Choosing Your Own Luck: Strategic Risk Taking and Effort in Contests*

Kyungmin Kim   R. Vijay Krishna   Dmitry Ryvkin

March 15, 2025

Abstract

We consider the problem of optimal contest design in an environment where contestants choose not only their effort, but also the distribution of shocks affecting their output. The presence of such strategic risk taking has a stark effect on contest design: The winner-take-all contest, whereby the entire prize budget is allocated to the top performer, maximises the expected effort (or output) of the agents for a wide variety of cost functions, including those with fixed costs or scale effects. The results also extend to settings where the designer values greater variability in output.

JEL Classification Codes: C72, C73
Keywords: contest; strategic risk taking; effort; prize allocation

*Kim: Department of Economics, Emory University (email: kyungmin.kim@emory.edu); Krishna: Department of Economics, Florida State University (email: rvk3570@gmail.com); Ryvkin: School of Economics, Finance and Marketing, RMIT University (email: d.ryvkin@gmail.com).

1. Introduction

Contests—allocation mechanisms based on ordinal performance comparisons—are used extensively to motivate agents in organisations and other settings.[1] A central question in the theoretical literature on contests has been that of optimal contest design: How should a fixed budget be distributed across performance ranks? In particular, how does increasing prize inequality affect agents' effort? As elaborated later, the literature provides a number of insightful results. However, existing studies typically consider a restricted setting where players' activities are summarised by a scalar value (effort or output), and they invariably produce nuanced results that depend on the fine details of the model. In particular, the winner-take-all contest could maximise or minimise agents' efforts depending on the structure of effort costs (Moldovanu and Sela, 2001; Fang, Noe, and Strack, 2020) or exogenous noise (Drugov and Ryvkin, 2020).

In this paper, we address the contest design problem in a novel environment where players not only exert effort, but also engage in strategic risk taking by "choosing their own luck." Specifically, as in the standard all-pay contest with complete information, each player selects his effort x_i at cost $c(x_i)$. Concurrently, he also chooses an *arbitrary* unbiased random noise ε_i, so that the final (realised) output is a non-negative random variable $Y_i = x_i + \varepsilon_i \geqslant 0$. This latter element of the model is our main addition to the standard all-pay contest. Of course, the literature has considered *pure risk-taking* contests where each player chooses a distribution with a fixed mean (eg, Myerson, 1993; Ray and Robson, 2012; Fang and Noe, 2022; Fang et al., 2024). Our contribution in modelling is, therefore, to bring together two distinct contest models and study how strategic risk taking influences agents' effort choices, and vice versa, and their joint effect on the design of optimal contests.[2]

We interpret "luck" as the randomness that is inherent in innovative endeavours and problem-solving more generally. This interpretation applies not only when x_i and Y_i correspond to monetary values, but also when they represent physical values of a good: x_i captures the vertical or intrinsic quality of the good, while ε_i represents the horizontal or design aspect of the good. To wit, in 1829, the Liverpool and Manchester railway instituted a contest, the *Rainhill Trials*, for the best design for a locomotive engine that could pull trains between the two cities. A £500 reward awaited the winner (see Taylor, 1995). Within the broad parameters of the trials, contestants offered a wide

[1] Examples include promotions and bonuses (Bognanno, 2001; Baker, Jensen, and Murphy, 1988), sales contests (Lim, Ahearne, and Ham, 2009), forced ranking systems (Bretz Jr., Milkovich, and Read, 1992), and R&D competition (Terwiesch and Ulrich, 2009).

[2] A model in which each player chooses both effort and risk has been studied by Hvide (2002), Kräkel and Sliwka (2004), Gilpatric (2009), and Kim (2018). However, unlike us, they make use of structural assumptions on risk taking, restricting the set of possible distributions of Y_i (ie, the set of possible *joint* distributions of (x_i, ε_i)).

variety of locomotive designs.[3] More recently, the Biomedical Advanced Research and Development Authority (BARDA) announced the Patch Forward Prize, a $50 million competition to advance microneedle patch-based RNA vaccines for COVID-19, seasonal influenza, and pandemic influenza.[4] Such research contests require contestants to take calculated risks in their designs, because the quality or level of output is not guaranteed. The contestants face trade-offs similar to those of financial investors, as more innovative designs have a higher potential upside but are also more likely to fail, while tested designs perform more consistently. Crucially, the degree and the exact "shape" of risk is *endogenously chosen* by the contestants as part of their innovation process, in addition to effort.

Organisations often benefit from the ideas and innovations of "customer-facing" employees. To maximise this knowledge, firms like PwC have instituted in-house platforms (akin to an "app store") where employees provide solutions to common problems that might be useful to others in the firm (see Salvador and Sting, 2022). Incentives for PwC employees to innovate, in the form of monetary compensation as well as performance appraisals, depend on download rates and user evaluations. Here, too, employees come up with design ideas whose uptake is uncertain and depends on the competing ideas of coworkers.[5]

Yet another interpretation of endogenous risk taking is signal jamming. Indeed, as in a standard moral hazard environment, agents may be interested in obfuscating their true effort with nonproductive activities such as self-promotion, engaging in a form of (reverse) Bayesian persuasion. The assumption that noise is mean-preserving then serves as a disciplining constraint similar to the one used in the information design literature.

Our main result is that in contests with effort and strategic risk taking, the winner-take-all contest maximises agents' efforts for all "regular" cost functions (with at most one inflexion point). This is in stark contrast to the existing result for the model without risk taking, namely, that the winner-take-all contest is effort-minimizing (effort-maximising) if the cost function is convex (concave). In other words, in our environment with strategic risk taking, the winner-take-all contest maximisies expected effort regardless of the shape of the cost function.

A key observation for our analysis is that strategic risk taking reduces players' effort costs to produce a stochastic output Y_i in a way that their *virtual* (effective) cost function of output, ξ^*, is concave, regardless of the shape of the underlying effort

[3]The steam powered locomotives included various arrangements of water and fuel which resulted in a different combinations of speed, reliability, and load pulling capacity.

[4]Winners were announced in the spring of 2025.

[5]Gibbs, Neckermann, and Siemroth (2017) find that stronger incentives and rewards foster greater innovation and result in ideas of a higher quality.

cost function c. To see this, suppose a player wishes to produce output y_1 or y_2 with equal probability. If c is concave then randomising over efforts—exerting effort $x_i = y_1$ and $x_i = y_2$ each with probability $1/2$—is more economical than deterministically choosing effort $x_i = (y_1 + y_2)/2$ and then randomising over outputs by choosing $\varepsilon_i = \pm|y_2 - y_1|/2$ with equal probability. In this case, strategic risk taking is irrelevant, and $\xi^* = c$. Conversely, if c is convex then randomising over outputs is more economical than randomising over effort, so the player chooses a deterministic effort x_i and then randomises over outputs. In this case, the resulting virtual cost ξ^* is linear because at the risk-taking stage, the player faces the mean constraint that $\mathbf{E}[Y_i] = x_i$ and ξ^* reflects the corresponding shadow cost. This basic idea applies to the entire relevant region of c and also irrespective of the structure of c. Therefore, the virtual cost of output ξ^* is always concave.

To analyze how an increase in prize inequality affects agents' output choices, it is useful to decompose the overall effect into the following two components: The *prize effect* measures the change in output in response to a change in the prize structure, keeping the virtual cost of output fixed, while the *virtual cost effect* represents the additional impact on output from the equilibrium adjustment of the virtual cost function keeping the prize schedule fixed.

The prize effect rasises effort (in a stochastic sense) as well as its variability. This result directly follows from the existing result on the case of concave costs and the concavity of ξ^*. It then follows that if the virtual cost ξ^* is independent of the prize structure then the optimality of WTA contests is immediate. This is indeed the case when c is concave or convex: If c is concave then ξ^* always coincides with c. If c is convex then ξ^* is an affine function that is independent of the prize schedule.

Beyond the simplest concave or convex cost cases, the virtual cost ξ^* does depend on the prize schedule, in which case one must examine whether the virtual cost effect works in the same direction as the prize effect. Unfortunately, the virtual cost effect is technically challenging to analyze in general, because it is an equilibrium object that can be determined only simultaneously with the equilibrium distribution of output. The resulting complexity prevents us from considering all possible cost functions and also forces us to employ different approaches for different structures. Still, we manage to demonstrate the optimality of WTA contests for the two representative cases, with c (i) initially convex and then concave and (ii) initially concave and then convex. The existing literature has restricted attention to the case where c is either globally concave or globally convex, not only because of their tractability but also because they are sufficient for a nuanced result—namely, that the optimal contest depends on the structure of c. Yet, cost structures (i) and (ii) are relevant for many applications, and the approach we develop in this paper allows us to consider them in a unified manner.

We focus on maximising players' expected effort (equivalently, output). However, our results suggest that the WTA contest is likely to be optimal for a class of objectives that include maximising expected output and the highest output as special cases. This is because we establish a general stochastic order result for the cases where c is concave, convex, or convex-concave: for those three cases, we show that the equilibrium distribution of output from the WTA contest dominates that from any other contest in the increasing convex order. For the concave-convex case, we cannot obtain a similar stochastic order result, because the virtual cost effect opposes the prize effect for certain objectives. Nevertheless, due to the difference in equilibrium expected effort, the WTA contest often remains optimal even if c is concave-convex and the principal's objective is to maximise the highest expected output.

More generally, our results suggest that to the extent real world contests are *not* WTA, it is because the contest designer cares about distributional aspects of output or effort. For instance, a designer may not like a large variability in output which forces the designer to trade off output and variability. Importantly, these tradeoffs for the designer *do not* depend on the shape of the contestants' cost functions.

Related literature. Finding the effort-maximising prize allocation is a classical problem in the contest literature. The latest and most comprehensive treatments (using three different contest models) are by Moldovanu and Sela (2001) for incomplete-information contests with private types; by Fang, Noe, and Strack (2020) for complete-information contests without noise; and by Drugov and Ryvkin (2020) for complete-information contests with exogenous noise à la Lazear and Rosen (1981). In all these models, effort is the *only* choice variable for contestants. Moreover, the aforementioned studies arrive at nuanced results, where the WTA contest is optimal in some cases, but prize sharing is optimal in others, depending on details such as the shape of the cost function (in the first two) or the distribution of noise (in the last, where costs are always assumed to be convex).

The two most closely related papers to ours are those on the effects of prize allocation on flexible risk taking in the absence of effort (ie, with an exogenous mean output). Fang and Noe (2022) consider a principal facing a selection problem: heterogeneous contestants compete for promotion by flexibly selecting stochastic output as a mean-preserving spread of their ability. The authors show that less competitive promotion policies—effectively, more equitable prize schedules—reduce risk taking and lead to improved selection in equilibrium. These results are echoed by Fang et al. (2024) who show, both theoretically and experimentally, that increasing prize inequality leads to more dispersion in output.

Methodologically, we leverage the technical results by Dworczak and Martini (2019). As illustrated in Section 4, the problem of finding the cost-minimising effort distribution for a given distribution of output is an instance of the Bayesian persuasion problem with a continuous state space studied by Dworczak and Martini (2019). We use their results to determine the structure of the virtual cost function ξ^*, which in turn allows us to recover the equilibrium distribution of effort.

The rest of the paper is structured as follows. The model is formally set up in Section 2. In Section 3, we reproduce, for completeness, the arguments from Fang, Noe, and Strack (2020) underlying the results for contests without risk taking and establish our main result for two tractable cases (c either globally concave or globally convex) with elementary methods. In Section 4, we provide the key reformulation of the model via the virtual cost and use this formulation to characterise the equilibrium for four representative cost structures—concave, convex, convex-concave, and concave-convex. Finally, Section 5 contains our comparative statics results with respect to prize inequality and the optimality of winner-take-all. We discuss some extensions in Section 6 and conclude in Section 7.

2. The Model

We build upon the standard complete information all-pay contest. There are n $(\geqslant 2)$ players, each choosing *effort* $x_i \in \mathbb{R}_+$ at a cost according to the common cost function $c \in \mathbb{R}_+^{\mathbb{R}_+}$. We assume that c is twice differentiable, strictly increasing (with $c'(x) > 0$ for $x > 0$), has a finite number of inflexion points, and satisfies $c(0) = 0$ and $c^{-1}(1) < \infty$. The last condition says that there is a finite level of effort whose cost equals the maximum possible benefit from the contest (normalised to one). Reflecting the possibility of mixing, we represent each player i's choice of effort as a non-negative random variable X_i. The associated expected cost of effort is given by $\mathbf{E}[c(X_i)]$.

Strategic risk taking is modelled as follows: Concurrently with effort X_i, each player i chooses a random variable ε_i leading to *output* $Y_i = X_i + \varepsilon_i$, subject to two constraints: (i) $\mathbf{E}[\varepsilon_i \mid X_i] = 0$, and (ii) $Y_i \geqslant 0$ almost surely. In other words, each player can add any unbiased noise to X_i, as long as the resulting output Y_i is non-negative. By definition, Y_i is feasible from X_i if, and only if, Y_i is a non-negative mean-preserving spread of X_i. Such a pair (X_i, Y_i) is said to be *admissible*. As usual, we use X_{-i} and Y_{-i} to denote strategy profiles excluding player i.

A *contest* is defined by a vector $\mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{R}_+^n$, where v_k represents the prize to the player who produces the k-th highest output. We assume that prizes are monotonically decreasing in rank, and the total prize budget is normalised to one. In addition, because setting $v_n > 0$ (ie, giving surplus to the worst performer) is always detrimental to players' incentives, we restrict attention to the prize vectors such that $v_n = 0$. Let $\mathscr{V} := \{\mathbf{v} \in \mathbb{R}_+^n : v_1 \geqslant \ldots \geqslant v_n = 0, \sum_{k=1}^n v_k = 1\}$ denote the set of all prize vectors (contests) that satisfy these restrictions. The usual *winner take all* (WTA) contest corresponds to $\mathbf{v}^{\text{WTA}} = (1, 0, ..., 0)$, while the "punish the bottom" (PTB) contest has $\mathbf{v}^{\text{PTB}} = (\frac{1}{n-1}, \dots, \frac{1}{n-1}, 0)$.

Given $\mathbf{v} \in \mathscr{V}$, a player's payoff is

$$u_i(X_i, Y_i, X_{-i}, Y_{-i}) = \sum_{k=1}^{n} v_k \cdot \mathbf{P}[Y_i \text{ is ranked } k\text{-th}] - \mathbf{E}[c(X_i)].$$

For notational simplicity, we ignore ties, which will arise with zero probability in equilibrium.

As usual, a *Nash equilibrium* is a profile of admissible effort-output combinations, $(X_i^*, Y_i^*)_{i=1}^n$, such that $u_i(X_i^*, Y_i^*, X_{-i}^*, Y_{-i}^*) \geqslant u_i(X_i, Y_i, X_{-i}^*, Y_{-i}^*)$ for all i and for all admissible $(X_i, Y_i)_{i=1}^n$. Following the literature, we focus on symmetric equilibria and use (X^*, Y^*) to denote a symmetric equilibrium strategy, with marginal distributions (F^*, G^*). The next proposition, whose proof is in Appendix E, records existence.

Proposition 2.1. For each prize schedule $\mathbf{v} \in \mathscr{V}$, the contest has a symmetric equilibrium (X^*, Y^*).

As in a few recent studies (Vojnović, 2015; Fang, Noe, and Strack, 2020; Drugov and Ryvkin, 2020), we adopt the *majorisation order* over $\mathscr{V}$ to compare prize schedules in terms of the level of inequality.[6] For $\mathbf{v}, \mathbf{w} \in \mathscr{V}$, $\mathbf{w}$ *majorises* $\mathbf{v}$—or, $\mathbf{w}$ is *more unequal* than $\mathbf{v}$—if $\sum_{i=1}^k (w_i - v_i) \geqslant 0$ for all $k = 1, \dots, n$. Clearly, $\mathbf{v}^{\text{WTA}}$ majorises all $\mathbf{v} \in \mathscr{V}$, while $\mathbf{v}^{\text{PTB}}$ is majorised by any $\mathbf{v} \in \mathscr{V}$. Therefore, $\mathbf{v}^{\text{WTA}}$ is the most unequal contest, while $\mathbf{v}^{\text{PTB}}$ is the most equal contest, in $\mathscr{V}$.

An elementary way to reduce inequality is via a *Pigou-Dalton (PD) transfer*, which reduces the prize to the i-th place and raises the prize to the j-th place by the same amount for $i < j$. Formally, if $\mathbf{w}, \mathbf{v} \in \mathscr{V}$ are such that $v_i = w_i - \delta$ and $v_j = w_j + \delta$ for some $i < j$, with $v_k = w_k$ for all other $k \neq i, j$, then $\mathbf{w}$ is more unequal than $\mathbf{v}$, and $\mathbf{v}$ is obtained from $\mathbf{w}$ via a PD transfer. Importantly, if $\mathbf{w}$ majorises $\mathbf{v}$, then $\mathbf{v}$ can be obtained from $\mathbf{w}$ via a finite sequence of such PD transfers. Therefore, in many instances, in order to prove a comparative static result for the majorisation order it is sufficient to prove it only for an arbitrary PD (or reverse PD) transfer.

3. First Pass: Prior Results and Elementary Analyses

We begin by reproducing prior benchmark results (in particular, those by Fang, Noe, and Strack, 2020) in the standard setting *without* strategic risk taking. We then establish our

[6]For a comprehensive discussion of the majorisation order and its applications, see, eg, Marshall, Olkin, and Arnold (2011).

contest design results for the simplest cases of concave or convex costs via elementary analyses.

3.1. Contests without Risk Taking

Fix a $\mathbf{v} \in \mathscr{V}$ and consider the standard complete information all-pay contest in which each player's output is given by his effort, ie, $Y_i = X_i$. As is well known, this contest has a unique symmetric (mixed strategy) equilibrium in which the distribution of effort $F_i = F$ is continuous and supported on $[0, c^{-1}(v_1)]$, and all players earn zero rents (cf. Barut and Kovenock, 1998). To characterize the equilibrium, suppose a player exerts effort x, while all other players randomise according to F. The indicative player's payoff is then given by $\Phi(F(x); \mathbf{v}) - c(x)$, where $\Phi(F(x); \mathbf{v})$ represents the player's expected winnings from the contest.

The (benefit) function $\Phi(\cdot; \mathbf{v}) : [0,1] \to \mathbb{R}_+$ can be written explicitly as

$$\Phi(q; \mathbf{v}) = \sum_{k=1}^{n} \binom{n-1}{k-1} q^{n-k}(1-q)^{k-1} v_k. \tag{3.1}$$

To understand the structure of $\Phi(q; \mathbf{v})$, suppose a player outperforms every other contestant with probability q (which corresponds to $F(x)$). In order to be ranked k-th and receive prize v_k, the player must be above $n-k$ players while also being below $k-1$ players. For a given set of other players' identities, the probability of this event is $q^{n-k}(1-q)^{k-1}$, and the binomial coefficient $\binom{n-k}{k-1}$ in [3.1] counts the number of ways the other players' identities can be selected. For any $\mathbf{v} \in \mathscr{V}$, $q \mapsto \Phi(q; \mathbf{v}) \in [0, v_1]$ is a continuous and strictly increasing function; thus, $\Phi^{-1}(t; \mathbf{v})$ is also a continuous and strictly increasing function of t over $[0, v_1]$.

Because players earn zero rent in equilibrium, it must be that $\Phi(F(x); \mathbf{v}) = c(x)$ for all $x \in \operatorname{supp}(F)$, which implies that the symmetric equilibrium distribution of effort F is given by

$$F(x) = \begin{cases} \Phi^{-1}(c(x); \mathbf{v}), & c(x) \leqslant v_1 \\ 1, & c(x) > v_1. \end{cases} \tag{3.2}$$

In equilibrium, each player's expected winnings are $1/n$. The zero-rent condition now implies that each individual contestant's expected cost is *also* equal to $1/n$, that is,

$$\frac{1}{n} = \int \Phi(F(x); \mathbf{v})\, \mathrm{d}F(x) = \int c(x)\, \mathrm{d}F(x) \qquad \text{for all } \mathbf{v} \in \mathscr{V}. \tag{3.3}$$

For a given $\mathbf{v} \in \mathscr{V}$, let $X(\mathbf{v})$ denote the random variable corresponding to the equilibrium effort and $F(x;\mathbf{v})$ denote the corresponding distribution, as derived in [3.2]. The following comparative statics then hold as shown by Fang, Noe, and Strack (2020), whose argument we reproduce here for completeness.[7]

Lemma 3.1. Consider prize schedules $\mathbf{v}, \mathbf{w} \in \mathscr{V}$ such that $\mathbf{w}$ is more unequal than $\mathbf{v}$.

(a) If $c(x)$ is concave then $X(\mathbf{w})$ dominates $X(\mathbf{v})$ in the increasing convex order.
(b) If $c(x)$ is convex then $X(\mathbf{v})$ dominates $X(\mathbf{w})$ in the increasing concave order.
(c) If $c(x)$ is linear then $X(\mathbf{w})$ dominates $X(\mathbf{v})$ in the convex order.

Proof. It is sufficient to consider $\mathbf{v}$ and $\mathbf{w}$ such that $\mathbf{w}$ is obtained by a reverse PD transfer from $\mathbf{v}$. Using [3.1], it can be shown that $\Phi(q;\mathbf{w})$ crosses $\Phi(q;\mathbf{v})$ once from below. Together with [3.2], this implies that $F(x;\mathbf{w})$ crosses $F(x;\mathbf{v})$ once from above.

Suppose c is concave and $\mathbf{E}[X(\mathbf{v})] > \mathbf{E}[X(\mathbf{w})]$. Then, the fact that $F(\cdot;\mathbf{w})$ crosses $F(\cdot;\mathbf{v})$ once from above implies that $X(\mathbf{v})$ dominates $X(\mathbf{w})$ in the *increasing concave* order (see, eg, Theorem 4.A.22 of Shaked and Shanthikumar, 2007). But then $\mathbf{E}[c(X(\mathbf{v}))] > \mathbf{E}[c(X(\mathbf{w}))] = 1/n$, which contradicts [3.3], where the strict inequality is because c is strictly increasing. This establishes part (a). Part (b) for the convex case can be analogously proven, while the result for the linear case in part (c) follows from parts (a) and (b). □

Intuitively, greater inequality hurts low performers but helps high performers. Thus, more unequal prizes given players an incentive to "swing for the fences" because higher values of effort—those more likely to result in top prizes—become relatively more profitable. When c is concave, this additional dispersion lowers the expected cost of effort. However, in equilibrium the expected cost of effort stays fixed (cf. the zero-rent condition [3.3]); hence, players exert more effort. By contrast, if c is convex then additional dispersion raises overall effort costs, in which case the expected effort decreases.

Lemma 3.1 highlights the sensitivity of the optimal contest to the structure of the cost function in the absence of risk taking. If costs are concave then more prize inequality raises expected effort; therefore, the WTA contest maximises expected effort, while the PTB contest minimises it. If costs are convex then the effect of prize inequality on expected effort gets reversed, so the WTA contest *minimises* expected effort, while the PTB contest maximises expected effort.

[7] For random variables X_1 and X_2, X_2 *first order stochastically dominates* X_1 if $\mathbf{E}[u(X_2)] \geqslant \mathbf{E}[u(X_1)]$ for any increasing function $u \in \mathbb{R}^{\mathbb{R}}$; similarly, X_2 dominates X^1 in the *(increasing) convex order* if $\mathbf{E}[u(X_2)] \geqslant \mathbf{E}[u(X_1)]$ for any (increasing) convex function $u \in \mathbb{R}^{\mathbb{R}}$. The (increasing) concave order is defined similarly. Dominance in the convex order represents a mean-preserving spread, while dominance in the increasing convex order means that X_2 is both larger *and more variable* than X_1, in a stochastic sense.

3.2. Strategic Risk Taking for Concave or Convex Costs

Next, consider our model in which each player chooses effort X_i and noise ε_i, so his output is given by $Y_i = X_i + \varepsilon_i$ with $\mathbf{E}[\varepsilon_i \mid X_i] = 0$ and $Y_i \geqslant 0$ a.s.

Concave costs. Suppose c is strictly concave. Then, by Jensen's inequality, $\mathbf{E}[c(X_i)] \geqslant \mathbf{E}[c(Y_i)]$ whenever Y_i is a mean-preserving spread of X_i. This implies that no player has an incentive to engage in strategic risk taking; that is, it is optimal to set $X_i = Y_i$—instead of choosing $X_i \neq Y_i$ and then adding noise ε_i—and thus, the (indirect) cost of choosing Y_i is equal to $\mathbf{E}[c(Y_i)]$. It follows that the result for the contest without strategic risk taking continues to apply; in particular, the WTA contest delivers the greatest expected effort.[8]

Suppose c is strictly concave, and player i will choose output Y_i. If $Y_i = X_i + \varepsilon_i$ but $X_i \neq Y_i$ then, by Jensen's inequality, $\mathbf{E}[c(X_i)] > \mathbf{E}[c(Y_i)]$. This implies that it is optimal for player i to eschew strategic risk taking and directly produce Y_i (ie, setting $X_i = Y_i$). With equilibrium therefore devoid of risk taking, the result for the contest *without* risk taking continues to apply (cf, Lemma 3.1(a)); in particular, the WTA contest delivers the greatest expected effort .

Convex costs. An appeal to Jensen again shows that if c is strictly convex then a deterministic effort (ie, a degenerate X_i) is always optimal for each player. Let x_{d} denote the symmetric deterministic equilibrium effort. Given x_{d}, the game reduces to a pure risk-taking contest where each player chooses $Y_i \geqslant 0$ subject to $\mathbf{E}[Y_i] = x_{\mathrm{d}}$. For a fixed level of effort x_{d}, Myerson (1993) provides a general characterization for the pure risk-taking game, establishing that the unique symmetric equilibrium distribution G (of Y) satisfies

$$\Phi(G(y); \mathbf{v}) = \min\left\{\frac{y}{n x_{\mathrm{d}}}, v_1\right\}. \qquad [3.4]$$

As with F, we indicate the dependence of the equilibrium G on $\mathbf{v}$ only as needed.

To identify the equilibrium effort x_{d}, suppose all other players select x_{d} and the distribution G in [3.4]. Then, player i's problem is

$$\max_{x_i, G_i} \int \Phi(G(y); \mathbf{v})\, \mathrm{d}G_i(y) - c(x_i) \quad \text{s.t.} \quad x_i = \int y\, \mathrm{d}G_i(y). \qquad [3.5]$$

[8] If costs are only *weakly* concave, then there may exist equilibria with risk taking. However, all those are outcome(output)-equivalent to those *without* risk taking. For example, if c is linear then there exist both an equilibrium with no risk taking and an equilibrium with no effort randomisation; in fact, all mixtures between them, yielding the same output distribution, are also equilibria. The same caveat applies in the weakly convex case.

As is clear from [3.4], $\Phi(G(y);\mathbf{v})$ is globally concave in y, and so, given x_i, it is (weakly) optimal for player i to eschew strategic risk altogether. Thus, one (but not the only) solution to his problem in [3.5] is to take a degenerate distribution at x_i. This reduces the problem of finding the optimal effort level to

$$\max_{x_i}\left[\min\left\{\frac{x_i}{nx_{\mathrm{d}}},v_1\right\}-c(x_i)\right].$$

From here, it is straightforward that the symmetric equilibrium effort x_{d} is such that $x_{\mathrm{d}}c'(x_{\mathrm{d}})=1/n$.[9]

Note that the equilibrium effort x_{d} is independent of $\mathbf{v}$. This means that contest design does not affect players' effort choices, so the WTA contest produces the same expected effort (or output) as any other contest in $\mathscr{V}$. Notice that this neutrality result is in stark contrast with Lemma 3.1(b), namely, that for convex c, greater prize inequality disincentivises effort, and so the WTA contest minimises expected effort.

Let $Y(\mathbf{v})$ denote the equilibrium output in contest $\mathbf{v}\in\mathscr{V}$. We summarize the findings so far in the following result.[10]

Proposition 3.2. Consider our model of strategic risk taking, and suppose $\mathbf{w}$ is more unequal than $\mathbf{v}$.

(a) If $c(x)$ is concave then $Y(\mathbf{w})$ dominates $Y(\mathbf{v})$ in the increasing convex order.
(b) If $c(x)$ is convex then $Y(\mathbf{w})$ dominates $Y(\mathbf{v})$ in the convex order.

4. Equilibrium Characterisation

We provide herein a general characterisation of equilibrium in contests with risk taking. In particular, we find necessary and sufficient conditions for (F^*,G^*) to constitute an equilibrium, which can be used to identify the equilibrium distributions.

4.1. Necessary Conditions

Fix a contest $\mathbf{v}\in\mathscr{V}$, and let $\mathrm{MPS}(F)$ denote the set of non-negative mean-preserving spreads of F. Clearly, (F^*,G^*) is an equilibrium if and only if it solves the following

---

[9] There exists a unique value of x that satisfies $xc'(x)=1/n$—and so a unique symmetric equilibrium—because the strict monotonicity and convexity of c imply that (i) $xc'(x)$ is strictly increasing and (ii) $xc'(x)>c(x)>1/n$ for all sufficiently large x. Moreover, x_{d} is in the interior of $\mathrm{supp}(G)=[0,nv_1x_{\mathrm{d}}]$.

[10] Part (b) of the proposition follows from part (c) of Lemma 3.1 by noting that [3.4] is equivalent to [3.2] with a linear cost of effort.

problem:

$$[\mathbf{4.1}] \qquad \max_{F,G\in\Delta(\mathbb{R}_+)} \left[\int \Phi(G^*(y);\mathbf{v})\,\mathrm{d}G(y) - \int c(x)\,\mathrm{d}F(x) \right] \quad \text{s.t.} \quad G \in \mathrm{MPS}(F).$$

In other words, (F^*, G^*) should be each player's best response when the other players employ (F^*, G^*).

Strategic risk taking. A necessary condition for (F^*, G^*) to be an equilibrium is that, given F^*, the distribution of output G^* solves the following problem:

$$[\mathbf{4.2}] \qquad \max_{G\in\Delta(\mathbb{R}_+)} \int \Phi(G^*(y);\mathbf{v})\,\mathrm{d}G(y) \quad \text{s.t.} \quad G \in \mathrm{MPS}(F^*).$$

In other words, G^* should be a player's optimal mean-preserving spread of F^* when all other players choose G^*. This means that G^* is a symmetric equilibrium in a generalised pure risk-taking contest in which the players compete by choosing a mean-preserving spread of F^*. Myerson (1993) considers a special case of this contest when F^* is both exogenous *and* degenerate. In our model, as shown in Section 3.2, a degenerate F^* arises *endogenously* when the cost of effort is convex; in general, the (endogenous) equilibrium F^* can be non-degenerate.

In the generalised risk-taking contest, the equilibrium G^* cannot have an interior mass point: If G^* has a mass point at an interior y, then a contestant can make a discrete jump in his expected payoff by splitting the mass at y, eg, between $y+\delta$ and 0. Moreover, $\Phi(G^*;\mathbf{v})$ must be globally concave: If $\Phi(G^*;\mathbf{v})$ is not concave on some interval $[y_1, y_2]$ then a contestant has a profitable deviation whereby all the mass in the interval is moved to the end points y_1 and y_2 while preserving the mean. We record these observations next.

Lemma 4.1. In any equilibrium, $\Phi(G^*(y);\mathbf{v})$ is concave over $\mathbb{R}_+$.

We now relate the optimal distribution of output G^* (and *a fortiori*, the optimal level of risk taking) to the equilibrium effort level F^*. If $\Phi(G^*;\mathbf{v})$ is strictly concave (locally) at some y,[11] then it must be that $G^*(y) = F^*(y)$. Intuitively, local strict risk aversion leads a contestant to forgo further risk taking. Thus, for any interval over which $\Phi(G^*;\mathbf{v})$ is strictly concave, $F^* = G^*$. Together with the fact that $G^* \in \mathrm{MPS}(F^*)$, this result also implies that in any maximal interval where $\Phi(G^*;\mathbf{v})$ is affine, F^* and G^* must coincide at the extreme points and, therefore, share the same mean. This is because intervals of strict concavity and affinity alternate, and there can only be countably many such intervals. Finally, a related argument can be used to show that both

[11] A function $u : \mathbb{R} \to \mathbb{R}$ is *strictly concave at* y if $u(y) > \frac{1}{2}u(y+\delta) + \frac{1}{2}u(y-\delta)$ for all $\delta > 0$ sufficiently small; it is *locally affine at* y if there exists a $\delta > 0$ such that u is affine on $(y-\delta, y+\delta)$.

equilibrium distributions have bounded support.[12] The following result summarises these observations.

Lemma 4.2. In any equilibrium (F^*, G^*), the following hold:

(a) If $\Phi(G^*; \mathbf{v})$ is strictly concave on $[y_1, y_2]$, then $F^*(y) = G^*(y)$ for all $y \in [y_1, y_2]$.

(b) If (y_1, y_2) is a maximal open interval over which $\Phi(G^*; \mathbf{v})$ is affine, then $F^*(y_i) = G^*(y_i)$ for $i = 1, 2$, and $\int_{y_1}^{y_2} y \, \mathrm{d}(F^* - G^*) = 0$.

(c) $\operatorname{supp}(G^*)$ (and hence $\operatorname{supp}(F^*)$) is bounded.

Thus far, we have presented some necessary properties of equilibrium (F^*, G^*). However, Lemmas 4.1 and 4.2 do not describe how to relate (F^*, G^*) to the players' cost of effort c. We illustrate next the role of the cost function c in determining equilibrium effort and output.

Cost minimisation. Given the equilibrium distribution of output G^*, the equilibrium distribution of effort F^* in [**4.1**] must solve

$$\max_{F \in \Delta(\mathbb{R}_+)} \int [-c(x)] \, \mathrm{d}F(x) \quad \text{s.t.} \quad F \in \text{MPC}(G^*), \tag{4.3}$$

where MPC(G) denotes the set of mean-preserving contractions of G. Intuitively, each contestant should produce the target output distribution G^* in the most cost-effective way. Therefore, F^* should be the least costly mean-preserving contraction of G^*. In other words, [**4.3**] is a *cost-minimisation condition* that is necessary for each player's profit maximisation.

The problem in [**4.3**] is far from trivial because of the mean-preserving contraction constraint.[13] Fortuitously, this problem is studied by Dworczak and Martini (2019).[14] The key difference is that in Dworczak and Martini (2019) the distribution G^* is exogenous, whereas G^* is endogenously determined in our setting. Nevertheless, Dworczak and Martini's results imply the following for our cost minimisation problem.

[12] It is clear that no contestant would ever choose an effort level x such that $c(x) > 1$, so $\operatorname{supp}(F^*) \subseteq [0, c^{-1}(1)]$. The necessary argument for G^* is more subtle and can be found in Appendix A. The main idea is that, if $\operatorname{supp}(G^*)$ extends to a region beyond $\operatorname{supp}(F^*)$ then, from Lemma 4.2(b), $\Phi(G^*(y); \mathbf{v})$ is affine in that region; however, $\Phi(\cdot; \mathbf{v})$ is bounded by v_1, and hence the region has to be bounded.

[13] The constraint $F \in$ MPC(G^*) can be rewritten as $\int_0^y [F(t) - G^*(t)] \, \mathrm{d}t \leqslant 0$ for all $y \in \operatorname{supp}(G^*) = [0, \bar{y}]$, with equality at $y = \bar{y}$. This set of inequalities, one for each $y \in [0, \bar{y})$, is a linear constraint in the space of cumulative distribution functions; moreover, the objective is linear in F. This renders [**4.3**] an infinite dimensional linear programming problem; see Dentcheva and Ruszczyński (2003) for a general treatment of optimisation with stochastic order constraints.

[14] Dworczak and Martini (2019) consider a Bayesian persuasion problem in which the state is distributed over an interval according to G^*, and the sender's payoff is some function $u(\cdot)$ that depends only on the mean of the induced posterior. Since the set of feasible distributions of posterior means coincides with MPC(G^*), the sender's problem can be written as [**4.3**] with $u = -c$.

Lemma 4.3. Let (F^*, G^*) be a symmetric equilibrium, so that F^* solves [**4.3**]. Then, there exists a solution ξ^* to the problem

$$\text{[4.4]} \qquad \max_{\xi \in \mathbb{R}^{\mathbb{R}_+}} \int \xi(y)\,\mathrm{d}G^*(y) \qquad \text{s.t.} \quad \xi \leqslant c,\ \xi \text{ concave over } \operatorname{supp}(G^*).$$

Moreover, $\operatorname{supp}(F^*) \subseteq \{y : \xi^*(y) = c(y)\}$ and

$$\text{[4.5]} \qquad \int c\,\mathrm{d}F^* = \int \xi^*\,\mathrm{d}F^* = \int \xi^*\,\mathrm{d}G^*.$$

This result allows us to shift our attention from the direct expected cost of effort, $\int c\,\mathrm{d}F^*$, to the indirect cost of output, $\int \xi^*\,\mathrm{d}G^*$. This offers two significant advantages. First, given ξ^*, the objective function in [**4.1**] depends only, and linearly, on G^*. Second, certain structural properties of ξ^* can be immediately deduced from [**4.4**]. We will shortly illustrate how these advantages can be exploited to characterise the equilibrium distributions.

To understand [**4.5**], observe that the following (weak duality) always holds:

$$\int c\,\mathrm{d}F^* \geqslant \int \xi^*\,\mathrm{d}F^* \geqslant \int \xi^*\,\mathrm{d}G^*.$$

The first inequality follows from the constraint $c \geqslant \xi^*$, and the second one holds because ξ^* is concave and $G^* \in \mathrm{MPS}(F^*)$. Dworczak and Martini (2019) note that [**4.3**] is a linear programming problem and [**4.4**] is its dual. Because their regularity conditions are met in our model (in particular, c is Lipschitz and G^* has a compact support), Theorem 2 of Dworczak and Martini (2019) implies that *strong* duality holds, ie, $\int \xi^*\,\mathrm{d}G^* = \int c\,\mathrm{d}F^*$.

4.2. Virtual Costs and the Equilibrium

Given ξ^*, we can use strong duality in [**4.5**] to write the player's problem [**4.1**] as

$$\text{[4.6]} \qquad \max_{G \in \Delta(\mathbb{R}_+)} \int [\Phi(G^*(y); \mathbf{v}) - \xi^*(y)]\,\mathrm{d}G(y).$$

This is an *unconstrained* linear programming problem, and Lemma 4.1 implies $0 \in \operatorname{supp}(G^*)$; moreover, $G^*(0) = 0$.[15] Hence, it must be that $\Phi(G^*(y); \mathbf{v}) - \xi^*(y) = -\xi^*(0)$ for all $y \in \operatorname{supp}(G^*)$. Recall from Section 3.1 that in the all-pay contest without risk taking, the equilibrium distribution of effort F satisfies $\Phi(F(x); \mathbf{v}) = \min\{c(x), v_1\}$.

[15] The argument is similar to the standard one for all-pay contests without risk taking (see, eg, Hillman and Riley, 1989). If other contestants use a G^* with a mass point at 0, a player can achieve a discrete jump in winnings at an infinitesimal cost by shifting his G^* (and F^*) downward.

Thus, the above result shows that the equilibrium output in our contest coincides with the equilibrium effort (and output) in an all-pay contest *without risk taking* where the cost of effort is given by $\xi^*(y) - \xi^*(0)$. The following proposition, which is our main characterisation result of this section, formalises this discussion.

Proposition 4.4. Suppose (F^*, G^*) is a symmetric equilibrium, and let ξ^* denote a solution to [**4.4**]. Then, $\Phi(G^*(y); \mathbf{v}) = \min\{\xi^*(y) - \xi^*(0), v_1\}$.

In what follows, we refer to the function ξ^* as the *virtual cost function*. As shown in Lemma 4.1, $\Phi(G^*(y); \mathbf{v})$ is concave because G^* is an equilibrium in the generalised pure risk-taking contest, for a given F^*. At the same time, ξ^* is concave because F^* solves the cost minimisation problem [**4.3**], for a given G^*. The equilibrium brings these two pieces of the player's problem together, and the two concave functions are matched, up to an additive constant. This constant, as seen from [**4.6**], is the players' equilibrium rent $-\xi^*(0) \geqslant 0$.

We now illustrate how the results so far can be used to identify equilibria for the four leading cases of cost functions with at most one inflextion point.

Concave Costs. If c is concave then, as shown in panel (a) of Figure 1, $\xi^* = c$ is the solution to the dual problem [**4.4**]. This means that the expected cost of producing (stochastic) output Y_i is equal to $\mathbf{E}[\xi^*(Y_i)] = \mathbf{E}[c(Y_i)]$, which is consistent with our characterisation in Section 3.2, namely, that it is optimal for a player to not engage in risk taking and produce Y_i directly—by randomizing over effort levels—when c is concave. Given this, G^* can be obtained from Proposition 4.4 using $\Phi(G^*(y); \mathbf{v}) = \min\{c(y), v_1\}$, and $F^* = G^*$ solves [**4.3**].

Convex Costs. For expositional ease, we consider the case where c is strictly convex.[16] Then, as depicted in panel (b) of Figure 1, ξ^* is affine and tangent to c at some $m^* > 0$, that is, $\xi^*(y) = \xi^*(0) + c'(m^*)y$. Given this structure, (F^*, G^*) can be found by applying Lemma 4.2 and Proposition 4.4. In particular, since $\operatorname{supp}(F^*) \subseteq \{y : \xi^*(y) = c(y)\}$, any equilibrium necessarily entails a deterministic effort, that is, F^* is degenerate at m^*. Finally, the equilibrium output is distributed according to G^* such that $\Phi(G^*(y); \mathbf{v}) = \min\{c'(m^*)y, v_1\}$ and $m^* = \int y \,\mathrm{d}G^*(y) = 1/[nc'(m^*)]$, as in Section 3.2.

Convex-Concave Costs. Now, suppose c is initially convex and then concave, as depicted in panel (c) of Figure 1. For simplicity, we assume that c is strictly convex below the inflexion point x^ι and strictly concave above x^ι and refer to such cost functions

[16] The subsequent analysis can be extended to accommodate weakly convex functions with affine regions, but it renders the resulting exposition substantially cumbersome without providing any new insights. For example, if c is (only) weakly convex then ξ^* may coincide with c over an interval. In this case, similar to the weakly concave case explained in fn. 8, there still exists a deterministic-effort equilibrium (which is the unique equilibrium if c is strictly convex), but there may exist other equilibria in which the players randomise over the interval such that $\xi^* = c$.

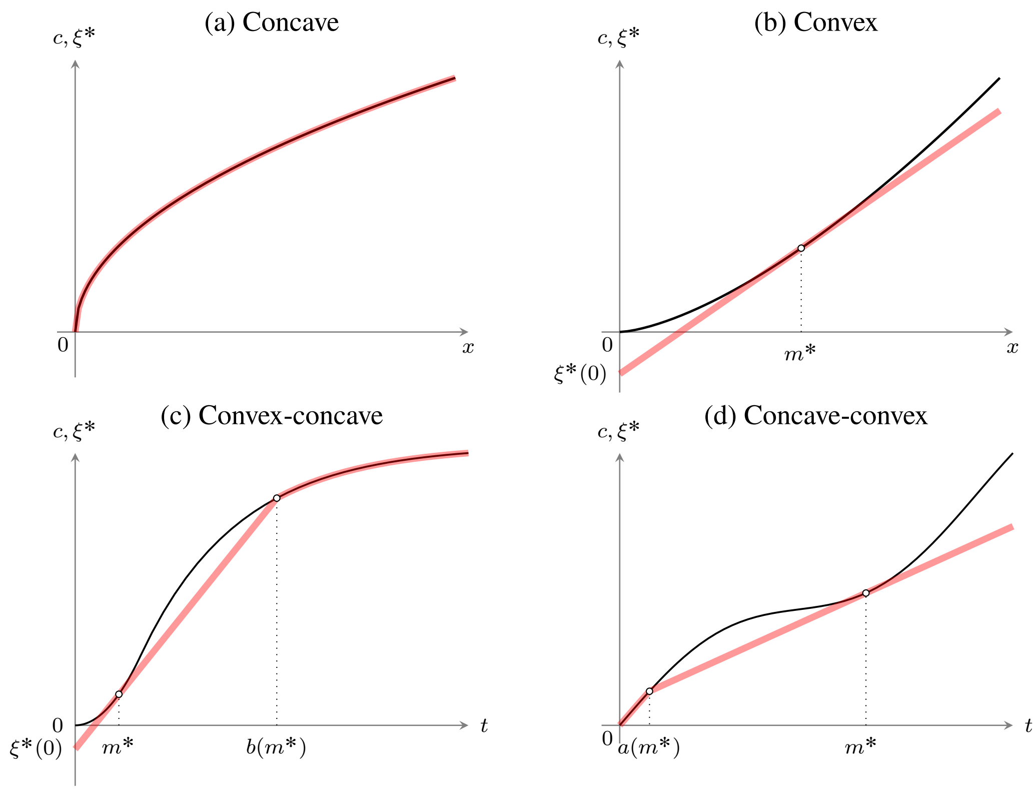

Figure 1 – Cost functions (black, solid) and virtual cost functions (red, translucent).

as *convex-concave*. In this case, as shown in Figure 1(c), the virtual cost function ξ^* has an affine-concave structure. Formally, for each $m \leqslant x^\iota$, let $b(m)(\geqslant x^\iota)$ denote the value such that[17]

$$c(b(m)) - c(m) = c'(m)[b(m) - m],$$

and define $\xi_{\text{xv}}(\cdot; m)$ as the affine-concave function such that

[4.7]
$$\xi_{\text{xv}}(y; m) := \begin{cases} c'(m)(y-m) + c(m) & \text{if } y < b(m) \\ c(y) & \text{if } y \geqslant b(m)\ . \end{cases}$$

In other words, $\xi_{\text{xv}}(\cdot; m)$ is initially affine and tangent to c at m and follows c once it meets c again at $b(m)$. The equilibrium virtual cost function ξ^* belongs to this family $\{\xi_{\text{xv}}(\cdot; m) : m \leqslant x^\iota\}$ of affine-concave functions parametrized by m. We use m^* to denote the value such that ξ^* coincides with $\xi_{\text{xv}}(\cdot; m^*)$. By Proposition 4.4, G^* is such that $\Phi(G^*; \mathbf{v})$ has the same affine-concave structure as ξ^*. In addition, F^* is a non-degenerate distribution, assigning positive probability to m^* and coinciding with G^* above $b(m^*)$.

[17] If $b(m)$ is not well-defined, we simply set $b(m) = \infty$ and interpret ξ_{xv} as a globally affine function.

Concave-Convex Costs. Finally, consider the case where c is initially concave and then convex, as depicted in panel (d) of Figure 1. Similarly to the convex-concave case, we assume that c is strictly concave below the inflexion point x^ι and strictly convex above x^ι. Then, as shown in Figure 1(d), the virtual cost function ξ^* typically has a concave-affine structure.[18] For each $m \geqslant x^\iota$, let $a(m)$ denote the value such that

$$c(m) - c(a(m)) = c'(m)[m - a(m)],$$

and define $\xi_{\text{vx}}(\cdot; m)$ as the concave-affine function such that

$$\xi_{\text{vx}}(y; m) = \begin{cases} c(y) & \text{if } y \leqslant a(m) \\ c'(m)(y - m) + c(m) & \text{if } y > a(m). \end{cases} \quad [\mathbf{4.8}]$$

The virtual cost function ξ^* belongs to this one-parameter family of affine-concave functions, $\{\xi_{\text{vx}}(\cdot; m) : m \geqslant x^\iota\}$, and we let m^* denote the value such that $\xi^* = \xi_{\text{vx}}(\cdot; m^*)$. From here, it follows that $\Phi(G^*; \mathbf{v})$ has the same concave-affine structure as ξ^*, and F^* now coincides with G^* until $a(m^*)$ and assigns all remaining probability to m^*.

4.3. Sufficient Conditions

As illustrated above, the necessary conditions can be used to characterise the equilibrium distributions (F^*, G^*). Of course, this approach is valid only when the necessary conditions are also *sufficient* for a symmetric equilibrium, which is our final result of this section.

Proposition 4.5. Suppose there exist distributions $F^*, G^* \in \Delta(\mathbb{R}_+)$ such that $F^* \in \text{MPC}(G^*)$ and $\Phi(G^*(y); \mathbf{v}) = \min\{\xi^*(y) - \xi^*(0), v_1\}$ for all y, where $\xi^* \in \mathbb{R}^{\mathbb{R}_+}$ solves [**4.4**]. Then (F^*, G^*) is a symmetric equilibrium.

Proof. Let $V(F, G)$ denote the player's expected utility from the objective in [**4.1**] for some $F, G \in \Delta(\mathbb{R}_+)$, with $F \in \text{MPC}(G)$. Then we have

$$\begin{aligned} V(F, G) &= \int \Phi(G^*; \mathbf{v}) \, \mathrm{d}G - \int c \, \mathrm{d}F \leqslant \int [\xi^* - \xi^*(0)] \, \mathrm{d}G - \int c \, \mathrm{d}F \\ &\leqslant -\xi^*(0) + \int \xi^* \, \mathrm{d}F - \int c \, \mathrm{d}F \leqslant -\xi^*(0) = V(F^*, G^*). \end{aligned}$$

[18] The concave-convex cost function can be effectively concave (if the inflexion point x^ι is sufficiently large) or effectively convex (if x^ι is sufficiently small). For those cases, the previous analyses of the concave and convex cases apply effectively unchanged. In what follows, we focus on the case where c is neither effectively concave nor effectively convex. Formally, we assume that $c(x^\iota) < v_1$ and $c(\hat{x}) > 1/n$ where $\hat{x}$ denotes the value that minimises average costs $c(x)/x$ (ie, such that $c'(\hat{x}) = c(\hat{x})/\hat{x}$).

The first inequality holds because $\Phi(G^*; \mathbf{v}) = \min\{\xi^* - \xi^*(0), v_1\} \leqslant \xi^* - \xi^*(0)$; the second inequality holds because ξ^* is concave and $F \in \mathrm{MPC}(G)$; and the third one holds due to the constraint $\xi^* \leqslant c$ in [4.4]. $\square$

5. Comparison of Contests and Optimality of Winner-Take-All

This section analyses the effects of increasing prize inequality on the equilibrium distribution of output. To be specific, let $Y^*(\mathbf{v})$ denote an equilibrium output in contest $\mathbf{v} \in \mathscr{V}$. We then ask: What happens to $Y^*(\mathbf{v})$ if $\mathbf{v}$ becomes more unequal? In particular, how does $Y^*(\mathbf{v}^{\mathrm{WTA}})$ compare to $Y^*(\mathbf{v})$ for any $\mathbf{v} \in \mathscr{V}$?

5.1. Decomposition

Recall that the equilibrium output $Y^*(\mathbf{v})$ is fully determined by the prize schedule $\mathbf{v} \in \mathscr{V}$ and the virtual cost function ξ^* (Proposition 4.4), where ξ^* itself varies with $\mathbf{v}$. For what follows, it is convenient to define the function

$$\Gamma(t; \mathbf{v}) = \begin{cases} \Phi^{-1}(t; \mathbf{v}), & t \leqslant v_1 \\ 1, & t > v_1. \end{cases}$$

Let $\tilde{\xi}^*(y; \mathbf{v}) := \xi^*(y; \mathbf{v}) - \xi^*(0; \mathbf{v})$ denote the virtual cost net of the equilibrium rent.[19] Then, using Proposition 4.4, we can write the equilibrium distribution of output as

$$G^*(y; \mathbf{v}) = \Gamma\big(\tilde{\xi}^*(y; \mathbf{v}); \mathbf{v}\big). \tag{5.1}$$

Equation [5.1] shows explicitly that $\mathbf{v}$ affects the equilibrium distribution of output in two distinct ways: (i) by determining the function $\Gamma(\cdot; \mathbf{v})$ (equivalently, the benefit function $\Phi(\cdot; \mathbf{v})$) and (ii) by influencing the virtual cost function $\tilde{\xi}^*(y; \mathbf{v})$.

Consider two prize schedules $\mathbf{v}, \mathbf{w} \in \mathscr{V}$. The change in the distribution of output from $G^*(\cdot; \mathbf{v})$ to $G^*(\cdot; \mathbf{w})$ can be decomposed as follows:

$$\begin{aligned} G^*(y; \mathbf{w}) - G^*(y; \mathbf{v}) = {} & \underbrace{\Gamma(\tilde{\xi}^*(y; \mathbf{w}); \mathbf{w}) - \Gamma(\tilde{\xi}^*(y; \mathbf{w}); \mathbf{v})}_{\text{prize effect}} \\ & + \underbrace{\Gamma(\tilde{\xi}^*(y; \mathbf{w}); \mathbf{v}) - \Gamma(\tilde{\xi}^*(y; \mathbf{v}); \mathbf{v})}_{\text{virtual cost effect}} \end{aligned} \tag{5.2}$$

where $\Gamma(\tilde{\xi}^*(y; \mathbf{w}); \mathbf{v})$ denotes the distribution of output in the contest with prize schedule

[19] Unless confusion arises, we will refer to both ξ^* and $\tilde{\xi}^*$ as "virtual cost", understanding that they differ by a constant.

$\mathbf{v}$ and the cost of output $\tilde{\xi}^*(y; \mathbf{w})$. In [5.2], the *prize effect* captures the change in output in response to a change in the prize structure (from $\Gamma(\cdot; \mathbf{v})$ to $\Gamma(\cdot; \mathbf{w})$), *keeping the virtual cost fixed*; while the *virtual cost effect* measures the additional impact on output from the equilibrium adjustment in the virtual cost function (from $\tilde{\xi}^*(\cdot; \mathbf{v})$ to $\tilde{\xi}^*(\cdot; \mathbf{w})$), *keeping the benefit function fixed*.

The prize effect is familiar from the model without risk taking, and its properties directly follow from Lemma 3.1 and the fact that the virtual cost function is always concave.

Proposition 5.1. For any $\mathbf{v}, \mathbf{w} \in \mathscr{V}$ such that $\mathbf{w}$ is more unequal than $\mathbf{v}$, $\Gamma(\tilde{\xi}^*(y; \mathbf{w}); \mathbf{w})$ dominates $\Gamma(\tilde{\xi}^*(y; \mathbf{w}); \mathbf{v})$ in the increasing convex order. Moreover, $\Gamma(\tilde{\xi}^*(y; \mathbf{w}); \mathbf{w})$ dominates $\Gamma(\tilde{\xi}^*(y; \mathbf{w}); \mathbf{v})$ in the convex order if and only if $\xi^*(y; \mathbf{w})$ is affine over the support of $G^*(y; \mathbf{w})$.

Thus, in *any* contest with risk taking, the prize effect is the force that moves the distribution of output upwards in the increasing convex order (or, in a special case, the convex order). As we shall see below, the virtual cost effect is more subtle and depends on the structure of c. As observed in Section 4.2, the shape of the cost c determines the players' equilibrium rent $-\xi^*(0)$. Intuitively, when this rent is positive, players respond to increased prize inequality by sacrificing a portion of their rent, which allows them to increase *both effort as well as risk-taking* (in a stochastic sense). However, if the equilibrium rent is fixed at 0 for some cost function, effort can no longer (stochastically) rise when prizes become more unequal because the expected cost must remain at $1/n$. In such a situation, players face more delicate tradeoffs for optimally combining effort and risk taking. Consequently, the virtual cost effect can only be characterised in expectation. In the sequel, we study the virtual cost effect for the four representative classes of cost functions discussed in Section 4.2.

5.2. Concave or Convex Costs

We first illustrate how the decomposition in [5.2] can be combined with Proposition 5.1 to recover Proposition 3.2.

If c is concave then, as illustrated by panel (a) of Figure 1, $\xi^* = c$ is the solution to the dual problem. This means that the equilibrium virtual cost is independent of $\mathbf{v}$, and hence the virtual cost effect in [5.2] is zero. Part (a) of Proposition 3.2—if c is concave then $Y(\mathbf{w})$ dominates $Y(\mathbf{v})$ in the increasing convex order—is then implied by Proposition 5.1.

Next, consider the case where c is convex. Then, as depicted in panel (b) of Figure 1, ξ^* is affine and tangent to c; that is, $\xi^*(y) = \xi^*(0) + c'(m^*)y$ for some $m^*(> 0)$.

By Proposition 4.4, the equilibrium output distribution G^* is such that $\Phi(G^*(y); \mathbf{v}) = \min\{\tilde{\xi}^*(y), v_1\} = \min\{c'(m^*)y; v_1\}$. The value of m^* can now be determined from the following:

$$\frac{1}{n} = \int \Phi(G^*(y); \mathbf{v}) \, \mathrm{d}G^*(y) = c'(m^*) \int y \, \mathrm{d}G^*(y) = c'(m^*)m^*,$$

where the first equality is because each contestant's expected winnings should be $1/n$, the second from Proposition 4.4 and because $\tilde{\xi}^*(y) = c'(m^*)y$, and the last because G^* has mean m^*, as noted in Section 4.2.

A crucial observation is that m^*—and hence, ξ^*—is independent of $\mathbf{v}$, which implies that the virtual cost effect is again zero. Then, part (b) of Proposition 3.2—if c is convex then $Y(\mathbf{w})$ dominates $Y(\mathbf{v})$ in the convex order—also directly follows from Proposition 5.1.

5.3. Convex-Concave Costs

We now examine the case where c is convex-concave. In this case, the virtual cost ξ^* varies with $\mathbf{v}$, and thus the virtual cost effect is not zero.

Characterising the virtual cost effect is technically challenging because ξ^* is determined globally, and simultaneously with the equilibrium output distribution G^*. This limits the generality of our comparative statics result; in particular, we can no longer accommodate arbitrary changes in prize inequality. However, the optimality of WTA still holds as strongly as in the concave case, as formally stated in the following result.

Proposition 5.2. If c is strictly convex-concave then the unique equilibrium output in the WTA contest dominates the largest equilibrium output from any other contest in the increasing convex order.

Notice that Proposition 5.2 takes into account potential equilibrium multiplicity of the convex-concave case. In Appendix C, we present an example in which there are multiple equilibria (Example C.1) but demonstrate that the WTA contest necessarily has a unique equilibrium (Proposition C.3). Finally, we show that multiple equilibria are always clearly ranked by the location of the mass point m^*; in particular, the lower m^* is, the larger the equilibrium output distribution is in the sense of first order stochastic dominance (Lemma C.2). This implies that for the purposes of Proposition 5.2, it suffices to focus on the equilibrium with the lowest value of m^* for each $\mathbf{v}$. In what follows, unless otherwise noted, we refer to the equilibrium with the lowest m^* as the equilibrium under $\mathbf{v} \in \mathscr{V}$ and use $m^*(\mathbf{v})$ to denote this value of m^*.

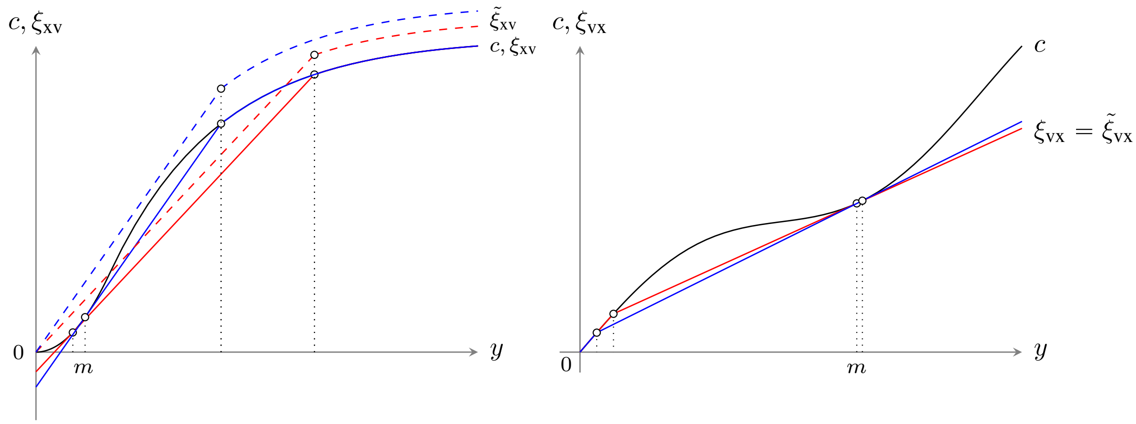

Figure 2 – How the virtual cost functions vary according to m in the convex-concave case (left) and in the concave-convex case (right).

To understand Proposition 5.2, recall from Section 4.2 that in the convex-concave case, the virtual cost function $\tilde{\xi}^*$ is fully determined by m^*, as in [4.7]. Therefore, the virtual cost effect, which characterises how a change in prize inequality affects the output distribution via the virtual cost function, can be identified by knowing (i) how the change in prizes affects m^* and (ii) how a change in m^* affects the output distribution. The following result provides a clear answer for part (ii).

Lemma 5.3. Suppose c is convex-concave, and let $\tilde{\xi}_{\text{xv}}(\cdot;m) := \xi_{\text{xv}}(\cdot;m) - \xi_{\text{xv}}(0;m)$, where $\xi_{\text{xv}}(\cdot;m)$ is defined in [4.8]. Then, an increase of m lowers the distribution $\Gamma(\tilde{\xi}_{\text{xv}}(\cdot;m);\mathbf{v})$ in the sense of first-order stochastic dominance.

Proof. Observe from [4.8] that $\tilde{\xi}_{\text{xv}}$ has the following structure:

$$\tilde{\xi}_{\text{xv}}(y;m) = \begin{cases} c'(m)y & \text{if } y < b(m) \\ c(y) + c'(m)m - c(m) & \text{if } y \geqslant b(m)\ . \end{cases}$$

The result follows because m lies in the convex region of c, so both $c'(m)$ and $c'(m)m - c(m)$ rise in m, and $\Gamma(t;\mathbf{v})$ is increasing in t. □

The left panel of Figure 2 illustrates Lemma 5.3. An increase of m rotates $\xi_{\text{xv}}(\cdot;m)$ around m counterclockwise (from the solid red curve to the solid blue curve). However, $\tilde{\xi}_{\text{xv}}(\cdot;m) = \xi_{\text{xv}}(\cdot;m) - \xi_{\text{xv}}(0;m)$ uniformly rises because $\xi_{\text{xv}}(\cdot;m)$ falls faster at 0 than at any other value; see the dashed curves. This means that if m increases then the players face unambiguously higher (virtual) costs and, therefore, necessarily produce a lower output.

It remains to study part (i), namely, how $m^*(\mathbf{v})$ depends on $\mathbf{v}$. Intuition suggests that $m^*(\mathbf{v})$ falls as $\mathbf{v}$ becomes more unequal: As illustrated in Section 3.1, given ξ^*, increasing prize inequality rotates the equilibrium distribution clockwise, making output more dispersed, i.e., assigning more probability to relatively low or high values. Because m^* is the only low effort level chosen, this tends to push the value of m^* down. We establish this result formally for a restricted class of reverse PD transfers such that the last positive prize is reduced in favor of another prize. We refer to such transfers as *bottom-reducing*.

Lemma 5.4. If $\mathbf{w}$ is obtained from $\mathbf{v}$ via a bottom-reducing transfer then $m^*(\mathbf{w}) \leqslant m^*(\mathbf{v})$.

Note that the restriction to bottom-reducing transfers in Lemma 5.4 is sufficient, but not necessary. The special property of bottom-reducing transfers that enables us to obtain the result is that, by construction, such transfers have the largest discouraging impact on the low end of the effort distribution; that is, they especially reduce the incentive to choose low levels of effort/output.

We conclude this subsection by describing how the results so far can be used to prove Proposition 5.2. Lemmas 5.3 and 5.4 together imply that the virtual cost effect induced by any bottom-reducing transfer raises the output distribution in the sense of first-order stochastic dominance (ie, in the increasing order). Recall that the prize effect always increases the output distribution in the increasing convex order (Proposition 5.1). Because first-order stochastic dominance implies dominance in the increasing convex order and the latter is transitive, it follows that $G^*(\cdot;\mathbf{w}) = \Gamma(\tilde{\xi}^*(\cdot;\mathbf{w});\mathbf{w})$ dominates $G^*(\cdot;\mathbf{v}) = \Gamma(\tilde{\xi}^*(\cdot;\mathbf{v});\mathbf{v})$ in the increasing convex order. Proposition 5.2 ensues because from any $\mathbf{v} \in \mathscr{V}$, $\mathbf{v}^{\mathrm{WTA}}$ can be obtained via a finite sequence of bottom-reducing transfers.[20]

5.4. Concave-Convex Costs

Finally, we consider the case where c is concave-convex, as defined in Section 4.2. In this case, as in the convex-concave case, the virtual cost ξ^* depends on $\mathbf{v}$, so it is crucial to characterise the virtual cost effect. One immediate observation, however, is that we can no longer obtain a stochastic order result. This is because the virtual cost effect is not uniform, as illustrated in the right panel of Figure 2. If m^* increases then the virtual cost function ξ^* rotates counterclockwise around m^* (from the red curve to the blue curve), as in the convex-concave case. But now $\xi^*(0) = 0$; therefore, unlike in the convex-concave case, $\tilde{\xi}^* = \xi^*$ rotates as well: It falls if $y < m^*$ but rises otherwise.

[20]Clearly, the same conclusion holds whenever $\mathbf{w}$ can be obtained from $\mathbf{v}$ via a sequence of bottom-reducing transfers, even if $\mathbf{w} \neq \mathbf{v}^{\mathrm{WTA}}$. This suggests that Proposition 5.2 can be generalised somewhat. However, there may exist multiple equilibria, so the comparison can be made only between the largest equilibrium outputs under $\mathbf{w}$ and $\mathbf{v}$, not between any pair of equilibrium outputs.

Importantly, the virtual cost increases for high levels of output; therefore, this rotation acts counter to the prize effect which calls for a more dispersed output distribution. Nevertheless, we can still obtain the following general comparative statics result, which implies the optimality of WTA contests in terms of expected output.

Proposition 5.5. If c is concave-convex then the equilibrium expected output $\mathbf{E}[Y^*(\mathbf{v})]$ increases as $\mathbf{v}$ becomes more unequal.[21]

We prove this result by showing that the *expected* virtual cost effect is zero, which, given Proposition 5.1, is sufficient for Proposition 5.5. Recall that the virtual cost function $\tilde{\xi}^* = \xi^*$ coincides with $\xi_{\text{vx}}(\cdot; m^*)$ for some $m^* > x^\iota$, where $\xi_{\text{vx}}(\cdot; m)$ is defined by [**4.8**]. As in the convex-concave case, the virtual cost effect can be determined by understanding (i) how $m^*(\mathbf{v})$ is affected by prize inequality and (ii) how m^* affects the output distribution.

For part (ii), consider the expectation $\int y \, \mathrm{d}\Gamma(\xi_{\text{vx}}(y; m^*); \mathbf{v})$ as a function of m^*. As discussed above, if m^* increases then $\xi_{\text{vx}}(y; m^*)$, and hence also $\Gamma(\xi_{\text{vx}}(y; m^*); \mathbf{v})$, falls if $y < m^*$ and rises otherwise. However, the key observation is that in equilibrium these opposing effects exactly cancel each other out, so the *expected* output stays constant. To see this formally, observe that the players' equilibrium payoff is zero and thus, as in the case of concave costs, each player's expected virtual cost should equal his expected winnings:

$$\int \xi_{\text{vx}}(y; m^*(\mathbf{v})) \, \mathrm{d}\Gamma(\xi_{\text{vx}}(y; m^*(\mathbf{v})); \mathbf{v}) = \frac{1}{n}.$$

The strictly concave region of $\xi_{\text{vx}}(y; m^*)$ between 0 and $a(m^*)$ is unaffected by a marginal change in m^*; therefore, the change in $\int y \, \mathrm{d}\Gamma(\xi_{\text{vx}}(y; m^*); \mathbf{v})$ is fully determined by the integral over the affine part of ξ_{vx}. But the equality above implies that the integral over the affine part remains constant as well.

For Proposition 5.5, it is not necessary to characterize part (i) (ie, how $m^*(\mathbf{v})$ depends on $\mathbf{v}$), which is why Proposition 5.5 accommodates arbitrary changes in prize inequality, unlike Proposition 5.2. While it is challenging to characterise the behaviour of $m^*(\mathbf{v})$ for an arbitrary PD transfer, we establish, similar to Proposition 5.2, a clear comparative static for a restricted set of transfers. Specifically, we show that

[21] In the concave-convex case, there always exists a unique equilibrium, so the equilibrium expected output is well-defined for any $\mathbf{v} \in \mathscr{V}$. This result can be obtained as follows: From Lemma 4.2(b), m^* should be the expected output over the affine region of ξ^*, ie, it must satisfy $H(m^*) = 0$, where

$$H(m) := \int_{a(m)}^{\infty} (y - m) \, \mathrm{d}\Gamma(\xi_{\text{vx}}(y; m); \mathbf{v}).$$

It can be shown that H crosses 0 once because $H'(m) < 0$ whenever $H(m) = 0$.

$m^*(\mathbf{w}) \geqslant m^*(\mathbf{v})$ for *top-improving* transfers that raise v_1 while lowering some v_j, $j > 1$, by the same amount (Lemma D.1 in Appendix D).

5.5. Implications for Contest Design

Focusing on the performance of the WTA contest as compared to any $\mathbf{v} \in \mathscr{V}$, the results in this section thus far can be summarised as follows:

- If c is concave or convex-concave then $Y^*(\mathbf{v}^{\mathrm{WTA}})$ dominates $Y^*(\mathbf{v})$ in the increasing convex order.
- If c is convex then $Y^*(\mathbf{v}^{\mathrm{WTA}})$ dominates $Y^*(\mathbf{v})$ in the convex order.
- If c is concave-convex then $\mathbf{E}[Y^*(\mathbf{v}^{\mathrm{WTA}})] \geqslant \mathbf{E}[Y^*(\mathbf{v})]$.

Clearly, the above results imply that if the principal's objective is to maximise expected output (or effort) then the WTA contest is optimal, regardless of the structure of the cost function, in our model of strategic risk taking. As illustrated before, this is in stark contrast to the results for the all-pay contest without risk taking, namely, that the WTA contest is effort-maximising if c is concave but effort-minimising if c is convex. Note also that in the model without risk taking, the effort-maximising contest depends on the fine structure of c if c is convex-concave or concave-convex.

In fact, our results suggest that the WTA contest is likely to be optimal for a more general class of performance measures. To see this formally, let $Y_{(1)} \geqslant \ldots \geqslant Y_{(n)}$ denote the order statistics from n iid draws of the random variable Y, and let $\mathbf{a} \in \mathbb{R}^n_+$ be a vector of weakly declining weights. Then, one can consider the problem of comparing the *weighted sum of order statistics* $Y_{\mathbf{a}} := \sum_{i=1}^n a_i Y_{(i)}$ for different prize structures. This formulation includes both the highest output (when $\mathbf{a} = (1, 0, \ldots, 0)$) and the expected output (when $a_1 = \ldots = a_n = 1/n$) as special cases, but also allows for all intermediate cases where the principal assigns weakly higher weights to higher order statistics (such as considering both expected output and highest output, or selecting a few best ones). Importantly, the random variable $Y_{\mathbf{a}}$ has the following property.

Proposition 5.6. Let $\mathbf{v}$ and $\mathbf{w}$ be two prize schedules such that the equilibrium output distribution $Y(\mathbf{w})$ dominates $Y(\mathbf{v})$ in the increasing convex order. Then, the following hold:

(a) $Y_{(1)}(\mathbf{w})$ dominates $Y_{(1)}(\mathbf{v})$ in the increasing convex order;
(b) $\mathbf{E}[Y_{\mathbf{a}}(\mathbf{w})] \geqslant \mathbf{E}[Y_{\mathbf{a}}(\mathbf{v})]$.

Combined with our earlier results, part (a) implies that the distribution of the highest output is higher in the increasing convex order if the cost function c is concave, convex, or convex-concave. For the same cost functions, part (b) implies that the WTA contest maximises $\mathbf{E}[Y^*_{\mathbf{a}}(\mathbf{v})]$ for *any* vector of declining weights $\mathbf{a} \in \mathbb{R}^n_+$. If c is concave-convex then the result still holds for a broad class of cost functions and objectives (because of the expected output result in Proposition 5.5), but we have not established it in general. The main obstacle is illustrated in Proposition D.3 in Appendix D, where we show that in the case of the highest output with concave-convex costs, the expected prize effect and the expected virtual cost effect always oppose each other.

Finally, while our main focus so far has been on output, it is also of interest to consider the impact of contest design on the equilibrium *effort*, $X^*(\mathbf{v})$. The characterisation in Section 3.2 implies that (i) $X^*(\mathbf{v}^{\mathrm{WTA}})$ dominates $X^*(\mathbf{v})$ in the increasing convex order for any $\mathbf{v} \in \mathscr{V}$ if costs are strictly concave, and (ii) $X^*(\mathbf{v}) = x_{\mathrm{d}}$ is independent of $\mathbf{v}$ when costs are strictly convex. For concave-convex costs, we can establish the same result by showing that $F^*(\cdot;\mathbf{w})$ crosses $F^*(\cdot;\mathbf{v})$ once from above if $\mathbf{w}$ is obtained from $\mathbf{v}$ by a *top-improving* transfer whereby the top prize increases and one of the lower prizes is decreased by the same amount (see Proposition D.2). For convex-concave costs, the single-crossing argument no longer works. However, adopting the same strategy as in Section 5.3, we can show that the unique equilibrium distribution from $\mathbf{v}^{\mathrm{WTA}}$ dominates any other equilibrium from any other contest in the increasing convex order: we construct an effort distribution that corresponds to $\Gamma(\tilde{\xi}(\cdot;\mathbf{w});\mathbf{v})$ and show that it is dominated by $F^*(\cdot;\mathbf{w})$, but dominates $F^*(\cdot;\mathbf{v})$, in the increasing convex order (see Proposition C.5).

6. Discussion

This section discusses two generalizations of our model.

Costly risk taking/reduction. In our model, a player's choice of effort (X_i) is costly, but his choice of risk (ε_i) is not.[22] This latter assumption is in line with the literature on risk-taking contests (eg Myerson, 1993; Hvide, 2002; Ray and Robson, 2012).[23] In fact, it is not even clear which should be costly between increasing risk (dispersing ε_i) and decreasing risk (contracting ε_i): If risk is associated with creativity or gambling, then it is conceivable that choosing a more dispersed distribution is more costly. In the

[22]The choice of risk is costless only conditional on X_i, and the mean-preserving constraint limits how much risk a player can take; in other words, risk taking has an indirect cost because a player should exert more effort to take more risk. Note also that this does not imply that a player's indirect cost of choosing an output random variable Y_i, denoted by $C(Y_i)$, is determined by its mean $\mathbf{E}[Y_i]$. As explained above, if c is convex then a degenerate X_i is always optimal, in which case $C(Y_i) = c(\mathbf{E}[Y_i])$. In general, however, $\delta_{\mathbf{E}[Y_i]} \in \mathrm{MPC}(G_i)$ and so

$$C(Y_i) = \min_{F \in \mathrm{MPC}(G_i)} \int c(x)\,\mathrm{d}F(x) \leqslant c(\mathbf{E}[Y_i])\,.$$

[23]One exception is Gilpatric (2009), who studies a model in which each contestant can pay to control the variance of the output. His main model assumes that increasing the variance (beyond the natural level) is costly; but he also illustrates how his results change if lowering the variance is costly.

context of production and supply management, however, a key objective is to reduce errors, implying that a less dispersed distribution is harder to obtain.

It is beyond the scope of this paper to fully incorporate the cost of risk taking (or reduction) into our model. However, at least conceptually, it is straightforward to accommodate a "small" cost of risk taking/reduction into our main comparative statics result. As shown above, unless c is fully convex over the relevant region, the optimality of $\mathbf{v}^{\text{WTA}}$ is strict and so unlikely to be affected by a small cost of risk. If c is fully convex, however, $\mathbf{v}^{\text{WTA}}$ performs just as well as any other contest in $\mathscr{V}$ in terms of expected output, in which case its optimality depends on the nature of risk costs. If risk taking (increasing risk) is costly then, as in Fang, Noe, and Strack (2020), $\mathbf{v}^{\text{WTA}}$ would perform worse than any other contest, and $\mathbf{v}^{\text{PTB}}$ would be effort-maximizing. This is a consequence of the fact that in our model with convex costs, the equilibrium output distribution under $\mathbf{v}^{\text{WTA}}$ (respectively, $\mathbf{v}^{\text{PTB}}$) is a mean-preserving spread (respectively, contraction) of—and so more (respectively, less) costly than—that from any other contest in $\mathscr{V}$ (see Proposition 3.2). Conversely, if risk reduction (decreasing risk) is costly then the optimality of $\mathbf{v}^{\text{WTA}}$ would strengthen: Reversing the previous logic, $\mathbf{v}^{\text{WTA}}$ would yield a strictly higher effort than other contests even when c is convex.

More general effort cost structures. It is natural to ask whether our results can be extended to a broader class of cost functions permitting multiple inflexion points in the relevant region.[24] The main difficulty lies in the equilibrium characterisation of the virtual cost function ξ^*. When there is at most one inflexion point, the entire family of potential virtual costs has a simple structure with a one-dimensional parameterisation (say, by the location of the point of tangency between the affine segment of ξ^* and c). This allows us to provide a comprehensive characterisation for the set of equilibria and obtain clear comparative statics results.

When c has multiple inflexion points, the structure of ξ^* becomes more complex, with possibly multiple affine segments, either adjacent or alternating with strictly concave segments. Importantly, because ξ^* is determined globally, and simultaneously with the equilibrium output distribution G^*, the number of its affine segments is not fixed and can change with $\mathbf{v}$. That is, ξ^* can "jump" discontinuously with $\mathbf{v}$, at which point G^* also jumps, and we can no longer make use of "local" methods as in Section 5 to analyse global comparative statics. Moreover, our techniques do not allow us to establish equilibrium uniqueness nor rank (possible) multiple equilibria by expected output.

Nevertheless, we can conduct *local* comparative statics when ξ^* is regular, in the following sense. Suppose c has finitely many inflexion points and ξ^* consists of

[24]Having more inflexion points beyond the support of G^* has no impact on our equilibrium analysis and comparative statics.

some number of alternating affine and strictly concave segments such that each strictly concave segment is an interval of positive length. In that case, due to the optimality of ξ^* in the dual problem [4.4], the objective $\int \xi \, \mathrm{d}G^*$ is stationary at ξ^* in each of its affine segments independently. That is, small variations in the points of tangency between ξ^* and c do not change the structure of ξ^*, and local methods similar to those in Section 5 can be used to show that a (particular type of) small increase in prize inequality raises equilibrium expected output. As mentioned above, these local results do not translate into global ones because of the possibility of multiple equilibria and discontinuous changes in the structure of ξ^*.

That said, we are unable to find a counterexample where the optimality of WTA contests for expected output does not hold. We, therefore, conjecture that this result holds in general. Proving it likely requires other methods and is left for future research.

7. Conclusion

We conclude by discussing broader implications of our results and a few potential extensions.

Our model produces an unusually clear and robust prediction, namely, that the winner-take-all contest is optimal for a large class of cost functions and principal objectives. This implies that our model can be more easily falsified than other models. Specifically, the observation that some other contest induces more efforts than the WTA contest immediately falsifies our model. However, it does not falsify Moldovanu and Sela, 2001 or Fang, Noe, and Strack (2020) because the phenomenon does arise under convex costs in these models. It also does not falsify Drugov and Ryvkin (2020) because the same happens (under convex costs) if the (exogenous) shocks are heavy-tailed.

Our analysis is particularly relevant for contest environments where agents are engaged in complex and creative tasks with uncertain outcomes, such as research and innovation contests, architectural design contests, or competition for promotion or bonuses in suitable types of organisations. Therefore, our results suggest that the winner-take-all contest is more likely to be prevalent in those than in other environments where an agent's output mostly depends on his own effort or an agent cannot control noise to his output.

In many cases (eg, when costs are convex), a risk-averse principal who cares about aggregate, or average, output, will also face a trade-off between risk and aggregate efficiency and may prefer to use prize sharing to reduce the variance of effort. For example, for a public research funding agency whose main mission is to support basic research and grow a wide research ecosystem (such as the NSF in the US or the ARC in Australia), it would make sense to fund many projects. The same applies to private foundations focusing on broad agendas, such as the Russel Sage Foundation or the Bill and Melinda Gates Foundation. A similar trade-off is faced by managers in organisations where stakeholders expect stable revenue streams.

A natural extension of our approach is to consider agents with private heterogeneous abilities. In addition to the usual contest design problem, an important application of such a setting is selection contests where the principal's objective is to reward (eg, promote) more able agents. Our techniques allow for a generalisation of Fang and Noe (2022) to continuous distributions of prior abilities. Another application we can generalise is to political competition, similar to Myerson (1993) where we can endogenise politicians' aggregate investments, ie, the "budgets" that politicians have to cultivate minorities. The introduction of endogenous risk taking can also help us contribute to better understanding the moral hazard problem, especially in the context of innovation contests. For instance, in the model of Che and Gale (2003), agents compete in a contest by first making costly investments to determine (private) output and then by participating in a mechanism chosen by the principal. A central assumption in Che and Gale (2003) is that the private research investments completely determine research output. A natural extension of our approach would be to consider such research contests, but where the agents can choose effort *as well as strategic risk*. This clearly changes the incentives for effort and may provide a more robust comparison of different mechanisms and research contest formats.

References

Aliprantis, Charalambos and Kim C. Border (1999). *Infinite Dimensional Analysis: A Hitchhiker's Guide, 2/e*. Springer, New York.

Baker, George P., Michael C. Jensen, and Kevin J. Murphy (1988). "Compensation and incentives: Practice vs. theory." In: *The Journal of Finance* 43.3, pp. 593–616.

Barut, Yasar and Dan Kovenock (1998). "The symmetric multiple prize all-pay auction with complete information." In: *European Journal of Political Economy* 14.4, pp. 627–644.

Baye, Michael R., Guoqiang Tan, and Jianxin Zhou (1993). "Characterizations of the Existence of Equilibria in Games with Discontinuous and Non-quasiconcave Payoffs." In: *Review of Economic Studies* 60.4, pp. 935–948.

Billingsley, Patrick (2012). *Probability and Measure*. Anniversary Edition. New York, NY: Wiley.

Bognanno, Michael L. (2001). "Corporate tournaments." In: *Journal of Labor Economics* 19.2, pp. 290–315.

Bretz Jr., Robert D., George T. Milkovich, and Walter Read (1992). "The current state of performance appraisal research and practice: Concerns, directions, and implications." In: *Journal of Management* 18.2, pp. 321–352.

Carothers, Neal L. (2000). *Real Analysis*. New York: Cambridge University Press.

Che, Yeon-Koo and Ian Gale (2003). "Optimal Design of Research Contests." In: *American Economic Review* 93.3, pp. 646–671.

Chew, Soo Hong, Edi Karni, and Zvi Safra (1987). "Risk Aversion in the Theory of Expected Utility with Rank Dependent Probabilities." In: *Journal of Economic Theory* 42, pp. 370–381.

Dentcheva, Darinka and Andrzej Ruszczyński (2003). "Optimization with Stochastic Dominance Constraints." In: *SIAM Journal on Optimization* 14.2, pp. 548–566.

Drugov, Mikhail and Dmitry Ryvkin (2020). "Tournament Rewards and Heavy Tails." In: *Journal of Economic Theory* 190, pp. 105–116.

Dworczak, Piotr and Giorgio Martini (2019). "The simple economics of optimal persuasion." In: *Journal of Political Economy* 127.5, pp. 1993–2048.

Fang, Dawei and Thomas Noe (2022). "Less competition, more meritocracy?" In: *Journal of Labor Economics* 40.3, pp. 669–701.

Fang, Dawei, Thomas Noe, and Philipp Strack (2020). "Turning up the heat: The discouraging effect of competition in contests." In: *Journal of Political Economy* 128.5, pp. 1940–1975.

Fang, Dawei et al. (2024). “Winning ways: How rank-based incentives shape risk-taking decisions.” In: *Working Paper*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4856341.

Gibbs, Michael, Susanne Neckermann, and Christoph Siemroth (2017). “A Field Experiment in Motivating Employee Ideas.” In: *Review of Economics and Statistics* 99.4, pp. 577–590.

Gilpatric, Scott M. (2009). “Risk taking in contests and the role of carrots and sticks.” In: *Economic Inquiry* 47.2, pp. 266–277.

Hillman, Arye L. and John G. Riley (1989). “Politically contestable rents and transfers.” In: *Economics & Politics* 1.1, pp. 17–39.

Hvide, Hans K. (2002). “Tournament rewards and risk taking.” In: *Journal of Labor Economics* 20.4, pp. 877–898.

Kim, Duk Gyoo (2018). “The second-tier trap: Theory and experimental evidence.” In: *International Journal of Economic Theory* 14.4, pp. 323–349.

Kleiner, Andreas, Benny Moldovanu, and Philipp Strack (2021). “Extreme Points and Majorization: Economic Applications.” In: *Econometrica* 89.4, pp. 1557–1593.

Kräkel, Matthias and Dirk Sliwka (2004). “Risk taking in asymmetric tournaments.” In: *German Economic Review* 5.1, pp. 103–116.

Lazear, Edward P and Sherwin Rosen (1981). “Rank-order Tournaments as Optimum Labor Contracts.” In: *Journal of Political Economy* 89.5, pp. 841–864.

Lim, Noah, Michael J. Ahearne, and Sung H. Ham (2009). “Designing sales contests: Does the prize structure matter?” In: *Journal of Marketing Research* 46.3, pp. 356–371.

Machina, Mark J (1982). “‘Expected Utility’ Analysis without the Independence Axiom.” In: *Econometrica* 50.2, pp. 277–323.

Marshall, Albert W., Ingram Olkin, and Barry C. Arnold (2011). *Inequalities: Theory of majorization and its applications*. Springer.

Milgrom, Paul and Ilya Segal (2002). “Envelope Theorems for Arbitrary Choice Sets.” In: *Econometrica* 70.2, pp. 583–601.

Moldovanu, Benny and Aner Sela (2001). “The optimal allocation of prizes in contests.” In: *American Economic Review* 91.3, pp. 542–558.

Myerson, Roger B. (1993). “Incentives to cultivate favored minorities under alternative electoral systems.” In: *American Political Science Review* 87.4, pp. 856–869.

Pollard, David (2002). *A User’s Guide to Measure Theoretic Probability*. New York, NY: Cambridge University Press.

Ray, Debraj and Arthur Robson (2012). "Status, Intertemporal Choice, and Risk-Taking." In: *Econometrica* 80.4, pp. 1505–1531.

Reny, Philip J. (2020). "Nash Equilibrium in Discontinuous Games." In: *Annual Review of Economics* 12, pp. 439–470.

Salvador, Fabrizio and Fabian J. Sting (2022). "How Your Company Can Encourage Innovation from All Employees." In: *Harvard Business Review* September.

Shaked, Moshe and J. George Shanthikumar (2007). *Stochastic Orders*. New York, NY: Springer.

Taylor, Curtis R. (1995). "Digging for Golden Carrots: An Analysis of Research Tournaments." In: *American Economic Review* 85.4, pp. 872–890.

Terwiesch, Christian and Karl T. Ulrich (2009). *Innovation tournaments: Creating and selecting exceptional opportunities*. Harvard Business Press.

Vojnović, Milan (2015). *Contest Theory: Incentive Mechanisms and Ranking Methods*. Cambridge University Press. ISBN: 9781316472903.

A. Proofs for Section 4

Proof of Lemma 4.1. By definition, G^* should be a solution to

$$\max_G \int \Phi(G^*(y); \mathbf{v})\, \mathrm{d}G(y) \quad \text{s.t. } G \in \text{MPS}(F^*). \tag{A.1}$$

To avoid triviality, assume that F^* is not a degenerate distribution at 0. Suppose $\Phi(G^*(y); \mathbf{v})$ is not concave over $\mathbb{R}_+$. Then, there exist y_1 and y_2 such that $0 \leqslant y_1 < y_2$ and

$$\frac{\Phi(G^*(y_2); \mathbf{v}) - \Phi(G^*(y_1); \mathbf{v})}{y_2 - y_1}(y - y_1) + \Phi(G^*(y_1); \mathbf{v}) > \Phi(G^*(y); \mathbf{v}) \tag{A.2}$$

for all $y \in (y_1, y_2)$.

Consider an alternative distribution $\hat{G}$ that coincides with G^* outside of $[y_1, y_2]$ and assigns the remaining probability $G^*(y_2) - G^*(y_1)$ to y_1 and y_2 so that the mean of G^* is preserved. Formally,

$$\hat{G}(y) := \begin{cases} G^*(y) & y < y_1 \\ G^*(y_1) + (1-\beta)[G^*(y_2) - G(y_1)] & y \in [y_1, y_2) \\ G^*(y) & y \geqslant y_2 \end{cases}$$

where $\beta := \int_{y_1}^{y_2} \frac{y - y_1}{y_2 - y_1}\, \mathrm{d}G^*(y)$. Notice that $(1-\beta)y_1 + \beta y_2 = \int_{y_1}^{y_2} y\, \mathrm{d}G^*(y)$, which ensures

$\int y \, \mathrm{d}\hat{G}(y) = \int y \, \mathrm{d}G^*(y)$. By construction, $\hat{G}$ is a mean-preserving spread of G^*, so $\hat{G} \in \text{MPS}(G^*) \subseteq \text{MPS}(F^*)$.

Restricting attention to the interval $[y_1, y_2]$ (on which $\hat{G}$ differs from G^*), we have

$$\begin{aligned}
&\int_{y_1}^{y_2} \Phi(G^*(y); \mathbf{v}) \, \mathrm{d}\hat{G}(y) \\
&\quad = (1-\beta)\Phi(G^*(y_1); \mathbf{v}) + \beta\Phi(G^*(y_2); \mathbf{v}) \\
&\quad = \Phi(G^*(y_1); \mathbf{v}) \int_{y_1}^{y_2} \frac{y_2 - y}{y_2 - y_1} \, \mathrm{d}G^*(y) + \Phi(G^*(y_2); \mathbf{v}) \int_{y_1}^{y_2} \frac{y - y_1}{y_2 - y_1} \, \mathrm{d}G^*(y) \\
&\quad = \int_{y_1}^{y_2} \left[\frac{\Phi(G^*(y_2); \mathbf{v}) - \Phi(G^*(y_1); \mathbf{v})}{y_2 - y_1} (y - y_1) + \Phi(G^*(y_1); \mathbf{v}) \right] \mathrm{d}G^*(y) \\
&\quad > \int_{y_1}^{y_2} \Phi(G^*(y); \mathbf{v}) \, \mathrm{d}G^*(y)
\end{aligned}$$

where the inequality is due to [**A.2**]. This implies that

$$\int \Phi(G^*(y); \mathbf{v}) \, \mathrm{d}\hat{G}(y) - \int \Phi(G^*(y); \mathbf{v}) \, \mathrm{d}G^*(y) > 0,$$

which contradicts the requirement that G^* solves [**A.1**]. □

Proof of Lemma 4.2. Given that other players play G^*, a player who has chosen effort $x \in \mathbb{R}_+$ faces the following problem:

$$[\textbf{A.3}] \qquad V(x; G^*) := \max_G \int \Phi(G^*(y); \mathbf{v}) \, \mathrm{d}G(y) \quad \text{s.t.} \quad \int y \, \mathrm{d}G(y) = x.$$

Since $\Phi(G^*; \mathbf{v})$ is globally concave (Lemma 4.1), $\int \Phi(G^*; \mathbf{v}) \, \mathrm{d}\tilde{G} \geqslant \int \Phi(G^*; \mathbf{v}) \, \mathrm{d}G$ whenever $G \in \text{MPS}(\tilde{G})$. This implies that δ_x (the degenerate distribution at x) is always a solution to the above problem, that is, $V(x; G^*) = \Phi(G^*(x); \mathbf{v})$ for all $x \in \mathbb{R}_+$.

Suppose $\Phi(G^*; \mathbf{v})$ is strictly concave at x' (see Footnote 11). Together with the global concavity of $\Phi(G^*; \mathbf{v})$, this implies that there exists an affine function ℓ such that $\ell(y) \geqslant \Phi(G^*(y); \mathbf{v})$, with equality holding only when $y = x'$. Then, for any distribution G such that $\int y \, \mathrm{d}G(y) = x'$ and $\text{supp}(G) \neq \{x'\}$, we have

$$\int \Phi(G^*; \mathbf{v}) \, \mathrm{d}G < \int \ell \, \mathrm{d}G = \ell(x) = \Phi(G^*(x'); \mathbf{v}).$$

This suggests that if $\Phi(G^*; \mathbf{v})$ is strictly concave at x' then $\delta_{x'}$ is uniquely optimal for an individual player.

Next, suppose $\Phi(G^*; \mathbf{v})$ is affine on $[y_1, y_2]$ that contains x'. Without loss, assume that $[y_1, y_2]$ is a maximal interval over which $\Phi(G^*; \mathbf{v})$ is affine. Let ℓ denote the affine function that coincides with $\Phi(G^*;\mathbf{v})$ on $[y_1,y_2]$. Since $\Phi(G^*;\mathbf{v})$ is globally concave, we have $\ell(y) \geqslant \Phi(G^*(y);\mathbf{v})$, with equality holding only on $[y_1,y_2]$. Consider any distribution G whose mean is x'. If $\operatorname{supp}(G) \subseteq [y_1,y_2]$ then

$$\int \Phi(G^*;\mathbf{v})\,\mathrm{d}G = \int \ell\,\mathrm{d}G = \ell(x') = \Phi(G^*;\mathbf{v}) = V(x';G^*).$$

Otherwise,

$$\int \Phi(G^*;\mathbf{v})\,\mathrm{d}G < \int \ell\,\mathrm{d}G = V(x';G^*).$$

Therefore, any solution G to [**A.3**] has $\operatorname{supp}(G) \subseteq [y_1,y_2]$.

Suppose $\Phi(G^*;\mathbf{v})$ is strictly concave at x. Then, the above results imply that a player with $x' < x$ will never choose a distribution G such that the upper bound of $\operatorname{supp}(G)$ exceeds x: If $\Phi(G^*;\mathbf{v})$ is strictly concave around x' then the player would simply choose x'. If $\Phi(G^*;\mathbf{v})$ is affine around x' then the player may induce $y > x'$ but will never go beyond x. Similarly, a player with $x' > x$ will never choose a distribution G such that the lower bound of $\operatorname{supp}(G)$ falls short of x. These together imply that $G^*(x) = F^*(x)$, establishing part (a).

Part (b) trivially holds if $\Phi(G^*;\mathbf{v})$ is globally affine over its support. Suppose not. Then, part (a) implies that $F^*(y_i) = G^*(y_i)$ for $i = 1,2$ whenever $[y_1,y_2]$ is a maximal interval over which $\Phi(G^*;\mathbf{v})$ is affine. Combining this with $F^* \in \text{MPS}(G^*)$, it also follows that $\int_{y_1}^{y_2} y\,\mathrm{d}F^* = \int_{y_1}^{y_2} y\,\mathrm{d}G^*$.

It remains to prove that the upper bound of $\operatorname{supp}(G^*)$ is bounded. Let $\overline{y} \leqslant \infty$ denote the upper bound. Lemma 4.1 and parts (a) and (b) imply that $\operatorname{supp}(G^*)$ is an interval starting from 0 that can be partitioned so that over each $[y_1,y_2]$ in the partition, either $\Phi(G^*;\mathbf{v})$ is strictly concave and $G^* = F^*$, or $\Phi(G^*;\mathbf{v})$ is affine and $\int_{y_1}^{y_2} y\,\mathrm{d}G^*(y) = \int_{y_1}^{y_2} y\,\mathrm{d}F^*(y)$. Let $\overline{x}_F (\leqslant c^{-1}(1))$ denote the upper bound of $\operatorname{supp}(F^*)$. If $G^*(\overline{x}_F) = F^*(\overline{x}_F) = 1$ then $\overline{y} = \overline{x}_F$. Otherwise, $\Phi(G^*(y);\mathbf{v})$ must be affine for all $y \in \operatorname{supp}(G^*) \cap (\overline{x}_F, \infty)$. Thus, $\Phi(G^*(y);\mathbf{v})$ has a positive slope for $y > \overline{x}_F$, and hence there exists a finite $\hat{y}$ such that $\Phi(G^*(\hat{y});\mathbf{v}) = v_1$, and $\overline{y} \leqslant \hat{y}$. □

Proof of Lemma 4.3. Our maintained assumptions on c ensure that the conditions of regularity in Dworczak and Martini (2019) hold. In addition, by Lemmas 4.1 and 4.2, $\operatorname{supp}(G^*) = [0,\overline{y}]$ for some $\overline{y}$. The desired result is then immediate from Theorem 2 and Proposition 1 of Dworczak and Martini (2019), when $-c$ and $-\xi^*$ are mapped to u and p, respectively, in their problem. □

B. Proofs for Section 5

Proof of Lemma 5.4. Fix $\mathbf{v} \in \mathscr{V}$, and let $\mathbf{v}^\delta$ denote the prize vector obtained from $\mathbf{v}$ via a bottom-reducing transfer of size δ from $j = \max\{i : v_i > 0\}$ to some $i < j$. We show that $m^*(\delta) := m^*(\mathbf{v}^\delta)$ is strictly decreasing. Recall that we focus on the smallest value of m^* such that $H(m^*, \delta) = 0$, where

$$H(m, \delta) := \int_0^{b(m)} (y - m)\,\mathrm{d}\Gamma(\tilde{\xi}_{\mathrm{xv}}(y; m); \mathbf{v}^\delta).$$

For such m^*, we have $H_m(m^*, \delta) \leqslant 0$. Since $H(m^*(\delta), \delta) = 0$ holds for any δ, the desired result holds if $H_\delta(m^*(0), 0) < 0$. We now prove this inequality.

We first make a few useful observations. For all $y < b(m)$, $\tilde{\xi}_{\mathrm{xv}}(y; m) = c'(m)y$, and hence $\Phi(\Gamma(c'(m)y; \mathbf{v}^\delta); \mathbf{v}^\delta) = c'(m)y$. Differentiating both sides with respect to δ and evaluating them at $\delta = 0$, we have

$$\phi_{ij}(\Gamma(c'(m)y; \mathbf{v})) + \Phi'(\Gamma(c'(m)y; \mathbf{v}); \mathbf{v})\gamma(y; m, \mathbf{v}) = 0 \tag{B.1}$$

where

$$\phi_{ij}(q) := \binom{n-1}{i-1} q^{n-i}(1-q)^{i-1} - \binom{n-1}{j-1} q^{n-j}(1-q)^{j-1} \tag{B.2}$$

and

$$\gamma(y; m, \mathbf{v}) := \frac{\partial}{\partial \delta}\Gamma(c'(m)y; \mathbf{v}^\delta)\bigg|_{\delta=0}.$$

Let q_0 be the unique interior point at which $\phi_{ij}(q_0) = 0$. Then, $\phi_{ij}(q) < 0$ if $q \in (0, q_0)$, while $\phi_{ij}(q) > 0$ if $q \in (q_0, 1)$. Combined with [B.1], this implies that $\gamma(y; m, \mathbf{v}) > 0$ if $q \in (0, q_0)$, while $\gamma(y; m, \mathbf{v}) < 0$ if $q \in (q_0, 1)$.

Integrating by parts, we obtain

$$H(m, \delta) = (b(m) - m)\Gamma(c'(m)b(m); \mathbf{v}^\delta) - \int_0^{b(m)} \Gamma(c'(m)y; \mathbf{v}^\delta)\,\mathrm{d}y.$$

Differentiating $H(m, \delta)$ with respect to δ and evaluating the derivative at $(m, \delta) = (m^*, 0)$, we obtain

$$H_\delta(m^*, 0) = (b(m^*) - m^*)\gamma(b(m^*); m^*, \mathbf{v}) - \int_0^{b(m^*)} \gamma(y; m^*, \mathbf{v})\,\mathrm{d}y.$$

Combining [B.1] with the fact that $\Phi'(\Gamma(c'(m)y; \mathbf{v}); \mathbf{v})\,\mathrm{d}\Gamma(c'(m)y; \mathbf{v}) = c'(m)\,\mathrm{d}y$ for

$y < b(m^*)$ yields

$$-\int_0^{b(m^*)} \gamma(y; m^*, \mathbf{v})\, \mathrm{d}y = \frac{1}{c'(m^*)} \int_0^{\Gamma(c'(m^*)b(m^*);\mathbf{v})} \phi_{ij}(q)\, \mathrm{d}q$$
$$< \frac{1}{c'(m^*)} \int_0^1 \phi_{ij}(q)\, \mathrm{d}q = 0$$

where the inequality holds because $\phi_{ij}(q) > 0$ for $q \in (q_0, 1)$. There are the following two cases to consider: (i) $\gamma(b(m^*); m^*, \mathbf{v}) \leqslant 0$ and (ii) $\gamma(b(m^*); m^*, \mathbf{v}) > 0$. The result ($H_\delta(m^*, 0) < 0$) is straightforward in the former case.

Consider the case where $\gamma(b(m^*); m^*, \mathbf{v}) > 0$, which, by the result above, is equivalent to $\Gamma(c'(m^*)b(m^*); \mathbf{v}) < q_0$; we use the properties of bottom-reducing transfers for this part of the proof. Using the condition $H(m^*, 0) = 0$, $H_\delta(m^*, 0)$ can be rewritten as

$$H_\delta(m^*, 0) = \frac{\gamma(b(m^*); m^*, \mathbf{v})}{\Gamma(c'(m^*)b(m^*); \mathbf{v})} \int_0^{b(m^*)} \Gamma(c'(m^*)y; \mathbf{v})\, \mathrm{d}y - \int_0^{b(m^*)} \gamma(y; m^*, \mathbf{v})\, \mathrm{d}y$$
$$= \frac{\gamma(b(m^*); m^*, \mathbf{v})}{\Gamma(c'(m^*)b(m^*); \mathbf{v})} \int_0^{b(m^*)} \gamma(y; m^*, \mathbf{v}) \left[\frac{\Gamma(c'(m^*)y; \mathbf{v})}{\gamma(y; m^*, \mathbf{v})} - \frac{\Gamma(c'(m^*)b(m^*); \mathbf{v})}{\gamma(b(m^*); m^*, \mathbf{v})} \right] \mathrm{d}y.$$

For $H_\delta(m^*, 0) < 0$, it is sufficient for $\Gamma(c'(m^*)y; \mathbf{v})/\gamma(y; m^*, \mathbf{v})$ to be *increasing* in y—as it implies that the bracketed term is negative for any $y \leqslant b(m^*)$—or equivalently, that $R(q; \mathbf{v}) := -q\Phi'(q; \mathbf{v})/\phi_{ij}(q)$ is increasing in q for $q < q_0$, where we have used [**B.1**] and set $q = \Gamma(c'(m^*)y; \mathbf{v})$. Using the definitions of Φ and ϕ_{ij}, it can be shown that

[**B.3**] $$q\Phi'(q; \mathbf{v}) = (n-1)q^{n-j}(1-q)^{j-1} \sum_{k=1}^{j} \binom{n-2}{k-1} z^{j-k} \Delta v_k$$

and

$$\phi_{ij}(q) = \binom{n-1}{i-1} q^{n-j}(1-q)^{j-1}(z^{j-i} - z_0^{j-i})$$

where $z = q/(1-q)$, $z_0 = q_0/(1-q_0)$, and $\Delta v_k = v_k - v_{k+1} \geqslant 0$. Now, $R(q; \mathbf{v})$ can be written as

$$R(q; \mathbf{v}) = -\frac{q\Phi'(q; \mathbf{v})}{\phi_{ij}(q)} = \frac{(n-1)\sum_{k=1}^{j} \binom{n-2}{k-1} z^{j-k} \Delta v_k}{\binom{n-1}{i-1}(z_0^{j-i} - z^{j-i})}.$$

Clearly, the numerator is increasing in z and the denominator is decreasing in z; therefore, $R(q; \mathbf{v})$ is increasing in z, and hence in q, as required. □

Proof of Proposition 5.5. It suffices to show that at $m = m^*(\mathbf{v})$

$$\frac{\mathrm{d}}{\mathrm{d}m}\int y \,\mathrm{d}\Gamma(\xi_{\mathrm{vx}}(y;m);\mathbf{v})\bigg|_{m=m^*} = \frac{\mathrm{d}}{\mathrm{d}m}\int (1-\Gamma(\xi_{\mathrm{vx}}(y;m);\mathbf{v}))\,\mathrm{d}y\bigg|_{m=m^*}$$
$$= -\int \frac{\mathrm{d}}{\mathrm{d}m}\Gamma(\xi_{\mathrm{vx}}(y;m);\mathbf{v})\,\mathrm{d}y\bigg|_{m=m^*}$$
$$= -\int \Gamma'(\xi_{\mathrm{vx}}(y;m^*);\mathbf{v})\frac{\mathrm{d}\xi_{\mathrm{vx}}(y;m^*)}{\mathrm{d}m}\,\mathrm{d}y = 0$$

where the first equality is via integration by parts. Given the concave-affine structure [4.8] of ξ_{vx}, we have

$$\xi'_{\mathrm{vx}}(y;m) = \begin{cases} c'(y) & \text{if } y < a(m) \\ c'(m) & \text{if } y > a(m). \end{cases}$$

and

$$\frac{\mathrm{d}\xi_{\mathrm{vx}}(y;m)}{\mathrm{d}m} = \begin{cases} 0 & \text{if } y < a(m) \\ c''(m)(y-m) & \text{if } y > a(m). \end{cases}$$

In addition, because $G^*(y;\mathbf{v}) = \Gamma(\xi_{\mathrm{vx}}(y;m^*);\mathbf{v})$ for any $y \in \mathrm{supp}(G^*(\cdot;\mathbf{v}))$, we have

$$\frac{\mathrm{d}G^*(y;\mathbf{v})}{\mathrm{d}y} = \Gamma'(\xi_{\mathrm{vx}}(y;m^*);\mathbf{v})\xi'(y;m^*).$$

Combining all these leads to

$$\frac{\mathrm{d}}{\mathrm{d}m}\int y \,\mathrm{d}\Gamma(\xi_{\mathrm{vx}}(y;m);\mathbf{v})\bigg|_{m=m^*} = -\frac{c''(m^*)}{c'(m^*)}\int_{a(m^*)}^{\infty}(y-m^*)\,\mathrm{d}G^*(y;\mathbf{v}) = 0,$$

where the second equality follows from the fact that $F^* \in \mathsf{MPC}(G^*)$, and so $m^* = \mathbf{E}[Y^*(\mathbf{v}) \mid Y^*(\mathbf{v}) \geqslant a(m^*)]$. ☐

Proof of Proposition 5.6. We make use of the following result, which is an extension of Theorem 1 in Chew, Karni, and Safra (1987).

Lemma B.1. Let $V : \Delta([0,d]) \to \mathbb{R}_+$ be a functional defined as $V(H) := \int v(y)\,\mathrm{d}(\varphi(H(y)),$[25] where v is increasing, convex, and continuously differentiable, and $\varphi \in [0,1]^{[0,1]}$ is strictly increasing, onto, convex, Lipszhitz, and twice continuously differentiable. If G^1 dominates G^2 in the increasing convex order, then $V(G^1) \geqslant V(G^2)$.

[25]The functional V represents preferences in the *rank-dependent utility* model of decision making under risk.

Proof. We first show that V is concave. For each $\alpha \in [0,1]$, let $G_\alpha := \alpha G^1 + (1-\alpha)G^2$. Notice that $V(G) = v(d) - \int \phi(G(y))\,\mathrm{d}v(y)$ via integration by parts. Then, we have

$$\begin{aligned} V(G_\alpha) &= v(d) - \int \phi(G_\alpha)\,\mathrm{d}y \\ &\geqslant \alpha\left[v(d) - \int \phi(G^1)\,\mathrm{d}y\right] + (1-\alpha)\left[v(d) - \int \phi(G^2)\,\mathrm{d}y\right] \\ &= \alpha V(G^1) + (1-\alpha)V(G^2) \end{aligned}$$

where the inequality follows from the convexity of ϕ.

Next, we show that $V'(G_\alpha) := \frac{\mathrm{d}}{\mathrm{d}\alpha}V(G_\alpha) \geqslant 0$. Observe that

$$\begin{aligned} V'(G_\alpha) &= \frac{\mathrm{d}}{\mathrm{d}\alpha}\left[v(d) - \int \phi(\alpha G^1(y) + (1-\alpha)G^2(y))\,\mathrm{d}v(y)\right] \\ &= -\int \frac{\mathrm{d}}{\mathrm{d}\alpha}\left[\phi(\alpha G^1(y) + (1-\alpha)G^2(y))\right]\,\mathrm{d}v(y) \\ &= -\int \phi'(G_\alpha(y))[G^1(y) - G^2(y)]\,\mathrm{d}v(y) \\ &= \int\left[\int_0^y \phi'(G_\alpha(z))\,\mathrm{d}v(z)\right]\,\mathrm{d}(G^1(y) - G^2(y)) \end{aligned}$$

where the last equality is via integration by parts. The integrand in the final expression, $\int_0^y \phi'(G_\alpha(z))\,\mathrm{d}v(z)$, is increasing and convex in y, because

$$\frac{\mathrm{d}}{\mathrm{d}y}\int_0^y \phi'(G_\alpha(z))\,\mathrm{d}v(z) = \phi'(G_\alpha(y))v'(y)$$

and both ϕ and v are non-negative, increasing and convex. Then, $V'(G_\alpha) \geqslant 0$ follows from the fact that $G^1 \geqslant_{\text{icx}} G^2$.

For the final result $V(G^1) \geqslant V(G^2)$, notice that, since V is concave, the fundamental theorem of calculus applies, so

$$V(G^1) - V(G^2) = \int_0^1 V'(G_\alpha)\,\mathrm{d}\alpha.$$

Since $V'(G_\alpha) \geqslant 0$ for all α, $V(G^1) - V(G^2) \geqslant 0$. □

Recall that our goal is to show that if $Y^1 \geqslant_{\text{icx}} Y^2$ then $\mathbf{E}[Y^1_{\mathbf{a}}] \geqslant \mathbf{E}[Y^2_{\mathbf{a}}]$. Let G denote the distribution for a random variable Y and $G_{(i)}$ denote the distribution of $Y_{(i)}$ (the i-th order statistic from n i.i.d. draws of Y, where $Y_{(1)} \geqslant \ldots \geqslant Y_{(n)}$). The latter is given by $G_{(i)}(y) = \varphi_i(G(y))$ where

$$\varphi_i(q) = \sum_{j=n+1-i}^{n} \binom{n}{j} q^j (1-q)^{n-j}$$

for all $q \in [0,1]$. Note that

$$\mathbf{E}[Y_{\mathbf{a}}] = \sum_{i=1}^{n} a_i \, \mathbf{E}[Y_{(i)}] = \sum_{i=1}^{n} a_i \int y \, \mathrm{d}G_{(i)}(y) = \int y \, \mathrm{d} \sum_{i=1}^{n} a_i \varphi_i(G(y)).$$

Define

$$\varphi(q) := \sum_{i=1}^{n} a_i \varphi_i(q) = \sum_{i=1}^{n} a_i \sum_{j=n+1-i}^{n} \binom{n}{j} q^j (1-q)^{n-j} = \sum_{j=1}^{n} A_j \binom{n}{j} q^j (1-q)^{n-j},$$

with $A_j := \sum_{i=n+1-j}^{n} a_i$. Differentiating $\varphi(q)$, we obtain

$$\begin{aligned}
\varphi'(q) &= \sum_{j=1}^{n} A_j \binom{n}{j} q^{j-1} (1-q)^{n-1-j} [j(1-q) - (n-j)q] \\
&= \sum_{j=1}^{n} A_j \frac{n!}{(n-j)!(j-1)!} q^{j-1} (1-q)^{n-j} - \sum_{j=1}^{n} A_j \frac{n!}{(n-1-j)!j!} q^j (1-q)^{n-1-j} \\
&= n \sum_{j=0}^{n-1} (A_{j+1} - A_j) \binom{n-1}{j} q^j (1-q)^{n-1-j} = n \sum_{j=0}^{n-1} a_{n-j} \binom{n-1}{j} q^j (1-q)^{n-1-j}.
\end{aligned}$$

Differentiating once again leads to

$$\varphi''(q) = n(n-1) \sum_{j=0}^{n-2} (a_{n-j-1} - a_{n-j}) \binom{n-2}{j} q^j (1-q)^{n-2-j}.$$

For $a_1 = 0$, the result is trivial. If $a_1 > 0$, we can assume without loss that $\sum_{i=1}^{n} a_i = 1$, and then the above implies that $\varphi \in [0,1]^{[0,1]}$ is strictly increasing, convex, and onto, and the result follows from Lemma B.1. □

C. Further Results for the Convex-Concave Case

This appendix provides further characterization results for the convex-concave case. We begin with an example with multiple equilibria.

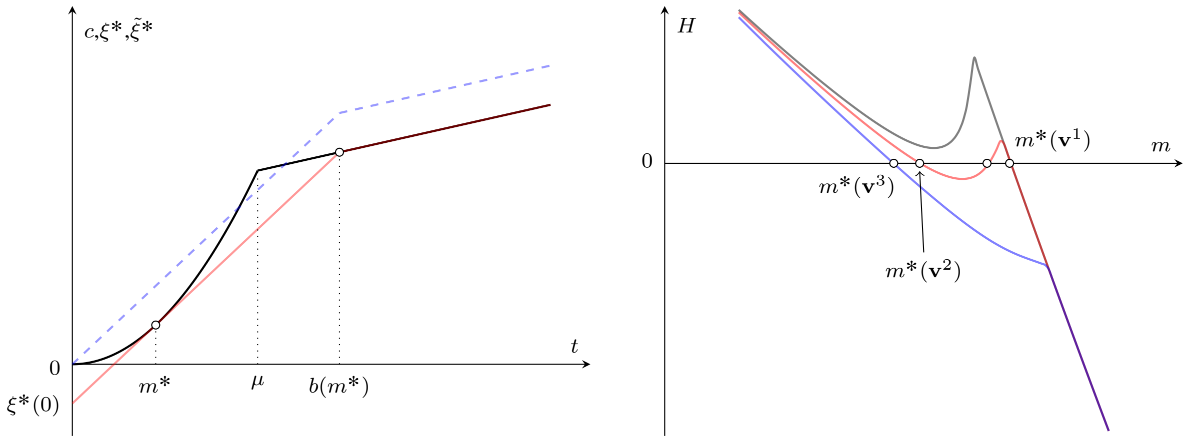

Figure 3 – This figure depicts Example C.1. The left panel shows the cost function c (black, solid) and virtual cost functions ξ^* (red, solid) and $\tilde{\xi}^*$ (blue, dashed). The right panel shows $H(m, \mathbf{v})$ for prize schedules $\mathbf{v}^1 = \frac{1}{4}(1.01, 1, 0.99, 0)$ (black), $\mathbf{v}^2 = \frac{1}{4}(1.025, 1, 0.975, 0)$ (red) and $\mathbf{v}^3 = \frac{1}{4}(1.05, 1, 0.95, 0)$ (blue). Parameter values: $\mu = 0.465$, $\alpha = 0.1$, $n = 4$.

Example C.1. For some $\mu > 0$ and $\alpha > 0$, consider the following cost function:

$$c(x) = \begin{cases} x^2 & \text{if } x \in [0, \mu] \\ \mu^2 + \alpha(x - \mu) & \text{if } x > \mu. \end{cases}$$

As depicted in the left panel of Figure 3, this function is strictly convex below μ and affine above μ. It is easy to see that the non-differentiability of c at μ does not affect the analysis, because it must be that $\xi^*(\mu) < c(\mu)$, so $\mu \notin \operatorname{supp}(F^*)$. In addition, the affine portion can be approximated by strictly concave curves.

Proposition C.3 below identifies two sufficient conditions for the equilibrium uniqueness in the convex-concave case: The first one requires that the marginal cost be sufficiently large in the concave region, and the second one requires that the prize schedule be such that the benefit function $\Phi(q; \mathbf{v})$ is convex. Therefore, to construct an example with multiple equilibria, we choose parameters so as to maximally violate these conditions, by choosing a low value of α and prize schedules close to the punish-the-bottom contest $\mathbf{v}^{\text{PTB}}$, for which $\Phi(q; \mathbf{v})$ is globally concave.

Specifically, we consider contests with $n = 4$ players and three prize schedules, $\mathbf{v}^1$, $\mathbf{v}^2$ and $\mathbf{v}^3$, that differ by bottom-reducing transfers see the caption to Figure 3). Similar to Footnote 21, define the following function:

$$H(m, \mathbf{v}) := \int_0^{b(m)} (y - m)\, \mathrm{d}\Gamma(\xi_{\text{xv}}(y; m); \mathbf{v}).$$

Following the same logic as in Footnote 21, $H(m^*, \mathbf{v}) = 0$ is necessary and sufficient for m^* to yield an equilibrium. As shown in the right panel of Figure 3, $H(\cdot;\mathbf{v})$ crosses 0 and, therefore, the equilibrium is unique in contests $\mathbf{v}^1$ and $\mathbf{v}^3$. However, in contest $\mathbf{v}^2$, $H(\cdot;\mathbf{v})$ crosses 0 three times, and thus there are three equilibria.

Note that, consistent with the proof of Proposition 5.2, the lowest equilibrium m^* shifts to the left. In addition, if we change $\mathbf{v}$ continuously from $\mathbf{v}^1$ to $\mathbf{v}^2$ via a bottom-reducing transfer, the lowest equilibrium m^* will jump discontinuously from $m^*(\mathbf{v}^1)$ to a lower point. With inequality rising further, three equilibria will exist until the local maximum of $H(m,\mathbf{v})$ falls below zero, at which point there is again a unique equilibrium similar to contest $\mathbf{v}^3$.

The following result shows that when there are multiple equilibria, they can be clearly ranked in terms of output.

Lemma C.2. In the convex-concave case, as m^* increases, the equilibrium output distribution G^* falls in the sense of first-order stochastic dominance, while the players' equilibrium expected payoff rises.

Proof. Since the equilibrium m^* necessarily lies in the convex region of c, the function $\tilde{\xi}^*(y) = c'(m^*)y$ for $y \leqslant b(m^*)$ and $\tilde{\xi}^*(y) = c(y) + c'(m^*)m^* - c(m^*)$ for $y > b(m^*)$ (uniformly) increases in m^*. Combined with the fact that $\Gamma(\cdot;\mathbf{v})$ is strictly increasing, this implies that $G^*(y) = \Gamma(\tilde{\xi}^*(y);\mathbf{v})$ rises in m^*. Since this result holds for any $y \geqslant 0$, G^* stochastically decreases. The players' equilibrium payoff is equal to $-\tilde{\xi}^*(0) = c'(m^*)m^* - c(m^*)$. This is increasing in m^*, because $-\,\mathrm{d}\tilde{\xi}^*(0)/\,\mathrm{d}m^* = c''(m^*)m^*$ and m^* always lies in the convex region of c. $\square$

We conclude this appendix by providing two sufficient conditions for equilibrium uniqueness in the convex-concave case.

Proposition C.3. In the convex-concave case, there is a unique equilibrium whenever any of the following conditions holds:

(a) $c'(b(m))b(m) \geqslant c'(m)m$ for all $m \in (0, x^\iota]$;

(b) $\Phi(q;\mathbf{v})$ is convex in q.

Proof. It is convenient to define a modified version of the function H:

$$\tilde{H}(m) = c'(m)\int_0^{b(m)} (y-m)\,\mathrm{d}\Gamma(\xi_{\mathrm{xv}}(y;m);\mathbf{v}).$$

Showing that $\tilde{H}$ is single-crossing from positive to negative in m is, of course, equivalent to showing the same for H. It is also convenient to let $B := \Gamma(\xi_{\mathrm{xv}}(b(m);m);\mathbf{v})$ and $M := \Gamma(\xi_{\mathrm{xv}}(m;m);\mathbf{v})$.

Since $\Phi(\Gamma(\xi_{\mathrm{xv}}(y;m);\mathbf{v});\mathbf{v}) = c'(m)y$ for $y \in [0, \min\{b(m), \overline{y}(m,\mathbf{v})\}]$ where $\overline{y}(m,\mathbf{v})$ is the smallest value such that $\Gamma(\xi_{\mathrm{xv}}(\overline{y}(m,\mathbf{v});m);\mathbf{v}) = 1$, we have

[**C.1**]
$$\begin{aligned}\tilde{H}(m) &= \int_0^{b(m)} c'(m)y \, \mathrm{d}\Gamma(\xi_{\mathrm{xv}}(y;m);\mathbf{v}) - c'(m)mB \\ &= \int_0^{B} \Phi(q;\mathbf{v}) \, \mathrm{d}q - c'(m)mB,\end{aligned}$$

where the second equality follows by changing the variable of integration to $q = \Gamma(\xi_{\mathrm{xv}}(y;m);\mathbf{v})$. The derivative of $\tilde{H}$ is

$$\tilde{H}'(m) = [\Phi(B;\mathbf{v}) - c'(m)m]\frac{\mathrm{d}B}{\mathrm{d}m} - [c''(m)m + c'(m)]B.$$

Note that, since m is in the convex region of c, $c''(m)m + c'(m) > 0$ always holds.

If $b(m) > \overline{y}(m,\mathbf{v})$ then $B = 1$ and so $\mathrm{d}B/\mathrm{d}m = 0$, in which case $\tilde{H}'(m) < 0$. From now on, we consider only the case where $b(m) \leqslant \overline{y}(m,\mathbf{v})$. Note that in that case, $\Phi(B;\mathbf{v}) = c'(m)b(m)$, so

[**C.2**]
$$\frac{\mathrm{d}B}{\mathrm{d}m} = \frac{1}{\Phi'(B;\mathbf{v})}[c''(m)b(m) + c'(m)b'(m)].$$

Part (a): A sufficient condition for $\tilde{H}'(m) < 0$ is $\mathrm{d}B/\mathrm{d}m \leqslant 0$, which is equivalent to $c''(m)b(m) + c'(m)b'(m) \leqslant 0$. We show that the condition in (a) ensures this inequality. Since $c(b(m)) - c(m) = c'(m)(b(m) - m)$, we have

[**C.3**]
$$c'(b(m))b'(m) - c'(m)b'(m) = c''(m)(b(m) - m).$$

It then follows that

$$\begin{aligned}c''(m)b(m) + c'(m)b'(m) &= c''(m)b(m) + c'(m)\frac{c''(m)(b(m) - m)}{c'(b(m)) - c'(m)} \\ &= \frac{c''(m)}{c'(b(m)) - c'(m)}[c'(b(m))b(m) - c'(m)m].\end{aligned}$$

The final expression is negative because $c'(b(m))b(m) - c'(m)m \geqslant 0$ by the given condition and $c'(b(m)) - c'(m) < 0$ due to the convex-concave structure of c and the definition of $b(m)$.

Part (b): From [C.3] and the fact that $b'(m) \leqslant 0$, we have

[**C.4**]
$$c''(m)b(m) + c'(m)b'(m) = c''(m)m + c'(b(m))b'(m) \leqslant c''(m)m,$$

implying

$$\frac{\mathrm{d}B}{\mathrm{d}m} \leqslant \frac{c''(m)m}{\Phi'(B;\mathbf{v})}.$$

Note also that $\Phi(M;\mathbf{v}) = \Phi(\Gamma(\xi_{\text{xv}}(y;m);\mathbf{v});\mathbf{v}) = c'(m)m$. It then follows that

$$\begin{aligned}
\tilde{H}'(m) &= [\Phi(B;\mathbf{v}) - c'(m)m]\frac{\mathrm{d}B}{\mathrm{d}m} - [c''(m)m + c'(m)]B \\
&< [\Phi(B;\mathbf{v}) - \Phi(M;\mathbf{v})]\frac{c''(m)m}{\Phi'(B;\mathbf{v})} - c''(m)mB \\
&\leqslant \frac{c''(m)m}{\Phi'(B;\mathbf{v})}[\Phi(B;\mathbf{v}) - \Phi(M;\mathbf{v}) - \Phi'(B;\mathbf{v})B] < 0,
\end{aligned}$$

where the last inequality follows from the convexity of $\Phi(q;\mathbf{v})$ in q. □

The following corollary of Proposition C.3 is important for our main contest design result, as the WTA contest satisfies the required condition and so necessarily has a unique equilibrium.

Corollary C.4. If $v_i - v_{i+1}$ is decreasing in i then there exists a unique equilibrium in the convex-concave case.

Proof. Differentiating $\Phi(q;\mathbf{v})$ with respect to q and arranging the terms, we arrive at

$$\Phi'(q;\mathbf{v}) = \sum_{k=1}^{n-1} \binom{n-1}{k} k q^{n-1-k}(1-q)^{k-1}(v_k - v_{k+1}).$$

Differentiating this again,

$$\Phi''(q;\mathbf{v}) = \sum_{k=1}^{n-2} \binom{n-1}{k+1} k(k+1) q^{n-2-k}(1-q)^{k-1}\left[(v_k - v_{k+1}) - (v_{k+1} - v_{k+2})\right].$$

This expression is necessarily positive for all q if $v_k - v_{k+1} \geqslant v_{k+1} - v_{k+2}$ for all $k = 1, ..., n-2$. □

We conclude this section by proving that the increasing convex order result in Proposition 5.2 holds for the equilibrium effort distributions as well.

Proposition C.5. If c is strictly convex-concave then the unique equilibrium effort in the WTA contest domintes any equilibrium effort in any other contest in the increasing convex order.

Proof. As in Section 5.3, it suffices to consider $\mathbf{v}, \mathbf{w} \in \mathscr{V}$ such that $\mathbf{w}$ is obtained from $\mathbf{v}$ via a bottom-reducing transfer. Note that, by Lemma 5.4, $m^*(\mathbf{w}) \leqslant m^*(\mathbf{v})$. Let

$F^*(\cdot;\mathbf{w})$ and $F^*(\cdot;\mathbf{v})$ represent the equilibrium distributions corresponding to $\mathbf{w}$ and $\mathbf{v}$, respectively.

Define a distribution $F_{\mathbf{v},\mathbf{w}}(\cdot)$ as follows: It has a mass point at $m^*(\mathbf{w})$ and follows $\Gamma(\tilde{\xi}^*(\cdot;\mathbf{w});\mathbf{v})$ for $x \geqslant \hat{b}$, where $\hat{b}$ is chosen such that $F_{\mathbf{v},\mathbf{w}}(\cdot)$ and $\Gamma(\tilde{\xi}^*(\cdot;\mathbf{w});\mathbf{v})$ have the same mean, that is,

$$\int_0^{\hat{b}} m^*(\mathbf{w})\,\mathrm{d}\Gamma(\tilde{\xi}^*(x;\mathbf{w});\mathbf{v}) + \int_{\hat{b}}^{\infty} x\,\mathrm{d}\Gamma(\tilde{\xi}^*(x;\mathbf{w});\mathbf{v}) = \int x\,\mathrm{d}\Gamma(\tilde{\xi}^*(x;\mathbf{w});\mathbf{v}),$$

which can be rewritten as

$$\int_0^{\hat{b}} (m^*(\mathbf{w}) - x)\ \mathrm{d}\Gamma(\tilde{\xi}^*(x;\mathbf{w});\mathbf{v}) = 0.$$

The value of $\hat{b}$ is well defined because the left-hand side is positive if $\hat{b} = m^*(\mathbf{w})$, decreasing in $\hat{b}(\geqslant m^*(\mathbf{w}))$, and negative if $\hat{b}$ is sufficiently large; this last result holds because $\Gamma(\tilde{\xi}^*(\cdot;\mathbf{w});\mathbf{v})$ first-order stochastically dominates $\Gamma(\tilde{\xi}^*(\cdot;\mathbf{v});\mathbf{v})$, so

$$m^*(\mathbf{w}) \leqslant m^*(\mathbf{v}) \leqslant \int x\,\mathrm{d}\Gamma(\tilde{\xi}^*(x;\mathbf{v});\mathbf{v}) \leqslant \int x\,\mathrm{d}\Gamma(\tilde{\xi}^*(x;\mathbf{w});\mathbf{v}).$$

Similar to [5.2], $F^*(x;\mathbf{w}) - F^*(x;\mathbf{v})$ can be decomposed as follows:

$$F^*(x;\mathbf{w}) - F^*(x;\mathbf{v}) = \underbrace{F^*(x;\mathbf{w}) - F_{\mathbf{v},\mathbf{w}}(x)}_{\text{prize effect}} + \underbrace{F_{\mathbf{v},\mathbf{w}}(x) - F^*(x;\mathbf{v})}_{\text{virtual cost effect}}.$$

We establish the result by showing that both the prize effect and the virtual cost effect raise the effort distribution in the increasing convex order; then, by transitivity, $F^*(\cdot;\mathbf{w})$ dominates $F^*(\cdot;\mathbf{v})$ in the increasing convex order.

For the virtual cost effect, recall that $\tilde{\xi}^*(\cdot;\mathbf{w})$ stays uniformly below $\tilde{\xi}^*(\cdot;\mathbf{v})$, and thus $\Gamma(\tilde{\xi}^*(\cdot;\mathbf{w});\mathbf{v})$ first-order stochastically dominates $G^*(\cdot;\mathbf{v}) = \Gamma(\tilde{\xi}^*(\cdot;\mathbf{v});\mathbf{v})$. By construction, $F_{\mathbf{v},\mathbf{w}}$ coincides with $\Gamma(\tilde{\xi}^*(\cdot;\mathbf{w});\mathbf{v})$ above $\hat{b}$ and is a step function below $\hat{b}$. Combining this with the fact that $m^*(\mathbf{w}) \leqslant m^*(\mathbf{v})$, it follows that $F_{\mathbf{v},\mathbf{w}}$ crosses $F^*(x;\mathbf{v})$ once from above, whether $\hat{b} \geqslant b(m^*(\mathbf{v}))$ or not. The desired result follows because

$$\int x\,\mathrm{d}F_{\mathbf{v},\mathbf{w}}(x) = \int y\,\mathrm{d}\Gamma(\tilde{\xi}^*(y;\mathbf{w});\mathbf{v}) \geqslant \int y\,\mathrm{d}\Gamma(\tilde{\xi}^*(y;\mathbf{v});\mathbf{v}) = \int x\,\mathrm{d}F^*(x;\mathbf{v}).$$

For the prize effect, recall that $G^*(\cdot;\mathbf{w}) = \Gamma(\tilde{\xi}^*(\cdot;\mathbf{w});\mathbf{w})$ crosses $\Gamma(\tilde{\xi}^*(\cdot;\mathbf{w});\mathbf{v})$ once from above (Proposition 5.1). Let $y^\dagger$ denote the interior crossing point between the two output distributions. First, consider the case where $\hat{b} \leqslant b(m^*(\mathbf{w}))$. If $\hat{b} \leqslant y^\dagger$ or $G^*(b(m^*(\mathbf{w}));\mathbf{w}) > \Gamma(\tilde{\xi}^*(\hat{b};\mathbf{w});\mathbf{v})$ then $F^*(\cdot;\mathbf{w})$ necessarily crosses $F_{\mathbf{v},\mathbf{w}}$ once from above. Since $F^*(\cdot;\mathbf{w})$ has a higher mean than $F_{\mathbf{v},\mathbf{w}}$, the former dominates the latter in the increasing convex order. If $\hat{b} > y^\dagger$ and $G^*(b(m^*(\mathbf{w}));\mathbf{w}) \leqslant \Gamma(\tilde{\xi}^*(\hat{b};\mathbf{w});\mathbf{v})$ then $F^*(\cdot;\mathbf{w})$ stays uniformly below $F_{\mathbf{v},\mathbf{w}}$ and, therefore, the former first-order stochastically dominates the latter.

Second, consider the case where $\hat{b} > b(m^*(\mathbf{w}))$. If $b(m^*(\mathbf{w})) \geqslant y^\dagger$ then $F^*(\cdot;\mathbf{w})$ always stays below $F_{\mathbf{v},\mathbf{w}}$, establishing first-order stochastic dominance. If $b(m^*(\mathbf{w})) < y^\dagger$ and $G^*(b(m^*(\mathbf{w}));\mathbf{w}) \geqslant \Gamma(\tilde{\xi}^*(\hat{b};\mathbf{w});\mathbf{v})$ then $F^*(\cdot;\mathbf{w})$ crosses $F_{\mathbf{v},\mathbf{w}}$ once from above, leading to dominance in the increasing convex order as before. We complete the proof by showing that if $\hat{b} > b(m^*(\mathbf{w}))$ then $G^*(b(m^*(\mathbf{w}));\mathbf{w}) \geqslant \Gamma(\tilde{\xi}^*(\hat{b};\mathbf{w});\mathbf{v})$. Toward a contradiction, suppose $G^*(b(m^*(\mathbf{w}));\mathbf{w}) < \Gamma(\tilde{\xi}^*(\hat{b};\mathbf{w});\mathbf{v})$. Recall that the mass point $m^*(\mathbf{w})$ is such that

$$m^*(\mathbf{w}) = \int_0^{b(m^*(\mathbf{w}))} \frac{t\,\mathrm{d}G^*(t;\mathbf{w})}{G^*(b(m^*(\mathbf{w}));\mathbf{w})} = \int_0^{\hat{b}} \frac{t\,\mathrm{d}\Gamma(\tilde{\xi}^*(t;\mathbf{w});\mathbf{v})}{\Gamma(\tilde{\xi}^*(\hat{b};\mathbf{w});\mathbf{v})}.$$

Integrating by parts, we obtain

$$b(m^*(\mathbf{w})) - \int_0^{b(m^*(\mathbf{w}))} \frac{G^*(t;\mathbf{w})}{G^*(b(m^*(\mathbf{w}));\mathbf{w})}\,\mathrm{d}t = \hat{b} - \int_0^{\hat{b}} \frac{\Gamma(\tilde{\xi}^*(t;\mathbf{w});\mathbf{v})}{\Gamma(\tilde{\xi}^*(\hat{b};\mathbf{w});\mathbf{v})}\,\mathrm{d}t.$$

Then, we arrive at the following contradiction:

$$\begin{aligned}
b(m^*(\mathbf{w})) - \hat{b} &= \int_0^{b(m^*(\mathbf{w}))} \frac{G^*(t;\mathbf{w})}{G^*(b(m^*(\mathbf{w}));\mathbf{w})}\,\mathrm{d}t - \int_0^{\hat{b}} \frac{\Gamma(\tilde{\xi}^*(t;\mathbf{w});\mathbf{v})}{\Gamma(\tilde{\xi}^*(\hat{b};\mathbf{w});\mathbf{v})}\,\mathrm{d}t \\
&\geqslant \int_0^{b(m^*(\mathbf{w}))} \frac{\Gamma(\tilde{\xi}^*(t;\mathbf{w});\mathbf{v})}{G^*(b(m^*(\mathbf{w}));\mathbf{w})}\,\mathrm{d}t - \int_0^{\hat{b}} \frac{\Gamma(\tilde{\xi}^*(t;\mathbf{w});\mathbf{v})}{\Gamma(\tilde{\xi}^*(\hat{b};\mathbf{w});\mathbf{v})}\,\mathrm{d}t \\
&\geqslant \int_0^{b(m^*(\mathbf{w}))} \frac{\Gamma(\tilde{\xi}^*(t;\mathbf{w});\mathbf{v})}{\Gamma(\tilde{\xi}^*(\hat{b};\mathbf{w});\mathbf{v})}\,\mathrm{d}t - \int_0^{\hat{b}} \frac{\Gamma(\tilde{\xi}^*(t;\mathbf{w});\mathbf{v})}{\Gamma(\tilde{\xi}^*(\hat{b};\mathbf{w});\mathbf{v})}\,\mathrm{d}t \\
&= \int_{\hat{b}}^{b(m^*(\mathbf{w}))} \frac{\Gamma(\tilde{\xi}^*(t;\mathbf{w});\mathbf{v})}{\Gamma(\tilde{\xi}^*(\hat{b};\mathbf{w});\mathbf{v})}\,\mathrm{d}t > b(m^*(\mathbf{w})) - \hat{b}.
\end{aligned}$$

Here, the first inequality is due to the fact that $G^*(\cdot;\mathbf{w})$ dominates $\Gamma(\tilde{\xi}^*(\cdot;\mathbf{w});\mathbf{v})$ in the increasing convex order; the second inequality follows from the assumption; and the last inequality holds because the integrand is strictly below one for all $t \in [b(m^*(\mathbf{w})), \hat{b})$. □

D. Further Comparative Statics Results for the Concave-Convex Case

The following result shows that in the concave-convex case, any top-improving transfer—raising the prize to the top performer—increases the tangency point m^*.

Lemma D.1. Suppose c is concave-convex. If $\mathbf{w}$ is obtained from $\mathbf{v}$ via a top-improving transfer then $m^*(\mathbf{w}) \geqslant m^*(\mathbf{v})$.

Proof. Fix $\mathbf{v} \in \mathscr{V}$ and $j > 1$ such that $v_j > v_{j+1}$. For each small and positive δ, let $\mathbf{v}^\delta$ denote the contest in $\mathscr{V}$ such that $\mathbf{v}^\delta$ is obtained from $\mathbf{v}^0$ by reducing v_j, while raising v_1 by δ, so that $\mathbf{v}^\delta = \mathbf{v}^0 + \delta(1, \ldots, -1, 0, \ldots, 0)$ where the -1 entry is in the j-th coordinate.

Extending the function H in Footnote 21, define H as follows:

$$H(m, \delta) := \int_{a(m)}^{\overline{y}(m,\delta)} (y - m)\, \mathrm{d}\Gamma(\xi_{\text{vx}}(y; m); \mathbf{v}^\delta)$$

where $\overline{y}(m, \delta)$ denotes the smallest value such that $\Gamma(\xi_{\text{vx}}(\overline{y}(m, \delta); m); \mathbf{v}^\delta) = 1$. For each $\delta \geqslant 0$, by definition, $H(m^*(\mathbf{v}^\delta), \delta) = 0$. In addition, as shown in Footnote 21, $H_m(m^*(\mathbf{v}^\delta), \delta) < 0$. By the implicit function theorem,

$$\frac{\mathrm{d}m^*(\mathbf{v}^0)}{\mathrm{d}\delta} = -\frac{H_\delta(m^*(\mathbf{v}^0), 0)}{H_m(m^*(\mathbf{v}^0), 0)}.$$

Therefore, for the desired result, it suffices to show that $H_\delta(m^*(\mathbf{v}^0), 0) \geqslant 0$.

Integrating $H(m, \delta)$ by parts,

$$H(m, \delta) = -(m - a(m))(1 - \Gamma(\xi_{\text{vx}}(a(m); m); \mathbf{v}^\delta)) + \int_{a(m)}^{\overline{y}(m,\delta)} (1 - \Gamma(\xi_{\text{vx}}(y; m); \mathbf{v}^\delta))\, \mathrm{d}y.$$

For brevity, let $a^* = a(m^*)$ and

$$\gamma(y; m, \mathbf{v}) := \frac{\partial}{\partial \delta} \Gamma(\xi_{\text{vx}}(y; m); \mathbf{v}^\delta)\Big|_{\delta=0}.$$

Then,

$$\text{[D.1]} \qquad H_\delta(m^*, 0) = (m^* - a^*)\Gamma_\delta(\xi_{\text{vx}}(a^*; m^*); \mathbf{v}) - \int_{a^*}^{\overline{y}(m^*,0)} \gamma(y; m^*, \mathbf{v})\, \mathrm{d}y.$$

Consider $y \in (a^*, \overline{y}(m^*, 0)]$. By the definition of Γ, we have

$$\Phi(\Gamma(\xi_{\text{vx}}(y; m^*); \mathbf{v}^\delta); \mathbf{v}^\delta) = \xi_{\text{vx}}(y; m^*) = c(m^*) + c'(m^*)(y - m^*).$$

Since the right-hand side is independent of δ, we have

$$\phi_{ij}(\Gamma(\xi_{\text{vx}}(y; m^*); \mathbf{v}); \mathbf{v}) + \Phi'(\Gamma(\xi_{\text{vx}}(y; m^*); \mathbf{v}); \mathbf{v})\gamma(y; m^*, \mathbf{v}) = 0$$

where ϕ_{1j} is defined in [B.2]. Meanwhile, we also have

$$\Phi'(\Gamma(\xi_{\mathsf{vx}}(y;m^*);\mathbf{v}^\delta);\mathbf{v}^\delta)\,\mathrm{d}\Gamma(\xi_{\mathsf{vx}}(y;m^*);\mathbf{v}^\delta) = c'(m^*)\,\mathrm{d}y.$$

Combining these two equations leads to

$$\begin{aligned}-\int_{a^*}^{\overline{y}(m^*,0)} \gamma(y;m^*,\mathbf{v})\,\mathrm{d}y &= \int_{a^*}^{\overline{y}(m^*,0)} \frac{\phi_{ij}(\Gamma(\xi_{\mathsf{vx}}(y;m^*);\mathbf{v});\mathbf{v})}{c'(m^*)}\,\mathrm{d}\Gamma(\xi_{\mathsf{vx}}(y;m^*);\mathbf{v}) \\ &= \frac{1}{c'(m^*)}\int_{\Gamma(\xi_{\mathsf{vx}}(a^*;m^*);\mathbf{v})}^{1} \phi_{ij}(q)\,\mathrm{d}q > 0.\end{aligned}$$

The inequality holds because $\phi_{1j}(q)$ is a single-crossing function of q, first negative, then positive, and integrates to zero on $[0,1]$. Let q_0 denote the crossing point. In order to sign [D.1], there are two cases to consider: (i) $\gamma(y;m^*,\mathbf{v}) \geqslant 0$ or, equivalently, $\Gamma(\xi_{\mathsf{vx}}(a^*;m^*);\mathbf{v}) \leqslant q_0$, in which case the result follows immediately; and (ii) $\gamma(y;m^*,\mathbf{v}) < 0$ or, equivalently, $\Gamma(\xi_{\mathsf{vx}}(a^*;m^*);\mathbf{v}) > q_0$, in which case the first term in [D.1] is negative and additional steps are needed.

We now consider case (ii). Recall that

$$H(m,\delta) = -(m-a(m))(1-\Gamma(\xi_{\mathsf{vx}}(a(m);m);\mathbf{v}^\delta)) + \int_{a(m)}^{\overline{y}(m,\delta)} (1-\Gamma(\xi_{\mathsf{vx}}(y;m);\mathbf{v}^\delta))\,\mathrm{d}y.$$

The condition $H(m^*,0)=0$ implies

$$m^* - a^* = \frac{1}{1-\Gamma(\xi_{\mathsf{vx}}(a^*;m^*);\mathbf{v})}\int_{a^*}^{\overline{y}(m^*,0)} (1-\Gamma(\xi_{\mathsf{vx}}(y;m^*);\mathbf{v}))\,\mathrm{d}y$$

which allows us to rewrite [D.1] as

$$\begin{aligned}&H_\delta(m^*,0)\\ &\quad= \frac{\gamma(a^*;m^*,\mathbf{v})}{1-\Gamma(\xi_{\mathsf{vx}}(a^*;m^*);\mathbf{v})}\int_{a^*}^{\overline{y}(m^*,0)} (1-\Gamma(\xi_{\mathsf{vx}}(y;m^*);\mathbf{v}))\,\mathrm{d}y - \int_{a^*}^{\overline{y}(m^*,0)} \gamma(y;m^*,\mathbf{v})\,\mathrm{d}y \\ &\quad= \int_{a^*}^{\overline{y}(m^*,0)} (1-\Gamma(\xi_{\mathsf{vx}}(y;m^*);\mathbf{v}))\left[\frac{\gamma(a^*;m^*,\mathbf{v})}{1-\Gamma(\xi_{\mathsf{vx}}(a^*;m^*);\mathbf{v})} - \frac{\gamma(y;m^*,\mathbf{v})}{1-\Gamma(\xi_{\mathsf{vx}}(y;m^*);\mathbf{v})}\right]\mathrm{d}y.\end{aligned}$$

Therefore, for $H_\delta(m^*,0) \geqslant 0$, it suffices to show that $\gamma(y;m^*,\mathbf{v})/(1-\Gamma(\xi_{\mathsf{vx}}(y;m^*);\mathbf{v}))$ is decreasing in y for $\Gamma(\xi_{\mathsf{vx}}(y;m^*);\mathbf{v}) > q_0$ or, equivalently, $R(q;\mathbf{v}) := \phi_{ij}(q)/\left[\Phi'(q;\mathbf{v})(1-q)\right]$ is increasing in q for $q > q_0$.

Similar to [B.3], we have

[D.2] $$\Phi'(q;\mathbf{v}) = (n-1)\sum_{i=1}^{n-1}\binom{n-2}{i-1}q^{n-i-1}(1-q)^{i-1}\Delta v_i.$$

Furthermore, from [B.2] for a transfer $(1,j)$ we have

[D.3] $$\phi_{1j}(q) = q^{n-1} - \binom{n-1}{j-1}q^{n-j}(1-q)^{j-1} = q^{n-1} - \left(\frac{q_0}{1-q_0}\right)^{j-1} q^{n-j}(1-q)^{j-1}.$$

Letting $z = q/(1-q)$ and $z_0 = q_0/(1-q_0)$, $R(q;\mathbf{v})$ can be written as

$$\begin{aligned} R(q;\mathbf{v}) = \frac{\phi_{1j}(q)}{\Phi'(q;\mathbf{v})(1-q)} &= \frac{q^{n-1} - z_0^{j-1}q^{n-j}(1-q)^{j-1}}{(n-1)\sum_{i=1}^{n-1}\binom{n-2}{i-1}q^{n-i-1}(1-q)^{i}\Delta v_i} \\ &= \frac{z^{n-1} - z_0^{j-1}z^{n-j}}{(n-1)\sum_{i=1}^{n-1}\binom{n-2}{i-1}z^{n-i-1}\Delta v_i}. \end{aligned}$$

The derivative of $R(q;\mathbf{v})$ with respect to z is, up to a positive multiplier,

$$\begin{aligned} \frac{\mathrm{d}R(q;\mathbf{v})}{\mathrm{d}z} \propto{}& [(n-1)z^{n-2} - (n-j)z_0^{j-1}z^{n-j-1}]\sum_{i=1}^{n-1}\binom{n-2}{i-1}z^{n-i-1}\Delta v_i \\ & - (z^{n-1} - z_0^{j-1}z^{n-j})\sum_{i=1}^{n-1}\binom{n-2}{i-1}(n-i-1)z^{n-i-2}\Delta v_i \\ ={}& \sum_{i=1}^{n-1}\binom{n-2}{i-1}z^{n-i-1}\Delta v_i z^{2n-i-j-2}\big[(n-1)z^{j-1} - (n-j)z_0^{j-1} \\ & - (n-i-1)(z^{j-1} - z_0^{j-1})\big] \\ ={}& \sum_{i=1}^{n-1}\binom{n-2}{i-1}z^{n-i-1}\Delta v_i z^{2n-i-j-2}[iz^{j-1} + (j-i-1)z_0^{j-1}] \\ >{}& \sum_{i=1}^{n-1}\binom{n-2}{i-1}z^{n-i-1}\Delta v_i z^{2n-i-j-2}(j-1)z_0^{j-1} > 0, \end{aligned}$$

where the first inequality is because $z > z_0$. Thus, $R(q;\mathbf{v})$ is increasing in z for $z > z_0$ and hence it is also increasing in q for $q > q_0$. □

Lemma D.1 describes the equilibrium adjustment in the virtual cost function $\xi^*_{\text{vx}}(y;\mathbf{v})$ in response to a top-improving transfer. It is easy to see that if the mass point shifts up, $m^*(\mathbf{w}) > m^*(\mathbf{v})$, then $F^*(\cdot;\mathbf{w})$ crosses $F^*(\cdot;\mathbf{v})$ once from above. Since we already know from Proposition 5.5 that the expected effort goes up as well, Theorem 4.A.22 from Shaked and Shanthikumar (2007) implies the following result.

Proposition D.2. Suppose c is concave-convex and prize schedules $\mathbf{v}, \mathbf{w} \in \mathscr{V}$ are such that $\mathbf{w}$ is obtained from $\mathbf{v}$ via a sequence of top-improving transfers. Then $X^*(\mathbf{w})$ dominates $X^*(\mathbf{v})$ in the increasing convex order.

We proceed to show that for top-improving transfers, the virtual cost effect lowers the expected highest output. Recall that the virtual cost effect is concerned with the change from $\Gamma(\xi^*_{\text{vx}}(y;\mathbf{v});\mathbf{v})$ to $\Gamma(\xi^*_{\text{vx}}(y;\mathbf{w});\mathbf{v})$ ([5.2]) and ξ^*_{vx} has the following concave-affine structure in the concave-convex case:

$$\xi_{\text{vx}}(y;m^*) = \begin{cases} c(y) & \text{if } y \leqslant a(m^*) \\ c'(m^*)(y-m^*)+c(m^*) & \text{if } y > a(m^*). \end{cases}$$

Given Lemma D.1, the following result suffices for our result.

Proposition D.3. If c is concave-convex then for any $n > 1$,

$$\frac{\mathrm{d}}{\mathrm{d}m^*}\int y \,\mathrm{d}\Gamma(\xi_{\text{vx}}(y;m^*(\mathbf{v}));\mathbf{v})^n \leqslant 0.$$

Proof. Applying integration by parts,

$$\overline{H}(m^*) := \int y\,\mathrm{d}\Gamma(\xi_{\text{vx}}(y;m^*);\mathbf{v})^n = \int \left(1-\Gamma(\xi_{\text{vx}}(y;m^*);\mathbf{v})^n\right)\,\mathrm{d}y.$$

Therefore,

$$\overline{H}'(m^*) = -n\int \Gamma(\xi_{\text{vx}}(y;m^*);\mathbf{v})^{n-1}\frac{\mathrm{d}\Gamma(\xi_{\text{vx}}(y;m^*);\mathbf{v})}{\mathrm{d}m^*}\,\mathrm{d}y.$$

Recall that $\mathrm{d}\Gamma(\xi_{\text{vx}}(y;m^*);\mathbf{v})/\,\mathrm{d}m^* < 0$ if $y \in (a(m^*), m^*)$, while $\mathrm{d}\Gamma(\xi_{\text{vx}}(y;m^*);\mathbf{v})/\,\mathrm{d}m^* > 0$ if $y > m^*$ (see the right panel of Figure 2). In addition, $\int \frac{\mathrm{d}\Gamma(\xi_{\text{vx}}(y;m^*);\mathbf{v})}{\mathrm{d}m^*}\,\mathrm{d}y = 0$ (see the proof of Proposition 5.5 in Appendix B). Combining these with the fact that

$\Gamma(\xi_{\text{vx}}(y; m^*); \mathbf{v})^{n-1}$ is increasing, it follows that

$$\begin{aligned}
&\int \Gamma(\xi_{\text{vx}}(y; m^*); \mathbf{v})^{n-1} \frac{\mathrm{d}\Gamma(\xi_{\text{vx}}(y; m^*); \mathbf{v})}{\mathrm{d}m^*} \,\mathrm{d}y \\
&= \int_{a(m^*)}^{\infty} \Gamma(\xi_{\text{vx}}(y; m^*); \mathbf{v})^{n-1} \frac{\mathrm{d}\Gamma(\xi_{\text{vx}}(y; m^*); \mathbf{v})}{\mathrm{d}m^*} \,\mathrm{d}y \\
&= \int_{a(m^*)}^{m^*} \Gamma(\xi_{\text{vx}}(y; m^*); \mathbf{v})^{n-1} \frac{\mathrm{d}\Gamma(\xi_{\text{vx}}(y; m^*); \mathbf{v})}{\mathrm{d}m^*} \,\mathrm{d}y \\
&\quad + \int_{m^*}^{\infty} \Gamma(\xi_{\text{vx}}(y; m^*); \mathbf{v})^{n-1} \frac{\mathrm{d}\Gamma(\xi_{\text{vx}}(y; m^*); \mathbf{v})}{\mathrm{d}m^*} \,\mathrm{d}y \\
&\geqslant \int_{a(m^*)}^{m^*} \Gamma(\xi_{\text{vx}}(m^*; m^*); \mathbf{v})^{n-1} \frac{\mathrm{d}\Gamma(\xi_{\text{vx}}(y; m^*); \mathbf{v})}{\mathrm{d}m^*} \,\mathrm{d}y \\
&\quad + \int_{m^*}^{\infty} \Gamma(\xi_{\text{vx}}(m^*; m^*); \mathbf{v})^{n-1} \frac{\mathrm{d}\Gamma(\xi_{\text{vx}}(y; m^*); \mathbf{v})}{\mathrm{d}m^*} \,\mathrm{d}y \\
&= \Gamma(\xi_{\text{vx}}(m^*; m^*); \mathbf{v})^{n-1} \int \frac{\mathrm{d}\Gamma(\xi_{\text{vx}}(y; m^*); \mathbf{v})}{\mathrm{d}m^*} \,\mathrm{d}y = 0,
\end{aligned}$$

which is equivalent to $\overline{H}'(m^*) \leqslant 0$.

□

E. Equilibrium Existence

The analysis in Section 3 establishes equilibrium existence for the case where there is no inflexion point (ie, c is either globally concave or globally convex). In addition, the argument in Lemma 3.1 can be used to obtain equilibrium existence for the convex-concave and concave-concave cases. In this appendix, we focus on the case where there are at least two inflexion points. Let K denote the number of inflexion points and use x_1^ι (x_K^ι) to denote the first (last) inflexion point. We define $r := \inf_{x \in [x_1^\iota, x_K^\iota]} c'(x) > 0$.

Lemma E.1. Let $\overline{y}$ denote the upper bound of $\operatorname{supp}(G^*)$. Then, $\overline{y} \leqslant y_m := \max\{c^{-1}(1) + \frac{1}{r}, nx_{\text{d}}\}$.

Proof. It suffices to consider the case where $\operatorname{supp}(G^*)$ extends beyond $c^{-1}(1)$. We know that the support of F^* cannot extend beyond $c^{-1}(1)$; therefore, $\xi^*(y)$ is affine for $y \geqslant c^{-1}(1)$. This last affine part of ξ^* is tangent to c at some $\overline{m} \leqslant c^{-1}(1)$, which is the last mass point of F^* (for, if not, then ξ^* can be improved in [4.4]). Now, there are several possibilities.

(i) The last mass point is also the first, and only, mass point of F^*. In this case, we either have the deterministic effort equilibrium and $\overline{y} = nx_{\text{d}}$, or $\overline{m} > x_1^\iota$. In the latter

case, there are two possibilities: (a) $\overline{m} < x_K^\iota$, which implies $c'(\overline{m}) \geqslant r$; and (b) $\overline{m} > x_K^\iota$, which implies $c'(\overline{m}) \geqslant r$ as well because in this case $c(x)$ is convex for $x \geqslant x_K^\iota$.
(ii) There is more than one mass point. In this case, $\overline{m} > x_1^\iota$ and either (a) or (b) holds as above, and hence $c'(\overline{m}) \geqslant r$ again holds.

We know that in the symmetric equilibrium $\Phi(G^*(y); \mathbf{v}) = \min\{\tilde{\xi}^*(y), v_1\}$; therefore, $\tilde{\xi}^*(\overline{y}) = v_1$. This equation takes the form

$$c(\overline{m}) + c'(\overline{m})(\overline{y} - \overline{m}) - \xi^*(0) = v_1,$$

which gives

$$\overline{y} = \overline{m} + \frac{v_1 + \xi^*(0) - c(\overline{m})}{c'(\overline{m})} < c^{-1}(1) + \frac{1}{r}.$$

We conclude that $\overline{y}$ is bounded above by $y_m = \max\{c^{-1}(1) + \frac{1}{r}, nx_{\mathrm{d}}\}$. □

We make use of the following result:

Theorem E.1. *Suppose that a game is compact, convex, quasi-symmetric, and diagonally quasiconcave. If the game has the local better-reply property on the diagonal, then it has a symmetric pure strategy Nash equilibrium.*

This is Theorem 1 in Baye, Tan, and Zhou (1993), reported as Theorem 4 in Reny (2020, p 453); we use the terminology from the latter. For a cdf G with support in $\mathbb{R}_+$, define a cost function

$$C(G) := \min_{F \in \mathrm{MPC}(G)} \int c \, \mathrm{d}F.$$

We consider a game where each player i chooses a cdf of output $G_i \in S_i = \{G \in [0,1]^{\mathbb{R}} : G \text{ is a cdf with } G(0-) = 0 \text{ and } G(y_m) = 1\}$, where y_m is defined in Lemma E.1.[26] The set S_i is a metric space with the L_1 metric, which metrises the topology of weak convergence (ie, the topology of convergence in distribution) on S_i (cf Machina, 1982). We observe that with this metric, (i) S_i is a compact metric space (cf Aliprantis and Border, 1999, Theorem 14.11, p. 482), and (ii) MPC(G) is a compact subset of S_i for any $G \in S_i$ (see Kleiner, Moldovanu, and Strack, 2021). The latter fact is immediate because MPC(G) is defined by linear inequalities, and hence is a closed subset of S_i. It is straightforward to see that S_i is also convex.

The payoffs in this game are

$$\pi_i(G_i, \mathbf{G}_{-i}) = \int \Psi(\mathbf{G}_{-i}(y); \mathbf{v}) \, \mathrm{d}G_i(y) - C(G_i).$$

[26]Since Lemma E.1 shows that $\mathrm{supp}(G^*)$ is bounded by y_m in *any* symmetric equilibrium, restricting the strategy spaces to S_i is without loss.

Here, $\mathbf{G}_{-i}$ denotes the vector of other players' strategies; and $\Psi(\mathbf{G}_{-i}(y);\mathbf{v})$ is the expected winnings of player i conditional on player i's output being y. With $\mathbf{G}_{-i} = (\hat{G},\ldots,\hat{G})$, the "symmetrised" payoffs in our game are given by

$$\Pi(G,\hat{G}) := \int \Phi(\hat{G}(y);\mathbf{v})\,\mathrm{d}G(y) - C(G). \tag{E.1}$$

Following Reny (2020, Theorem 4, p453), our contest is *compact* because the strategy spaces S_i are compact; it is convex because the strategy spaces S_i are convex (and locally convex) sets; it is quasi-symmetric because our contest is symmetric, as evinced from [E.1]; it is diagonally quasiconcave because the mapping $G \mapsto \Pi(G,\hat{G})$ is concave (see Lemma E.2). Our contest has the *local better-reply property* on the diagonal if for any $\bar{G}$ that is not a symmetric Nash equilibrium, there exists a $\delta > 0$ and a $G^\dagger$ such that

$$\left\|\hat{G} - \bar{G}\right\|_1 < \delta \qquad \text{implies} \qquad \Pi(G^\dagger,\hat{G}) > \Pi(\hat{G},\hat{G}).$$

Theorem E.2. *The game above has a pure strategy symmetric Nash equilibrium.*

Proof. By definition, our game is symmetric, and hence is quasi-symmetric. The strategy space is compact and convex, and the game is diagonally quasiconcave (Lemma E.2). Finally, it has the local better-reply on the diagonal property (Lemma E.6). Thus, all the conditions of Theorem E.1 above are met, and the existence of a pure strategy symmetric Nash equilibrium is established. □

The proofs establishing the various properties of the payoff functions described above are in Appendix E.1, with some auxiliary results relegated to Appendix E.2.

E.1. Lemmas Supporting Theorem E.2

Lemma E.2. The function $\Pi(G,\hat{G})$ defined in [E.1] is concave in G.

Proof. It suffices to show that $C(G)$ is convex in G. Towards this end, let F^t denote a solution to $\min \int c\,\mathrm{d}F$ s.t. $F \in \mathrm{MPC}(tG^1 + (1-t)G^0)$ for all $t \in [0,1]$. Then,

$$\begin{aligned} C(tG^1 + (1-t)G^0) &= \int c\,\mathrm{d}F^t \\ &\leqslant \int c\,\mathrm{d}(tF^1 + (1-t)F^0) = t\int c\,\mathrm{d}F^1 + (1-t)\int c\,\mathrm{d}F^0 \\ &= tC(G^1) + (1-t)C(G^0). \end{aligned}$$

The inequality holds because $F^1 \in \text{MPC}(G^1)$ and $F^0 \in \text{MPC}(G^0)$ implies $tF^1 + (1-t)F^0 \in \text{MPC}(tG^1 + (1-t)G^0)$. $\square$

Lemma E.3. The function $\Pi(G, G) = \int \Phi(G(y); \mathbf{v})\, \mathrm{d}G(y) - C(G)$ is continuous in G. In particular, the mapping $G \mapsto C(G)$ is continuous.

Proof. Note that $\Pi(G, G) = \frac{1}{n} - C(G)$, and hence it is sufficient to establish that $C(G)$ is continuous. Consider subset $S_0 := \{G \in S : G(0) = 0 \text{ and } G \text{ has full support on } [0, y_m]\}$. Then, S_0 is dense in S.

For any $G \in S_0$, Theorem 2 of Dworczak and Martini (2019) implies that

$$C(G) = \min_{F \in \text{MPC}(G)} \int c\, \mathrm{d}F = \max_{\xi \leqslant c,\, \xi \text{ concave}} \int \xi\, \mathrm{d}G$$

We claim that any optimal ξ is Lipschitz, with $\text{Lip}(\xi) \leqslant \text{Lip}(c)$. To see this, notice that by Proposition 2 of Dworczak and Martini (2019), the interval $[0, y_m]$ can be partitioned into a finite number of subintervals in which ξ is either affine and supports c or strictly concave and coincides with c. If $\xi(0) = c(0) = 0$, it must be that $\xi'(0) \leqslant c'(0)$ because otherwise the constraint $\xi \leqslant c$ would be violated at some $\varepsilon > 0$. If $\xi(0) < 0$ then the initial segment of ξ is affine and $\xi(y) < c(y)$ for $y \in [0, b)$, where $b = \min\{y : \xi(y) = c(y)\}$ is the point where ξ and c first meet (if there is no such point, there is a feasible improvement of ξ). In this case it must be that $\xi'(0) = c'(b)$. Indeed, if $\xi'(0) < c'(b)$ then the constraint $\xi \leqslant c$ would be violated at some $b - \varepsilon$; if $\xi'(0) > c'(b)$ then a strict improvement obtains by increasing $\xi(0)$ while keeping $\xi(b)$ fixed. (In particular, raise $\xi(0)$ until the affine segment connecting $(0, \xi(0))$ and $(b, \xi(b))$ is tangent to the graph of c or $\xi(0) = 0$, whichever happens first.) Clearly, $\xi'(0) \leqslant \text{Lip}(c)$ holds in all cases, and hence $\text{Lip}(\xi) \leqslant \text{Lip}(c)$ due to the concavity of ξ.

Recall also that in any symmetric equilibrium $-\xi(0)$ is the players' rent, and hence we cannot have $\xi(0) < -1$. Define the set $\Xi_c := \{\xi \in \mathbb{R}^{[0, y_m]} : -1 \leqslant \xi \leqslant c,\ \text{Lip}(\xi) \leqslant \text{Lip}(c),\ \xi \text{ concave, increasing}\}$. It is clear from the above discussion that for all $G \in S_0$,

$$C(G) = \max_{\xi \leqslant c,\, \xi \text{ concave}} \int \xi\, \mathrm{d}G = \max_{\xi \in \Xi_c} \int \xi\, \mathrm{d}G$$

For any $G_1, G_2 \in S_0$ and $\xi \in \Xi_c$, observe that $\int \xi\, \mathrm{d}(G_1 - G_2) = \int (G_2 - G_1)\, \mathrm{d}\xi$. Following the proof of Theorem 2 in Milgrom and Segal (2002), we note that $|C(G_1) - C(G_2)| \leqslant \max_{\xi \in \Xi_c} \left|\int (G_1 - G_2)\, \mathrm{d}\xi\right| \leqslant \text{Lip}(c)\, \|G_1 - G_2\|_1$. That is, C is Lipschitz on S_0.

By the Continuous Extension Theorem (Carothers, 2000, Theorem 8.16, p. 119) it follows that C has a unique uniformly continuous extension to S, which establishes the claim. $\square$

Lemma E.4. Let $G^\dagger$ be continuous. Then, the mapping $G \mapsto \Pi(G^\dagger, G)$ is continuous in G.

Proof. We have $\Pi(G^\dagger, G) = \int_0^{y_m} \Phi(G(y), \mathbf{v})\, \mathrm{d}G^\dagger(y) - C(G^\dagger)$; thus, it is sufficient to show that $\int_0^{y_m} \Phi(G(y), \mathbf{v})\, \mathrm{d}G^\dagger(y)$ is continuous in G. Integrating by parts, we see that $\int_0^{y_m} \Phi(G(y), \mathbf{v})\, \mathrm{d}G^\dagger(y) = \Phi(G(y_m), \mathbf{v}) \cdot 1 - \int_0^{y_m} G^\dagger(y)\, \mathrm{d}\Phi(G(y), \mathbf{v})$. Consider the mapping $T : S \to S$, where $TG(y) := \Phi(G(y), \mathbf{v})$. We claim that the mapping T is Lipschitz. To see this, recall that $q \mapsto \Phi(q, \mathbf{v})$ is Lipschitz of some rank A. Thus,

$$\begin{aligned}
\left\|TG - T\tilde{G}\right\|_1 &= \int_0^{y_m} \left|TG(y) - T\tilde{G}(y)\right| \mathrm{d}y \\
&= \int_0^{y_m} \left|\Phi(G(y), \mathbf{v}) - \Phi(\tilde{G}(y), \mathbf{v})\right| \mathrm{d}y \\
&\leqslant \int_0^{y_m} A \left|G(y) - \tilde{G}(y)\right| \mathrm{d}y \\
&\leqslant A \cdot \left\|G - \tilde{G}\right\|_1 \cdot y_m
\end{aligned}$$

which proves the claim.

Let (G_n) be a sequence in S that converges to G (in the L_1 metric), which implies that $TG_n \to TG$ in the L_1 metric (because T is Lipschitz). Because each $G \in S$ is the cdf of a measure μ_G on $[0, y_m]$, it follows that $\mu_{TG_n} \to \mu_{TG}$ in the topology of weak convergence (recall that $\mu_{G_n} \to_{w*} \mu_G$ if and only if $G_n \to G$ in the L_1 metric). Thus,

$$\begin{aligned}
\left|\int G^\dagger(y)[\mathrm{d}\Phi(G_n(y), \mathbf{v}) - \mathrm{d}\Phi(G(y), \mathbf{v})]\right| &= \left|\int G^\dagger(y)[\mathrm{d}\mu_{TG_n}(y) - \mathrm{d}\mu_{TG}(y)]\right| \\
&\to 0
\end{aligned}$$

because $G^\dagger$ is continuous and $\mu_{TG_n} \to_{w*} \mu_{TG}$. □

Lemma E.5. Suppose $\bar{G}$ is not a symmetric Nash equilibrium. Then, there exists a continuous $G^\dagger \in S$ such that $\Pi(G^\dagger, \bar{G}) > \Pi(\bar{G}, \bar{G})$.

Proof. Let $\psi(G) := \Pi(G, \bar{G})$. Suppose $\bar{G}$ is not a symmetric Nash equilibrium. Then there exists a $G \in S$ such that $\psi(G) - \psi(\bar{G}) = 4\varepsilon > 0$. By Lemma E.8, there exists a $G_0 \in S$ that does not have any points of discontinuity in common with $\bar{G}$ such that $\psi(G_0) - \psi(G) > -\varepsilon$. Lemma E.9 ensures the existence of a $G_1 \in S$ such that $\psi(G_1) - \psi(G_0) > -2\varepsilon$. Putting these together, we find that $\psi(G_1) - \psi(\bar{G}) = [\psi(G_1) - \psi(G_0)] + [\psi(G_0) - \psi(G)] + [\psi(G) - \psi(\bar{G})] > \varepsilon$. Setting $G^\dagger = G_1$ establishes the claim. □

Lemma E.6. Our contest has the *local better-reply property* on the diagonal.

Proof. Suppose $\bar{G}$ is not a symmetric Nash equilibrium. Then, from Lemma E.5 there exists a continuous $G^\dagger$ and an $\varepsilon > 0$ such that

$$\Pi(G^\dagger, \bar{G}) - \Pi(\bar{G}, \bar{G}) = 2\varepsilon > 0 \tag{E.2}$$

Define the functions $\varphi(G) := \Pi(G, G)$ and $\psi(G) := \Pi(G^\dagger, G)$. The continuity of φ is established in Lemma E.3, and the continuity of ψ is shown in Lemma E.4. Moreover, $\psi(\bar{G}) - \varphi(\bar{G}) = 3\varepsilon > 0$. The continuity of φ implies that there exists δ_φ such that $\|G - \bar{G}\|_1 < \delta_\varphi$ implies $|\varphi(G) - \varphi(\bar{G})| < \varepsilon$. Similarly, the continuity of ψ ensures the existence of δ_ψ such that $\|G - \bar{G}\|_1 < \delta_\psi$ implies $|\psi(G) - \psi(\bar{G})| < \varepsilon$. Let $\delta = \min\{\delta_\psi, \delta_\varphi\}$.

Then, for any G such that $\|G - \bar{G}\| < \delta$, we have $\psi(G) > \psi(\bar{G}) - \varepsilon = \varphi(\bar{G}) + \varepsilon > \varphi(G)$, where the second inequality is because $\psi(\bar{G}) - \varphi(\bar{G}) = 2\varepsilon$. In other words, $\Pi(G^\dagger, G) > \Pi(G, G)$ for any G such that $\|G - \bar{G}\| < \delta$, as required. □

E.2. Results on Approximating Functions

Lemma E.7. Let μ be a regular finite measure on $[0, y_m]$, and G an increasing function. For any $\delta, \varepsilon > 0$, there exists $H \in \mathbb{R}^{[0,y_m]}$ that is continuous, increasing, and satisfies (i) $\|G - H\|_1 < \delta$ and (ii) $\int_0^{y_m} |H(y) - G(y)| \, \mathrm{d}\mu(y) < \varepsilon$.

Proof. [27] Given $\varepsilon, \delta > 0$, there exists a finite partition $(A_j)_{j=1}^n$ of $[0, y_m]$ such that (i) $\mathrm{Leb}(A_j) < \delta$ and $\mu(\mathrm{int}\, A_j) < \varepsilon$ for all $j \leqslant n$ (where $\mathrm{Leb}(A)$ is the Lebesgue measure of A and $\mathrm{int}\, A_j = (x_j, x_{j+1})$ is the interior of the interval A_j), and (ii) $A_j = [x_j, x_{j+1})$ for all $j < n$ and $A_n = [x_n, x_{n+1} = y_m]$. In what follows, $\partial A_j = \{x_j, x_{j+1}\}$ will represent the boundary of the set of A_j. Notice that all mass points of μ with mass greater than ε (of which there are only finitely many) must necessarily be at the boundary of some A_j.

Define the function H piecewise: for $x \in A_j$, let $H(x) = G(x_j) + [G(x_{j+1}) - G(x_j)](x - x_j)/(x_{j+1} - x_j)$. Because G is increasing, it follows that H is also increasing. Moreover, H is continuous and $H(x_j) = G(x_j)$ for all $j = 1, \ldots, n+1$. Notice also that (a) for all $x \in A_j$, $|H(x) - G(x)| \leqslant |G(x_{j+1}) - G(x_j)|$, and (b) $H(x) = G(x)$ for

[27] This argument is based on a proof by PhilippeC posted on StackExchange.

$x \in \partial A_j$ for all j. This gives us the bound

$$\begin{aligned}&\int |H(x) - G(x)| \,\mathrm{d}\mu(x)\\ &\quad= \sum_{j=1}^{n}\left[\int_{\text{int}\, A_j} |H(x) - G(x)| \,\mathrm{d}\mu(x) + \overbrace{\int_{\partial A_j} |H(x) - G(x)| \,\mathrm{d}\mu(x)}^{=0}\right]\\ &\quad\leqslant \sum_{j=1}^{n}[G(x_{j+1}) - G(x_j)]\,\mu(\text{int}\, A_j) < \varepsilon\end{aligned}$$

because $\sum_{j=1}^{n}[G(x_{j+1}) - G(x_j)] = 1$. Replacing the measure μ with Lebesgue measure, and recalling that $\mathrm{Leb}(A_j) < \delta$ for all $j = 1, \ldots, n$, it follows that $\|H - G\|_1 < \delta$, as claimed. □

Lemma E.8. Let $\bar{G}, G \in S$ and $\varepsilon > 0$. Then, there exist $\delta > 0$ and $G_0 \in S$ with finite support such that $\|G - G_0\|_1 < \delta$ and $\Pi(G_0, \bar{G}) - \Pi(G, \bar{G}) > -\varepsilon$.

Proof. Notice that the mapping $G \mapsto C(G)$ is uniformly continuous, so there exists a δ (independent of G) such that $\|G_0 - G\|_1 < \delta$ implies $|C(G) - C(G_0)| < \varepsilon$. That G can be approximated from below by a simple distribution G_0 is shown, for example, by Aliprantis and Border (1999, Lemma 11.13, p. 403) or Pollard (2002, Lemma 11, p. 25). In particular, this G_0 can be chosen such that $\|G_0 - G\|_1 < \delta$. Moreover, by construction, we have $G_0 \leqslant G$. Thus, we find that

$$\begin{aligned}\Pi(G_0, \bar{G}) - \Pi(G, \bar{G}) &= \underbrace{\int \Phi(\bar{G}(y), \mathbf{v})[\mathrm{d}G_0 - \mathrm{d}G]}_{\geqslant 0} - \underbrace{\big(C(G_0) - C(G)\big)}_{<\varepsilon}\\ &> -\varepsilon\end{aligned}$$

where the integral inequality is because G_0 first-order stochastically dominates G, and we have used the continuity of C and the fact that $\|G_0 - G\|_1 < \delta$. □

Lemma E.9. Let $\bar{G}, G_0 \in S$ and $\varepsilon > 0$ such that $\bar{G}$ and G_0 have no common points of discontinuity. Then, there exists $\delta > 0$ and a continuous $G_1 \in S$ such that $\|G_1 - G_0\|_1 < \delta$ and $\Pi(G_1, \bar{G}) - \Pi(G_0, \bar{G}) > -2\varepsilon$.

Proof. The mapping $G \mapsto C(G)$ being uniformly continuous, there exists a δ (independent of G_0) such that $\left\|G_0 - \hat{G}\right\|_1 < \delta$ implies $\left|C(\hat{G} - C(G_0)\right| < \varepsilon$ for any $\hat{G} \in S$. By Lemma E.7, there exists a $G_1 \in S$ such that $\|G_0 - G_1\|_1 < \delta$ and $\int |G_1 - G_0| \,\mathrm{d}\Phi(\bar{G}, \mathbf{v}) < \varepsilon$. Because G_0 and $\bar{G}$ do not have any common points of discontinuity, we may integrate by parts (Billingsley, 2012, Theorem 18.4, p251) to obtain $\int \Phi(\bar{G}, \mathbf{v})\,\mathrm{d}[G_1 - G_0] =$

$-\int[G_1 - G_0]\,\mathrm{d}\Phi(\bar{G}, \mathbf{v})$. Thus,

$$\begin{aligned}\Pi(G_1, \bar{G}) - \Pi(G_0, \bar{G}) &= \int \Phi(\bar{G}(y), \mathbf{v})\,\mathrm{d}[G_1 - G_0] - \big(C(G_1) - C(G_0)\big) \\ &= \underbrace{\int [G_0 - G_1]\,\mathrm{d}\Phi(\bar{G}(y), \mathbf{v})}_{>-\varepsilon} - \underbrace{\big(C(G_1) - C(G_0)\big)}_{<\varepsilon} \\ &> -2\varepsilon\end{aligned}$$

as claimed. □