contest could increase as the elimination process is slowed down by adding additional stages. Thus, a naturally compelling extension is to allow for stochastic elimination in each stage, as this would further extend the horizon of the game. In other words, instead of committing to knocking out one player in each round of competition, elimination occurs with a fixed probability, such that a round of competition may end up with no real outcome. We now examine the ramifications of this alternative rule.

Consider a full-sequence winner-take-all contest with $N(\geq 3)$ risk-neutral contestants. If $l(\geq 2)$ contestants remain in a particular stage, then one of them is to be eliminated and $l-1$ contestants advance to the next stage. However, given the stochastic elimination rule, the process of elimination could imply many rounds of competition. In each round, contestants submit their effort outlays simultaneously, and the event of elimination occurs with a probability $p_l \in (0, 1]$. The complementary event takes place with a probability of $1 - p_l$, where no contestant is eliminated. The same rule governs the competition in subsequent rounds until the "knock-out" is realized. We consider this stochastic elimination process (from l contestants to $l-1$) to be a *subcontest* of the whole game. The contestants have to repeat their attempts until the "knock-out" takes place, and the timing of this is unforeseen. The entire contest thus consists of $N-1$ such subcontests.

We use $V(l)$ to denote the conditional expected payoff of a representative contestant when there are l contestants who survived from the previous stage, $l \in \{N, N-1, \ldots, 1\}$.[27] Assume an impact function $f(e_i) = e_i^r$, with $r \in (0, \frac{N}{N-1})$. For any $l \geq 2$, a representative remaining contestant i solves the following problem

$$\begin{aligned}\text{Max}_{e_l^i} V_l^i &= p_l \sum_{m=1}^{l-1} P_m\left(e_l^i, \mathbf{e}_l^{-i}\right) V(l-1) + (1-p_l)V(l) - e_l^i \\ &= \sum_{m=1}^{l-1} P_m\left(e_l^i, \mathbf{e}_l^{-i}\right) (p_l V(l-1)) + (1-p_l)V(l) - e_l^i, \end{aligned} \tag{15}$$

where $P_m(\cdot, \cdot)$ is defined in Sect. 2.2.

The l-person contest would be repeated until the "knock-out" occurs. This elimination process mimics an infinitely repeated game with a positive continuation probability. To avoid complication, we focus on the case that contestants play the symmetric stage Nash equilibrium in every round of competition in this *subcontest* (from l contestants to $l-1$). Let e_k^* denote the individual equilibrium effort and V_k^* denote the conditional individual payoff in stage k of a full-sequence contest of Sect. 3. As defined in Sects. 3.2 and 4.1, $N_k = N - k + 1$ contestants remain in the game and compete to advance further up the hierarchy.

Proposition 5 *Consider a full-sequence contest composed of* $N-1$ *subcontests, with each of which eliminating one contestant. In a subcontest* $(l \to (l-1))$, $l(\geq 2)$ *persons*

[27] As the final winner gets Γ_0 without having to exert any effort, we must have $V(1) = \Gamma_0$.

compete for $l-1$ *tickets, and the elimination probability is* $p_l \in (0,1]$, $l = 2, \ldots, N$. *In the symmetric Nash equilibrium, each contestant exerts an effort*

$$\tilde{e}_l^{p_l} = p_l e^*_{N-l+1}, \tag{16}$$

in each round of the subcontest contest $(l \to (l-1))$. *The payoff* $V(l)$ *of a representative contestant is given by*

$$V(l) = V^*_{N-l+1}, \quad l = 2, \ldots, N. \tag{17}$$

Proof See Appendix. □

From Proposition 5, the entire process of the *subcontest* $(l \to (l-1))$ thus generates a total expected effort

$$E_{l\to(l-1)} = \sum_{k=0}^{\infty}(1-p_l)^k l\tilde{e}_l^{p_l} = l\tilde{e}_l^{p_l}\frac{1}{p_l} = le^*_{N-l+1}, \quad l = 2, \ldots, N. \tag{18}$$

Equation (18) shows that the total expected effort supplied in any *subcontest* does not depend on the probability of elimination $p_l \in (0,1]$. This further implies that on average, stochastic elimination does not affect the total equilibrium effort generated by the contest for any elimination probability $p_l > 0$. We therefore obtain the following.

Theorem 7 *Allowing for stochastic elimination does not change the total equilibrium effort generated on average by the full-sequence winner-take-all "Pyramid" contest.*

Two comments must be made. Firstly, the result will not hold if more sophisticated equilibria, such as those involving trigger strategies, are allowed for in the stochastic elimination game. However, the symmetric Nash equilibrium has already been represented as the most favorable outcome (in the viewpoint of the contest organizer) as other equilibria could only reduce effort outlay. We therefore conclude that allowing for stochastic elimination does not benefit the contest organizer.

Secondly, our analysis showed that the efficiency of a Pyramid contest cannot be improved through random elimination. The reasoning, however, continues to apply to any contest sequence. Without loss of generality, consider a winner-take-all contest with sequence $\{N_1, N_2, \ldots, N_L\}$ where $N_1 = N > N_2 > \cdots > N_L = 1$. Let V_l be the conditional equilibrium individual payoff in stage l of this contest. Imagine that in each round of the *subcontest* $(N_l \to N_{l+1})$, $N_l - N_{l+1}$ of the N_l contestants are simultaneously eliminated with probability $p_l \in (0,1]$. Similar to the previous full-sequence contest setting, we have

$$V(N_l) = V_l, \quad l = 1, 2, \ldots, L, \tag{19}$$

where $V(N_l)$ is the expected payoff of a remaining contestant in any round of the *subcontest* $(N_l \to N_{l+1})$. Hence, Theorem 7 and its implications remain valid.

Equation (19) reveals why stochastic elimination cannot increase the rent-dissipation rate over the extended horizon. V_l represents a contestant's conditional expected payoff when N_l remain in the contest with the sequence $\{N_1, N_2, \ldots, N_L\}$ without stochastic elimination. However, we see that a contestant's conditional expected payoff in each round of the elimination process $V(N_l)$ equals exactly V_l. An additional round of competition with a *subcontest* does not vary the payoff a contestant expects from the future, i.e., the value of the ticket $V(N_{l+1})$, while this repeatable process discounts future prizes, which, as evidenced by (15), weakens the contestants' incentives to exert effort in each round of competition. Consequently, this repeatable process does not provide additional incentives to the contestants to supply effort.

4.3 Further extensions

In this paper, we allow the contest organizer to design the optimal contest using two instruments: the contest sequence and the allocation of a fixed total prize purse. When the impact function takes the form $f(e) = e^r$, and the size of r remains in a moderate range, we show (1) that the contest organizer must allocate the entire prize purse to a single final prize, regardless of the contest sequence; and (2) that a multi-stage contest elicits more effort than a single-stage contest. When a sufficiently "noisy" contest technology (r is sufficiently small) is in place, a "Pyramid" contest that eliminates one contestant in each stage would emerge as the optimum. Our results therefore provide a rationale for the multi-phase sequential competition that is widely observed in reality.

Our analysis provides important insights into the optimal design of multi-stage contests, but leaves open tremendous possibilities for future extension. First, we do not consider the cost of organizing the contest. The contest organizer may be concerned about the additional costs that could arise from additional stages, and this concern should indeed be taken into account in future research into the optimal design of multi-stage contests. Secondly, we do not consider the heterogeneity in the abilities and preferences of the contestants. One interesting but technically challenging extension is to allow for contestants of differing private types. While we do not believe that an extension in this direction will vary the main theme of our results, it would be intriguing to investigate the efficiency of a multi-stage contest in serving as a screening mechanism when contestants have private information about their abilities.[28] Finally, our model assumes that the contestants' effort affects only the outcome of the subcontest in the current stage. One may extend this model by allowing for "accumulatable" effort, in which case effort made in the current stage can be carried over into future stages such that it continues to influence a contestant's likelihood of winning.

Appendix

Proof of Proposition 1 The subgame perfect Nash equilibrium can be solved through backward induction. In stage l, a representative remaining contestant i rationally chooses his effort e_l^i to maximize his expected payoff (2). Since we consider the

[28] Rosen (1986) considers the effect of heterogeneous types. Only numerical cases are discussed due to tractability.

symmetric equilibrium, we assume all contestants other than i exert the same effort e_l' without loss of generality. Under this simplification,

$$P_m\left(e_l^i, e_l^{-i}\right) = \frac{(N-1)!}{(N-m)!}\left[\prod_{k=1}^{m-1}\frac{\left(e_l'\right)^r}{\left(e_l^i\right)^r + (N-k)\left(e_l'\right)^r}\right] \times \frac{\left(e_l^i\right)^r}{\left(e_l^i\right)^r + (N-m)\left(e_l'\right)^r}. \tag{A.1}$$

From (A.1), $\frac{\partial P_m(e_l^i, \mathbf{e}_l^{-i})}{\partial e_l^i}$ is given by

$$\begin{aligned}
&\frac{\partial P_m\left(e_l^i, \mathbf{e}_l^{-i}\right)}{\partial e_l^i} \\
&= \frac{(N-1)!}{(N-m)!}\left[\prod_{k=1}^{m-1}\frac{(e_l')^r}{(e_l^i)^r + (N-k)(e_l')^r}\right] \times \frac{r(e_l^i)^{r-1}(N-m)(e_l')^r}{[(e_l^i)^r + (N-m)(e_l')^r]^2} \\
&\quad - \left\{\frac{(N-1)!}{(N-m)!} \times \left[\prod_{k=1}^{m-1}\frac{(e_l')^r}{(e_l^i)^r + (N-k)(e_l')^r}\right] \times \frac{(e_l^i)^r}{(e_l^i)^r + (N-m)(e_l')^r}\right. \\
&\quad \left. \times \sum_{k=1}^{m-1}\frac{r(e_l^i)^{r-1}}{(e_l^i)^r + (N-k)(e_l')^r}\right\}.
\end{aligned} \tag{A.2}$$

In a symmetric equilibrium with $e_l^i = e_l$, we obtain

$$\frac{\partial P_m\left(e_l, \ldots, e_l\right)}{\partial e_l^i} = \frac{\left(1 - \sum_{g=0}^{m-1}\frac{1}{N_l - g}\right) r}{N_l e_l}. \tag{A.3}$$

From (2), the first order condition for an interior equilibrium is

$$\sum_{m=1}^{N_{l+1}}\left[\frac{\partial P_m(e_l, \ldots, e_l)}{\partial e_l^i}(V_{l+1} + W_l^m)\right] + \sum_{m=N_{l+1}+1}^{N_l}\left[\frac{\partial P_m(e_l, \ldots, e_l)}{\partial e_l^i} W_l^m\right] - 1 = 0. \tag{A.4}$$

(A.3) and (A.4) give that if $e_l > 0$, then it must satisfy $\frac{r\Phi_l}{N_l e_l} - 1 = 0$, where $\Phi_l = \sum_{m=1}^{N_{l+1}}[(1 - \sum_{g=0}^{m-1}\frac{1}{N_l - g})(V_{l+1} + W_l^m)] + \sum_{m=N_{l+1}+1}^{N_l}[(1 - \sum_{g=0}^{m-1}\frac{1}{N_l - g})W_l^m]$. Following this result, we have that $e_l = \frac{r\Phi_l}{N_l}$ if $\Phi_l > 0$, whereas $e_l = 0$ if $\Phi_l \leq 0$. These results can be alternatively written as in (4). Each symmetric contestant has the same chance of winning each component of the total stage-award $N_{l+1}V_{l+1} + \Gamma_l$ (including N_l stage prizes and N_{l+1} tickets to the next stage) in a symmetric equilibrium. Therefore, the equilibrium expected payoff of a representative contestant at stage l is given by $V_l = (N_{l+1}V_{l+1} + \Gamma_l)/N_l - e_l$. □

Proof of Theorem 1 The proof proceeds in two steps.

Step 1: The optimal prize allocation requires the entire prize purse $\Gamma \leq \Gamma_0$ to be allocated to the first prize in the last stage, i.e., $W_l^1 = 0$, for $l \in \{1, 2, \ldots, L-1\}$, and $W_L = \Gamma$. From (10), we have at the optimal prize allocation $V_1 = \sum_{l=1}^{L} \theta_l W_l^1$. Next, we show that for any $l \in \{1, 2, \ldots, L-1\}$, shifting resources from W_l^1 toward W_{l+1}^1 further reduces V_1. This property holds if and only if $\{\theta_l\}$ is a decreasing sequence. We therefore compare θ_l with θ_{l+1} for $l \in \{1, 2, \ldots, L-1\}$. Ignoring the common element $(\prod_{j=1}^{l} \frac{1}{N_j})(\prod_{j=1}^{l-1}[N_{j+1}(1-r)+r\sum_{g=0}^{N_{j+1}-1} \frac{N_{j+1}-g}{N_j-g}])$[29] contained in both of the two terms, we only need to compare $[(1-r)+\frac{r}{N_l}]$ to $\frac{1}{N_{l+1}}[N_{l+1}(1-r)+r\sum_{g=0}^{N_{l+1}-1} \frac{N_{l+1}-g}{N_l-g}][(1-r)+\frac{r}{N_{l+1}}]$. We have

$$\begin{aligned}\theta_l - \theta_{l+1} \propto & \left[(1-r)+\frac{r}{N_l}\right] \\ & - \frac{1}{N_{l+1}}\left[N_{l+1}(1-r)+r\sum_{g=0}^{N_{l+1}-1}\frac{N_{l+1}-g}{N_l-g}\right]\left[(1-r)+\frac{r}{N_{l+1}}\right] \\ = & \left[(1-r)+\frac{r}{N_l}\right]-\left[(1-r)+\frac{r\sum_{g=0}^{N_{l+1}-1}\frac{N_{l+1}-g}{N_l-g}}{N_{l+1}}\right]\left[(1-r)+\frac{r}{N_{l+1}}\right].\end{aligned} \tag{A.5}$$

Note that $\frac{N_{l+1}-g}{N_l-g} \leq \frac{N_{l+1}}{N_l}$ for $g \geq 0$. Thus we have

$$\begin{aligned}\text{RHS of (A.5)} \geq & \left[(1-r)+\frac{r}{N_l}\right]-\left[(1-r)+\frac{rN_{l+1}}{N_l}\right]\left[(1-r)+\frac{r}{N_{l+1}}\right] \\ = & \, r(1-r)\left(N_{l+1}-1\right)\left(\frac{1}{N_{l+1}}-\frac{1}{N_l}\right) \geq 0.\end{aligned}$$

Thus, we show that the weight on W_l^1 decreases along the path. It follows that in order to minimize V_1, the budget allocated to prizes in an earlier stage $(l < L)$ should be shifted toward W_L^1. Thus we conclude that the optimal prize allocation requires $W_L^1 = \Gamma$.

Step 2: $\Gamma = \Gamma_0$, i.e., the contest organizer uses up the entire budget on prizes. The total effort E can be reduced to the following form given the optimal prize allocation rule we obtained from step one.

[29] This term is positive for any $r \in (0, \frac{N}{N-1})$.

$$E = \Gamma \left\{ 1 - N \left(\prod_{j=1}^{L} \frac{1}{N_j} \right) \left(\prod_{j=1}^{L-1} \left[N_{j+1}(1-r) + r \sum_{g=0}^{N_{j+1}-1} \frac{N_{j+1}-g}{N_j - g} \right] \right) \right.$$
$$\left. \times \left[(1-r) + \frac{r}{N_L} \right] \right\}.$$

Note that

$$N \left(\prod_{j=1}^{L} \frac{1}{N_j} \right) \left(\prod_{j=1}^{L-1} \left[N_{j+1}(1-r) + r \sum_{g=0}^{N_{j+1}-1} \frac{N_{j+1}-g}{N_j-g} \right] \right) \left[(1-r) + \frac{r}{N_L} \right]$$
$$= \frac{N}{N_L} \left[(1-r) + \frac{r}{N_L} \right] \prod_{j=1}^{L-1} \left\{ \frac{1}{N_j} \left[N_{j+1}(1-r) + r \sum_{g=0}^{N_{j+1}-1} \frac{N_{j+1}-g}{N_j-g} \right] \right\}$$
$$< \frac{N}{N_L} \left[(1-r) + \frac{r}{N_L} \right] \prod_{j=1}^{L-1} \frac{1}{N_j} \left[N_{j+1}(1-r) + rN_{j+1} \right]$$
$$= (1-r) + \frac{r}{N_L} \leq 1,$$

we have that E increases strictly with Γ. As a consequence, the entire budget Γ_0 should be allocated to W_L^1 in order to maximize the total effort. □

Proof of Theorem 2 Denote by E_0 the total efforts in the original contest $\{N_l\}$, while by E_M the total effort in the hypothetical contest after one additional stage M is inserted. We only need to show $E_M > E_0$.

Denote by $\widetilde{V}_1$ the equilibrium expected payoff that the N contestants anticipate at the first stage of the contest after the additional stage is inserted. By Lemma 1, we only need to show $\widetilde{V}_1 < V_1$.

Under the optimal prize allocation rule we have characterized in Theorem 1, (4) and (5) lead to

$$V_1 = \frac{V_{J+1}}{\prod_{l=1}^{J} N_l} \times \prod_{l=1}^{J} \left[N_{l+1}(1-r) + r \sum_{g=0}^{N_{l+1}-1} \frac{N_{l+1}-g}{N_l - g} \right].$$

Similarly, we obtain the expected payoff $\widetilde{V}_1$ after stage M is hypothetically inserted, which is given by

$$\widetilde{V}_1 = \frac{V_{J+1}}{M\prod_{l=1}^{J} N_l} \times \left\{ \prod_{l=1}^{J-1} \left[N_{l+1}(1-r) + r \sum_{g=0}^{N_{l+1}-1} \frac{N_{l+1}-g}{N_l - g} \right] \right\}$$
$$\times \left[M(1-r) + r \sum_{g=0}^{M-1} \frac{M-g}{N_J - g} \right] \times \left[N_{J+1}(1-r) + r \sum_{g=0}^{N_{J+1}-1} \frac{N_{J+1}-g}{M-g} \right].$$

To establish $\widetilde{V}_1 < V_1$, we need to show the following sufficient and necessary condition is satisfied:

$$\frac{1}{M}\left[M(1-r) + r \sum_{g=0}^{M-1} \frac{M-g}{N_J - g} \right] \times \left[N_{J+1}(1-r) + r \sum_{g=0}^{N_{J+1}-1} \frac{N_{J+1}-g}{M-g} \right]$$
$$< N_{J+1}(1-r) + r \sum_{g=0}^{N_{J+1}-1} \frac{N_{J+1}-g}{N_J - g}. \tag{A.6}$$

Rearrange LHS of (A.6), and we obtain

$$\frac{1}{M}\left[M(1-r) + r \sum_{g=0}^{M-1} \frac{M-g}{N_J - g} \right] \times \left[N_{J+1}(1-r) + r \sum_{g=0}^{N_{J+1}-1} \frac{N_{J+1}-g}{M-g} \right]$$
$$= N_{J+1}(1-r)^2 + N_{J+1} r(1-r) \times \frac{\sum_{g=0}^{M-1} \frac{M-g}{N_J-g}}{M} + r(1-r) \times \sum_{g=0}^{N_{J+1}-1} \frac{N_{J+1}-g}{M-g}$$
$$+ r^2 \frac{\sum_{g=0}^{M-1} \frac{M-g}{N_J-g}}{M} \sum_{g=0}^{N_{J+1}-1} \frac{N_{J+1}-g}{M-g}.$$

Because $\frac{M-g}{N_J-g}$ is decreasing in g and $N_{J+1} < M$, $\frac{\sum_{g=0}^{M-1} \frac{M-g}{N_J-g}}{M} < \frac{\sum_{g=0}^{N_{J+1}-1} \frac{M-g}{N_J-g}}{N_{J+1}}$ must hold, which leads to

$$\text{LHS of (A.6)} < N_{J+1}(1-r)^2 + r(1-r) \sum_{g=0}^{N_{J+1}-1} \frac{M-g}{N_J - g} + r(1-r)$$
$$\times \sum_{g=0}^{N_{J+1}-1} \frac{N_{J+1}-g}{M-g} + r^2 \frac{\sum_{g=0}^{N_{J+1}-1} \frac{M-g}{N_J-g}}{N_{J+1}} \sum_{g=0}^{N_{J+1}-1} \frac{N_{J+1}-g}{M-g}.$$

By Chebyshev Sum Inequality, $\frac{1}{N_{J+1}}(\sum_{g=0}^{N_{J+1}-1}\frac{M-g}{N_J-g})(\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{M-g}) \leq \sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{N_J-g}$. Thus, we obtain

$$\begin{aligned}
&\text{LHS of (A.6)}\\
&< N_{J+1}(1-r)\left[(1-r) + r\frac{\sum_{g=0}^{N_{J+1}-1}\frac{M-g}{N_J-g}}{N_{J+1}} + r\frac{\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{M-g}}{N_{J+1}}\right]\\
&\quad + \left[r\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{N_J-g} - r(1-r)\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{N_J-g}\right]\\
&\leq N_{J+1}(1-r)\left[(1-r) + \frac{r}{N_{J+1}}\sum_{g=0}^{N_{J+1}-1}\left(\frac{M-g}{N_J-g} + \frac{N_{J+1}-g}{M-g} - \frac{N_{J+1}-g}{N_J-g}\right)\right]\\
&\quad + r\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{N_J-g}.
\end{aligned}$$

Simple algebra would verify $\frac{M-g}{N_J-g} + \frac{N_{J+1}-g}{M-g} - \frac{N_{J+1}-g}{N_J-g} < 1$. We therefore obtain

$$\begin{aligned}
\text{LHS of (A.6)} &< N_{J+1}(1-r)\left[(1-r) + \frac{rN_{J+1}}{N_{J+1}}\right] + r\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{N_J-g}\\
&= \text{RHS of (A.6)}.
\end{aligned}$$

□

Proof of Proposition 4 It is sufficient to show weights θ_l decrease with l when $r \in (1, \bar{r}_1)$. For this purpose, we only need to compare the weight θ_l on W_l^1 to the weight θ_{l+1} on W_{l+1}^1, where $\theta_l = (\prod_{j=1}^{l}\frac{1}{N_j})(\prod_{j=1}^{l-1}[N_{j+1}(1-r) + r\sum_{g=0}^{N_{j+1}-1}\frac{N_{j+1}-g}{N_j-g}])[(1-r) + \frac{r}{N_l}]$. Recall (A.5):$\theta_l - \theta_{l+1} \propto [(1-r) + \frac{r}{N_l}] - [(1-r) + \frac{r\sum_{g=0}^{N_{l+1}-1}\frac{N_{l+1}-g}{N_l-g}}{N_{l+1}}][(1-r) + \frac{r}{N_{l+1}}]$. Rearrange the RHS, and we obtain

$$\begin{aligned}
\text{RHS of (A.5)} &= \left(1 - \frac{N_l - 1}{N_l}r\right) - \left(1 - \frac{N_{l+1} - \sum_{g=0}^{N_{l+1}-1}\frac{N_{l+1}-g}{N_l-g}}{N_{l+1}}r\right)\\
&\quad\times\left(1 - \frac{N_{l+1}-1}{N_{l+1}}r\right).
\end{aligned}$$

Note $N_{l+1} - \sum_{g=0}^{N_{l+1}-1}\frac{N_{l+1}-g}{N_l-g} = (N_l - N_{l+1})\sum_{g=0}^{N_{l+1}-1}\frac{1}{N_l-g}$. We thus have

RHS of (A.5)

$$= \frac{(N_l - N_{l+1}) \sum_{g=0}^{N_{l+1}-1} \frac{1}{N_l - g}}{N_{l+1}} r - \frac{N_l - N_{l+1}}{N_l N_{l+1}} r - \frac{(N_{l+1}-1)(N_l - N_{l+1}) \sum_{g=0}^{N_{l+1}-1} \frac{1}{N_l - g}}{N_{l+1}^2} r^2$$

$$= \frac{(N_l - N_{l+1})r}{N_{l+1}^2} \left\{ \left[N_{l+1} - (N_{l+1}-1)r\right] \sum_{g=0}^{N_{l+1}-1} \frac{1}{N_l - g} - \frac{N_{l+1}}{N_l} \right\}.$$

Because $\frac{(N_l - N_{l+1})r}{N_{l+1}^2} > 0$, the sign of the last expression is the same as that of $[N_{l+1} - (N_{l+1}-1)r] \sum_{g=0}^{N_{l+1}-1} \frac{1}{N_l - g} - \frac{N_{l+1}}{N_l}$. In other words, $\theta_l - \theta_{l+1} \geq 0$ if and only if

$$\sum_{g=0}^{N_{l+1}-1} \frac{1}{N_l - g} \geq \frac{N_{l+1}}{N_l \left[N_{l+1} - (N_{l+1}-1)r\right]}. \tag{A.7}$$

Note that $\frac{1}{N_l - g}$ is increasing and convex in g. By Jensen's inequality, $\sum_{g=0}^{N_{l+1}-1} \frac{1}{N_l - g} \geq \frac{N_{l+1}}{N_l - \frac{N_{l+1}-1}{2}}$. Thus, (A.7) must hold if $\frac{N_{l+1}}{N_l - \frac{N_{l+1}-1}{2}} \geq \frac{N_{l+1}}{N_l[N_{l+1} - (N_{l+1}-1)r]}$, which is equivalent to

$$2N_l(N_{l+1}-1) - 2N_l(N_{l+1}-1)r \geq -(N_{l+1}-1).$$

When $r \leq \frac{2N+1}{2N} \leq \frac{2N_l+1}{2N_l}$, the last expression must hold.

□

Proof of Theorem 6 By Proposition 3, we write V_1 as $V_1 = \sum_{l=1}^{L} \theta_l W_l^1$. Because L is finite, a complete ranking among all θ_ls exists. To minimize V_1, the contest organizer only needs to allocate the entire prize purse to a prize W_l^1 with the smallest weight θ_l. We follow the logic in 4.1.1. An optimal contest must award the entire prize purse to the first winner in its last stage. It is either a multi-stage contest under the "hierarchical winner-take-all" rule or a single stage one. Thus, when $r \in (1, \frac{N}{N-1})$, there is no loss of generality to assume "hierarchical winner-take-all" in searching for optimal contest sequence.

Clearly, for any contest sequence with "hierarchical winner-take-all", if $N_L > 1$, we can always add another (hypothetical) stage with $N_{L+1} = 1$. The new hypothetical sequence will induce the same amount of effort as the original one does. Therefore, there is no loss of generality to consider contest sequence with $N_L = 1$ and winner-take-all for the optimal contest.

To show Theorem 6, we only need to show that when $r \in (1, \bar{r}_2]$, any contest sequence is dominated by the full sequence.

We first show that when $r \in (1, \bar{r}_2]$, if $N_{L-1} > 2$ then inserting a stage with $M = 2$ contestants increases effort induced. For this purpose, from (A.6) we need to show

$$\left[(1-r)+r\frac{1}{M}\sum_{g=0}^{M-1}\frac{M-g}{N_J-g}\right]\left[(1-r)+r\frac{1}{N_L}\sum_{g=0}^{N_L-1}\frac{N_L-g}{M-g}\right]$$
$$\leq\left[(1-r)+r\frac{1}{N_L}\sum_{g=0}^{N_L-1}\frac{N_L-g}{N_{L-1}-g}\right], \tag{A.8}$$

where $N_L=1$, $M=2$. (A.8) holds if

$$r\leq\frac{1-\frac{1}{N_L}\sum_{g=0}^{N_L-1}\frac{N_L-g}{M-g}-\frac{1}{M}\sum_{g=0}^{M-1}\frac{M-g}{N_{L-1}-g}+\frac{1}{N_L}\sum_{g=0}^{N_L-1}\frac{N_L-g}{N_{L-1}-g}}{1-\frac{1}{N_L}\sum_{g=0}^{N_L-1}\frac{N_L-g}{M-g}-\frac{1}{M}\sum_{g=0}^{M-1}\frac{M-g}{N_{L-1}-g}+\frac{1}{N_L}\frac{1}{M}\left(\sum_{g=0}^{M-1}\frac{M-g}{N_{L-1}-g}\right)\left(\sum_{g=0}^{N_L-1}\frac{N_L-g}{M-g}\right)}$$
$$=1+\frac{\frac{1}{N_{L-1}}-\frac{1}{4}\left(\sum_{g=0}^{1}\frac{2-g}{N_{L-1}-g}\right)}{\frac{1}{2}-\frac{1}{4}\sum_{g=0}^{1}\frac{2-g}{N_{L-1}-g}}.$$

As

$$1+\frac{\frac{1}{N_{L-1}}-\frac{1}{4}\left(\sum_{g=0}^{1}\frac{2-g}{N_{L-1}-g}\right)}{\frac{1}{2}-\frac{1}{4}\sum_{g=0}^{1}\frac{2-g}{N_{L-1}-g}}\geq 1+\frac{2}{N_{L-1}}-\frac{1}{2}\left(\sum_{g=0}^{1}\frac{2-g}{N_{L-1}-g}\right)$$
$$\geq 1+\frac{1}{4N}>\bar{r}_2=1+\frac{1}{8N^2},$$

(A.8) must hold when $r\in(1,\bar{r}_2]$.

We now consider a sequence with $N_{L-1}=2$. In this case, we show that if it is feasible to insert an additional stage with $M\in\{N_{J+1}+1,\ldots,N_J-1\}$ contestants between stage J and stage $J+1$, then this additional stage increases effort induced. For this purpose, we only need to show that (A.6) holds for $r\in(1,\bar{r}_2]$. Since $N_{L-1}=2$ and $N_L=1$, we must have $N_{J+1}\geq 2$ (A.6) holds if and only if

$$r\leq\frac{1-\frac{1}{N_{J+1}}\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{M-g}-\frac{1}{M}\sum_{g=0}^{M-1}\frac{M-g}{N_J-g}+\frac{1}{N_{J+1}}\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{N_J-g}}{1-\frac{1}{N_{J+1}}\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{M-g}-\frac{1}{M}\sum_{g=0}^{M-1}\frac{M-g}{N_J-g}+\frac{1}{N_{J+1}}\frac{1}{M}\left(\sum_{g=0}^{M-1}\frac{M-g}{N_J-g}\right)\left(\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{M-g}\right)}$$
$$=1+\frac{\frac{1}{N_{J+1}}\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{N_J-g}-\frac{1}{N_{J+1}}\frac{1}{M}\left(\sum_{g=0}^{M-1}\frac{M-g}{N_J-g}\right)\left(\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{M-g}\right)}{\left(\frac{N_J-M}{M}\sum_{g=0}^{M-1}\frac{1}{N_J-g}\right)\left(\frac{M-N_{J+1}}{N_{J+1}}\sum_{g=0}^{N_{J+1}-1}\frac{1}{M-g}\right)}. \tag{A.9}$$

Note that if $a_1>a_2>a_3>\cdots>a_n$, and $b_1>b_2>b_3>\cdots>b_n$, then we have[30]

$$n\sum_{g=1}^{n}a_ib_i-\left(\sum_{g=1}^{n}a_i\right)\left(\sum_{g=1}^{n}b_i\right)\geq(a_{n-1}-a_n)\left[\sum_{g=1}^{n}(b_i-b_n)\right]. \tag{A.10}$$

[30] The proof is available from the authors upon request.

From (A.10), we have

$$1+\frac{\frac{1}{N_{J+1}}\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{N_J-g}-\frac{1}{N_{J+1}}\frac{1}{M}\left(\sum_{g=0}^{M-1}\frac{M-g}{N_J-g}\right)\left(\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{M-g}\right)}{\left(\frac{N_J-M}{M}\sum_{g=0}^{M-1}\frac{1}{N_J-g}\right)\left(\frac{M-N_{J+1}}{N_{J+1}}\sum_{g=0}^{N_{J+1}-1}\frac{1}{M-g}\right)}$$

$$\geq 1+\frac{\frac{1}{N_{J+1}}\sum\limits_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{N_J-g}-\frac{1}{N_{J+1}}\frac{1}{M}\left(\sum_{g=0}^{M-1}\frac{M-g}{N_J-g}\right)\left(\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{M-g}\right)}{\frac{N_J-M}{N_J-\frac{M-1}{2}}\frac{M-N_{J+1}}{M-\frac{N_{J+1}-1}{2}}}$$

$$\geq 1+\frac{1}{N_{J+1}}\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{N_J-g}-\frac{1}{N_{J+1}}\frac{1}{N_{J+1}}\left(\sum_{g=0}^{N_{J+1}-1}\frac{M-g}{N_J-g}\right)$$
$$\times\left(\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{M-g}\right),$$

which can be further rewritten as

$$1+\left(\frac{1}{N_{J+1}}\right)^2\left[N_{J+1}\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{N_J-g}\right.$$
$$\left.-\left(\sum_{g=0}^{N_{J+1}-1}\frac{M-g}{N_J-g}\right)\left(\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{M-g}\right)\right]$$
$$\geq 1+\left(\frac{1}{N_{J+1}}\right)^2\left[\left(\frac{2}{N_J-M+2}-\frac{1}{N_J-M+1}\right)\right.$$
$$\left.\times\left(\sum_{g=0}^{N_{J+1}-1}\left(\frac{N_{J+1}-g}{M-g}-\frac{1}{M-N_{J+1}+1}\right)\right)\right]$$
$$\geq 1+\left(\frac{1}{N_{J+1}}\right)\left[\left(\frac{2}{N_J-M+2}-\frac{1}{N_J-M+1}\right)\right]$$
$$\times\left[(N_{J+1}-1)\left(\left(\frac{2}{N_J-M+2}-\frac{1}{N_J-M+1}\right)\right)\right]$$
$$=1+\frac{N_{J+1}-1}{N_{J+1}}\left[\frac{1}{N_J-M+2}\left(1-\frac{1}{(N_J-M+1)}\right)\right]^2.$$

Note that $1+\frac{N_{J+1}-1}{N_{J+1}}[\frac{1}{N_J-M+2}(1-\frac{1}{(N_J-M+1)})]^2\geq 1+\frac{1}{2}(\frac{1}{N}\frac{1}{2})^2 = 1+\frac{1}{8N^2}=\bar{r}_2$, because $N_{J+1}\geq 2$. We then obtain

$$1+\frac{\frac{1}{N_{J+1}}\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{N_J-g}-\frac{1}{N_{J+1}}\frac{1}{M}(\sum_{g=0}^{M-1}\frac{M-g}{N_J-g})(\sum_{g=0}^{N_{J+1}-1}\frac{N_{J+1}-g}{M-g})}{(\frac{N_J-M}{M}\sum_{g=0}^{M-1}\frac{1}{N_J-g})(\frac{M-N_{J+1}}{N_{J+1}}\sum_{g=0}^{N_{J+1}-1}\frac{1}{M-g})}$$
$$\geq 1+\frac{1}{8N^2}=\bar{r}_2.$$

Therefore, (A.9) must hold when $r\in(1,\bar{r}_2]$. Thus, when $N_{L-1}=2$, inserting an additional stage always increases effort induced. To summarize, we have shown that the full sequence dominates any other sequence when $r\in(1,\bar{r}_2]$. □

Proof of Proposition 5 Applying Proposition 1 to program (15) leads to that the equilibrium effort is

$$\tilde{e}_l^{p_l}=\frac{r}{l}\sum_{m=1}^{l-1}\left[\left(1-\sum_{g=0}^{m-1}\frac{1}{l-g}\right)p_lV(l-1)\right],\quad l=2,\ldots,N. \tag{A.11}$$

From (A.11) and (15), we have

$$V(l)=p_l\left(\frac{l-1}{l}V(l-1)\right)+(1-p_l)V(l)-\tilde{e}_l^{p_l},\quad l=2,\ldots,N.$$

We thus have

$$V(l)=\frac{l-1}{l}V(l-1)-\frac{r}{l}\sum_{m=1}^{l-1}\left[\left(1-\sum_{g=0}^{m-1}\frac{1}{l-g}\right)V(l-1)\right].$$

As $l=N_{N-l+1}$, we have

$$V(l)=\frac{N_{N-l+1}-1}{N_{N-l+1}}V(N_{N-l+1}-1)-\frac{r}{N_{N-l+1}}$$
$$\times\sum_{m=1}^{N_{N-l+1}-1}\left[\left(1-\sum_{g=0}^{m-1}\frac{1}{N_{N-l+1}-g}\right)V(N_{N-l+1}-1)\right].$$

As $V(1)=V_N^*=\Gamma_0$, we have

$$V(2)=\frac{N_{N-1}-1}{N_{N-1}}\Gamma_0-\frac{r}{N_{N-1}}\sum_{m=1}^{N_{N-1}-1}\left[\left(1-\sum_{g=0}^{m-1}\frac{1}{N_{N-1}-g}\right)\Gamma_0\right]$$
$$=\frac{N_{N-1}-1}{N_{N-1}}\Gamma_0-e_{N-1}^*$$
$$=V_{N-1}^*.$$

Similarly, we have

$$\begin{aligned} V(l) &= \frac{N_{N-l+1}-1}{N_{N-l+1}} V\left(N_{N-l+1}-1\right) - \frac{r}{N_{N-l+1}} \\ &\quad \times \sum_{m=1}^{N_{N-l+1}-1} \left[\left(1 - \sum_{g=0}^{m-1} \frac{1}{N_{N-l+1}-g} \right) V(N_{N-l+1}-1) \right] \\ &= \frac{N_{N-l+1}-1}{N_{N-l+1}} V(N_{N-l+1}-1) - e^*_{N-l+1} \\ &= V^*_{N-l+1}, \quad l = 3, \ldots, N. \end{aligned}$$

Since $V(l) = V^*_{N-l+1}$, (A.11) leads to that

$$\tilde{e}_l^{p_l} = p_l e^*_{N-l+1}, \quad l = 2, \ldots, N.$$

□

References

Amegashie, J.A.: The design of rent-seeking competitions: committees preliminary and final contest. Public Choice **99**, 63–76 (1999)

Amegashie, J.A.: Some results on rent-seeking contests with shortlisting. Public Choice **105**, 245–253 (2000)

Barut, Y., Kovenock, D.: The symmetric multiple prize all-pay auction with complete information. Eur J Political Econ **14**, 627–644 (1998)

Baye, M.R., Kovenock, D., de Vries, C.G.: Rigging the lobbying process: an application of all-pay auction. Am Econ Rev **86**, 289–293 (1993)

Baye, M.R., Kovenock, D., de Vries, C.G.: The solution to the Tullock rent-seeking game when $R > 2$: mixed-strategy equilibria and mean dissipation rates. Public Choice **81**, 362–380 (1994)

Baye, M.R., Kovenock, D., de Vries, C.G.: The incidence of overdissipation in rent-seeking contests. Public Choice **99**(3–4), 439–454 (1999)

Clark, D.J., Riis, C.: A multiple-winner nested rent-seeking contest. Public Choice **87**, 177–184 (1996)

Clark, D.J., Riis, C.: Influence and the discretionary allocation of several prizes. Eur J Political Econ **14**, 605–625 (1998)

Clark, D.J., Riis, C.: Competition over more than one prize. Am Econ Rev **88**, 276–289 (1998)

Fu, Q., Lu, J.: The beauty of “Bigness”: on optimal design of multi-winner contests. Games Econ Behav (2008a, forthcoming)

Fu, Q., Lu, J.: A Micro Foundation of Generalized Multi-Prize Contests: A Noisy Ranking Perspective, working paper (2008b)

Fullerton, R.L., McAfee, P.R.: Auctioning entry into tournaments. J Political Econ **107**, 573–605 (1999)

Gradstein, M., Konrad, K.A.: Orchestrating rent seeking contests. Econ J **109**, 535–545 (1999)

Harbaugh, R., Klumpp, T.: Early round upsets and championship blowouts. Econ Inq **43**, 316–329 (2005)

Konrad, K.: Strategy in contests-an introduction. WZB Discussion Papers (2007)

Konrad, K.A., Kovenock, D.: Multi-Battle Contests (April 2006). CEPR Discussion Paper No. 5645. Available at SSRN: http://ssrn.com/abstract=918292

Krishna, V., Morgan, J.: The winner-take-all principles in small tournaments. In: Baye, M. (ed.) Advances in Applied Microeconomics. Stamford: JAI Press (1998)

Matros, A.: Elimination Tournaments where Players Have Fixed Resources, working paper, University of Pittsburgh (2005)

Moldovanu, B., Sela, A.: The optimal allocation of prizes in contests. Am Econ Rev **91**, 542–558 (2001)

Moldovanu, B., Sela, A.: Contest architecture. J Econ Theory **126**, 70–96 (2006)

Moldovanu, B., Sela, A., Shi, X.: Contests for status. J Political Econ **115**(2), 338–363 (2007)
Rosen, S.: Prizes and incentives in elimination tournaments. Am Econ Rev **76**, 701–715 (1986)
Skaperdas, S.: Contest success functions. Econ Theory **7**, 283–290 (1996)
Yates, A.J., Heckelman, J.C.: Rent-setting in multiple winner rent-seeking contests. Eur J Political Econ **17**, 835–852 (2001)