Performance Responses to Competition Across Skill-Levels in Rank Order Tournaments: Field Evidence and Implications for Tournament Design

HARVARD | BUSINESS | SCHOOL

Performance Responses to Competition Across Skill-Levels in Rank Order Tournaments: Field Evidence and Implications for Tournament Design

Kevin J. Boudreau **Constance E. Helfat**
Karim R. Lakhani **Michael Menietti**

Working Paper

14-014

August 8, 2013

PERFORMANCE RESPONSES TO COMPETITION ACROSS SKILL-LEVELS IN RANK ORDER TOURNAMENTS: FIELD EVIDENCE AND IMPLICATIONS FOR TOURNAMENT DESIGN

Kevin J. Boudreau, Constance E. Helfat, Karim R. Lakhani, and Michael Menietti*

Abstract

Tournaments are widely used in the economy to organize production and innovation. We study individual contestant-level data on 2796 contestants in 774 software algorithm design contests with random assignment. Precisely conforming to theory predictions, the performance response to added contestants varies non-monotonically across contestants of different abilities; most respond negatively to competition; highest-skilled contestants respond positively. In counterfactual simulations, we interpret a number of tournament design policies (number of competitors, prize allocation and structure, divisionalization, open entry) as a means of reconciling non-monotonic incentive responses to competition, effectively manipulating the number and skills distribution of contestants facing one another.

JEL Codes: D02, J4, L2, M5

* Boudreau: London Business School, Strategy Department (email: kboudreau@london.edu); Helfat: Dartmouth University, Strategy Department (email: constance.e.helfat@tuck.dartmouth.edu); Lakhani: Harvard Business School: Department of Technology and Operations Management (email: klakhani@hbs.edu); Menietti: Harvard-NASA Tournament Laboratory (email: mmenietti@fas.harvard.edu). We are grateful to members of the TopCoder executive team for considerable attention, support, and resources in the carrying out of this project, including Jack Hughes, Rob Hughes, Mike Lydon, and Ira Heffan. For helpful comments, we thank seminar participants at Duke University, Georgia Tech and London Business School. The authors would like to acknowledge financial support from London Business School Research and Materials Development Grant and the NASA-Tournament Laboratory. All errors are our own.

1 Introduction

Tournaments and other rank-order incentive mechanisms have been used to model a wide range of settings: executive placement, elections, research and development and innovation contests, sports tournaments, and variable sales compensation–situations in which placing at the top of the performance rank-order leads to out-sized payoffs. Tournaments and contests have a long history as a means of achieving technological advances in industry (Brunt et al., 2011) and recently have been witnessed in conspicuous cases such as the X-prize for private space flight, DARPA challenges to develop automomonous vehicle technologies, and the Netflix contest to improve the company's algorithm to match users with preferred movies (Murray et al., 2012). Also in recent times, companies such as TopCoder and Innocentive have established fixed platforms and sets of regular contestants as "members" of those platforms to make it possible to carry out a regular and on-going stream of contests. Further, the US government recently passed legislation giving prize-based procurement authority to all federal agencies (Bershteyn and Roekel, 2011). Thus, rank order tournaments play an important role in organizing production, efforts and creative activity in the economy (Lazear and Rosen, 1981; Wright, 1983; Kremer and Williams 2010).

A number of core design questions around contests have been examined in the theoretical literature, including when they are efficient relative to alternative incentive schemes (e.g., Lazear and Rosen, 1981), and questions around the number and abilities of contestants (e.g., Fullerton and McAfee, 1999) and prize size and structure (e.g., Moldovanu and Sela, 2001, 2006).[1] The empirical literature examining these core questions of design remains somewhat less developed in large part because of data limitations: theoretical models typically make econometric demands that are rarely satisfied by existing data sources and it

[1]Szymanski (2003) evokes the core issues of contest and tournament design with these vivid examples from sports: "What is the optimal number of entrants in a race, or the optimal number of teams in a basketball league? What is the optimal structure of prizes for a golf tournament, or degree of revenue sharing for a football championship? How evenly balanced should the competing teams be in the NASCAR or Formula One championships? What is the maximum number of entrants per nation to the Olympic Games that should be permitted? What quota of qualifying teams to the soccer World Cup should be allocated to the developing nations?"

is in the relatively rare instances in which a regular stream of repeated contests have been carried out (rather than ad hoc events) where meaningful econometric comparisons might best be derived. Empirical work to date has nonetheless made considerable progress in establishing cornerstone findings such as higher prizes lead to higher performance (Ehrenberg and Bognanno, 1990a,b; Orszag, 1994) and competing with markedly superior opponents or "super-stars" decreases performance (Brown, 2011; Tanaka and Ishino, 2012). A range of experimental studies also provides evidence that our theoretical characterizations of strategic interactions in tournaments as incentive-providing mechanisms are to a considerable degree borne out under laboratory conditions (Dechenaux, et al., 2012).[2]

Among the most basic and important questions that have been addressed are those that concern how large and competitive a contest should be. How many contestants should be allowed, enabled, or facilitated to enter? The theoretical literature on innovation contests generally points to smaller contests as producing higher incentives,[3] where even just two contestants have been argued to produce the highest incentives (Fullerton and McAfee, 1999; Che and Gale, 2003). Absent any form of competition, contestants will have little incentive to exert effort to improve their work, but, beyond a minimum level of competition, the marginal return to added effort may diminish with a lower chance of winning. The broader theoretical literature on contests and tournaments has considered how the related issue of composition of contestants impacts contest performance. Roughly speaking, this research has shown that an increase in homogeneity among contestants increases aggregate effort (Konrad, 2009). Moldovanu and Sela (2006) establish a number of results on the preferred structure of competition for designers interested in maximizing aggregate effort or simply the highest effort. Within their model, they establish that if contestant costs are convex the optimal design depends on the particular cost function and distribution of abilities. Hence,

[2]A range of extensions beyond core questions of design have also been studied with both theory and experimental results, including the design of multi-stage tournaments (Fu and Lu 2009), implications of sabotage and "office politics" among contestants (e.g., Carpenter, et al. 2010) and implications of self-selection into open tournaments (Dohmen and Falk, 2011).

[3]See, for example, Glazer and Hassin, 1988; Lazear and Rosen, 1981; Taylor, 1995; Fullerton and McAfee, 1999; Che and Gale, 2003; Terwiesch and Xu, 2008.

optimal design is a "hard" problem in that no solution works over all environments, but the particular context needs to be considered.

Several field studies make important progress towards testing the empirical relationship between numbers of contestants and performance outcomes and have generally found a negative aggregate or average relationship at the contest level in areas such as sales compensation (Casas-Arce and Martínez-Jerez 2009), test-taking (Garcia and Tor 2009) and software algorithm development[4] (Boudreau et al., 2011).We have yet to observe fine-grained individual causal responses to better discern more nuanced patterns. Casual observation of contests and tournaments in the economy, however, readily reveals contests ranging from just a few to many (dozens or even hundreds) of contestants freely entering–and thus the possibility that a more nuanced view of incentive responses to competitive interactions may be warranted.

In this article, we clarify theoretical arguments for non-monotonic incentive and performance responses to competition across contestants of different skills or ability levels. To see the intuition of the model and arguments, it is useful to begin with the longstanding intuition of why two contestants in a tournament are better than one for producing high-quality outcomes. In winner-takes-all contests with only one participant, contestants will have little incentive to exert effort to improve their work because there are no parties against whom they will be evaluated. Thus, by adding some minimum level of competition and rivalry, probability of loss that can be lessened should lead to greater effort (Harris and Vickers, 1987). This is caused by effort-inducing *rivalry* or *racing.* While adding contestants beyond this point can dilute incentives by making tournaments less likely to win, following usual arguments, we clarify that for the strongest contestants, adding more contestants can produce effort-inducing rivalry. In contests with contestants of heterogeneous skills, the strongest contestants may gradually sense stimulating rivalry only with higher levels of competition. We illustrate the argument by building on the theoretical framework of Moldovanu and Sela (2001), which features a one-shot tournament with multiple prizes, contestants of

[4]Here we analyze similar data from the same empirical context, but studying variation in individual contestants' performance.

heterogeneous abilities, and flexible cost functions–a simple and basic set of features that are relatively general and common to real tournaments in a range of contexts including that studied here. The model's precise prediction is that there should be a sort of asymmetric, U-shaped incentive response to competition across the skills level distribution, with the lowest-skilled contestants negligibly responding to competition, the bulk of contestants at intermediate levels experiencing a negative response to competition, and the highest-ability contestants experiencing a more positive (less negative) response to competition. Where the earlier-mentioned stimulating effect of rivalry outweighs the profit- and effort-quashing effects of competition, added competition will in fact stimulate effort and performance.

Our main contribution is to estimate relationships between performance in these contests and competition levels across the full distribution of skill levels. We do this by studying data on software algorithm programming contests by TopCoder, a context in which fine-grained data are available on contestant ability levels and performance over a large number of comparable contests and where natural experimental variation created by random assignment can be exploited. We study data on 774 cash prize-based contests between 2005 and 2007, in which varying numbers of randomly-assigned individuals (roughly numbering between 15 and 20) competed to solve software algorithm problems over a 75 minute period; skill levels ranged widely, but skills distributions in each room were roughly constant. Our core analysis consists of precisely estimating the causal response to varying numbers of contestants across the skills distribution using an unconstrained, flexible nonparametric procedure. We find the specific single-valleyed non-monotonic relationship predicted by theory. We proceed to then estimate a structural model to consider counterfactual experiments and to more deeply interpret design policies of these contests. We find that a range of key contest design policies in this case (capping contest size at 20, allocating a fixed prize pool, creating two prizes per independent prize room, creating separate divisions and allowing open entry) serve to reconcile the non-monotonic incentive responses to competition documented here, serving to manipulate both the number and skills distribution of competitors who faces one another.

This article therefore most directly builds on the stream of field studies testing propositions of theoretical models of tournaments and particularly those examining effects of varying levels of competition in contexts of production contexts (Casas-Arce and Martínez-Jerez 2009; Garcia and Tor 2009; Brown, 2011; Boudreau et al., 2011; Brunt et al., 2011; Tanaka and Ishino, 2012). To this growing body of work we add fine-grained causal evidence demonstrating a non-monotonic response to competition across contestants at different skill levels, while also offering evidence outside the context of sports evidence. The theory and empirical findings presented here clarify that adding competition can stimulate effort and performance among highest-skilled contestants while depressing effort and performance over the larger bulk of contestants at lower skill levels. Where a contest organizer wishes to maximize performance and engagement across a wider cross-section of contestants (e.g., a sales contest, sports tournaments, executive compensation, contests geared to promoting many solutions, contests geared to promoting learning or engagement, etc.), the results illustrate that the non-monotonic responses to competition may create tradeoffs whereby nuanced approaches to managing numbers and skills distributions of contestants may be beneficial (i.e., the capping of entry, creating of divisions, etc.) The results contrast those of models with homogenous contestants in particular. These results also provide an explanation based on the strategic incentives for why contest organizers who are most interested in maximum outcomes, often choose to design and implement large "grand challenges" with open entry to large number of contestants despite the potential incentive-quashing effects of high levels of competition on many contestants. In this sense, these results clarify theories based on strategic incentives for holding large contests, complementing theories of large contests based on an interest of promoting large-scale "parallel" experimentation across many distinct technical approaches to a given problem by deploying large numbers of heterogeneous solvers (e.g., Terwiesch and Xu, 2008; Boudreau, et al., 2011). Likewise, a contest design might also consider these non-monotonic responses across skill levels to determine more nuanced attempts of minimizing costly effort while achieving some performance goal, particularly when efforts are costly or even wasteful (Tullock, 1980). It should also be noted that within these fine-grained data, our results illustrate the strong predictive power of economic models of strategic interactions and incentives, and particularly that of the framework developed by Moldovanu and Sela (2001). This possibility of a stimulating effect of competition on strategic investment incentives is analogous to, although based on distinct mechanisms and within a different institutional context, findings of the potentially stimulating effects of competition in dynamic industrial market competition on a vertical innovation quality ladder (Harris and Vickers, 1987; Aghion et al., 2005) and in models of patent races (Reinganum, 1989).

The article proceeds as follows. In Section 2 we develop our theory and empirical predictions. Section 3 describes the empirical context and data set. Section 4 presents results. Section 5 concludes.

2 Theory and Hypotheses Development

Anticipating key features of our empirical context, here we develop hypotheses of effects of competition on incentives and performance in a contest in relation to a one-shot tournament, with multiple prizes, with contestants of heterogeneous abilities. We build on a simple and tractable analytical framework fitting this description developed by Moldovanu and Sela (2001).[5]

Consider n-contestants competing in a simultaneous tournament for $p < n$ prizes. Prizes are strictly decreasing in value $V_1 > V_2 > \ldots > V_p$. Each player draws on an "ability" or skill level from zero up to some highest possible level, $a_i \in [0, m]$. Let skill be bounded at some $m < 1$. Ability is distributed randomly according to some cumulative distribution F_A. The distribution has continuous support and is differentiable; the density function is denoted f_A. Assume the distribution F_A is common knowledge, whereas a player's own ability is private information. Ability determines a player's marginal cost of submitting a solution of

[5] Moldovanu and Sela's further work in Moldovanu and Sela (2006) somewhat overlaps with the results here. They investigate a broader tournament framework allowing for two-stage elimination tournaments and consider the optimality of many aspects of design.

incrementally higher quality.

The quality of a player's solution submission or performance "score" is determined by the player's ability and his choice of effort level. Rather than consider the effort choice, it is useful to simply consider a player's choice of quality directly (based on both privately known ability and the choice of effort). This choice of solution quality is effectively a player's chosen "bid" in the contest. Henceforth, we refer to "bid" and the expected solution quality interchangeably as is customary.

Player i chooses a costly bid quality level, $b_i \in \Re_+$. The cost of bidding, $c(b_i, a_i)$, is increasing in the size of bid according to some function $\gamma(b_i)$, where $\gamma' > 0$ and a multiplier $(1 - a_i)$ associated with the ability of the player or $c(b_i, a_i) = (1 - a_i)\gamma(b_i)$. Higher skilled players have lower costs of supplying higher quality bids. Where players are risk-neutral and $r_i(b_i)$ is the rank of a player's bid b_i, the expected payoff to player i is as in the following expression: $\pi_i(b_i; a_i) = \sum_{j=1}^{p} Pr\left\{r_i(b_i) = j\right\} V_j - (1 - a_i)\gamma(b_i)$. This simple characterization of the contest implies an expected payoff that is simply the sum of prize values at different ranks, weighted by the probability of a bid placing at these ranks, less the cost of developing a bid of that quality level.

■ **Equilibrium.** Moldovanu and Sela (2001) find the symmetric equilibrium mapping abilities to bid quality levels $b : [0, m] \to \Re_+$. A symmetric, strictly increasing bid function is assumed to exist, allowing the probability term in the expected payoffs to be substituted with a probability in terms of the distribution F_A. Then first-order conditions yield a differential equation with a closed-form solution: the (proposed) equilibrium bid function, as in the following Proposition. (Refer to Moldovanu and Sela (2001) for the proof.)

Proposition I (Equilibrium "Bid" Quality). *Let $X = \{F_A, \gamma, \mathbf{V}, n\}$ be a tournament. Then the unique symmetric equilibrium bid function, where $P_{j,n}$ is the probability of ranking j^{th} in ability among n contestants, is as follows:*

$$b(a) = \gamma^{-1}\left(\sum_{j=1}^{p} V_j \int_0^a \frac{1}{1-z}\frac{\partial P_{j,n}}{\partial a}(z)\,dz\right). \tag{1}$$

Therefore, the bid quality generated, conditional on ability, relates quite intuitively to prize sizes, the marginal effect of varying ability levels of probabilities of placing, and the inverse of cost.[6]

■ **Comparative Statics.** Our chief interest and where we depart from past work is in examining comparative static implications of the theory. From the equilibrium bid expression (1), we develop predictions regarding the relationship between numbers of contestants and bidding behavior (effort and level of performance) across the ability distribution.

In our comparative static analysis, we begin by stressing effects of the *heterogeneity* of abilities and costs, rather than the particular shape of the cost function. Therefore, we simply presume the simple case of linear costs, i.e., $\gamma(x) = x$. This also has the benefit of allowing for simpler, more tractable analytical solutions, allowing us to establish a greater number of precise properties of outcomes.[7] (We clarify the implications of this assumption by also solving for the general case of convex costs.)

Proposition II (Responses to Competition by Ability). *Let* $X_n = \{F_A, \gamma, \mathbf{V}, n\}$ *and* $X_{n+1} = \{F_A, \gamma, \mathbf{V}, n+1\}$ *be tournaments differing in their number of contestants by one, with bid functions* b_n *and* b_{n+1}*. Let* $\Delta b_n = b_{n+1} - b_n$ *be the difference in bid quality level response to an added competitor. Where bid costs are linear and heterogeneous across contestants,* $\gamma(x;a) = (1-a)x > 0$,*, then* Δb_n *is "single-valleyed":* $\frac{\partial b_n}{\partial a}(a) < 0$ *for all ability levels up to some level* β *and* $\frac{\partial b_n}{\partial a}(a) > 0$ *for all* $a \geq \beta$*. Hence* Δb_n *is strictly quasi-convex in its shape. Further,* Δb_n *varies in its absolute level according to:* $\Delta b_n(0) = 0$*,* $\Delta b_n(a) < 0$ *for all* $0 \leq a < \alpha$ *and* $\Delta b_n(a) > 0$ *for all* $\alpha < a \leq m$*.*

[6] Although the probability of attaining a given rank is, in principle, determined by a number of complex structural features of the environment and strategic interactions, within the empirical analysis we can simply estimate this probability directly from the data.

[7] Proposition 2 is most directly related to lemma 2 in Moldovanu and Sela (2006). That establishes the results on the sign of the effect of competition, but not the quasi-convexity.

Proof. See Appendix.

Therefore, provided there are linear and heterogeneous costs of improving the bid quality level by contestants, we predict the response to competition across different ability levels should vary in a rather precise and particular way, as illustrated in Figure 1. The empirical predictions are as follows:

(i) The response to competition is zero at the origin among lowest-skilled contestants.

(ii) The response to competition decreases and becomes negative at higher levels of ability, up to a unique minimum at a point, β.

(iii) The response to competition then becomes more positive (less negative) at ability levels above, β, and continues to increase with ability level.

(iv) The response to competition finally increases to the point of becoming absolutely positive at a point $\alpha > \beta$.

(v) The response continues to increase with higher levels of ability until reaching the upper bound of ability, m.

<Figure 1>

The case of general convex and heterogeneous costs, i.e., $\gamma > 0$, $\gamma' > 0$, and $\gamma'' > 0$ (and following all earlier characterizations of the environment) is quite similar. The non-monotonic sign of the response to added contestants matches that of the linear case. The response to added contestants begins negative and stays negative until some level, α; then, the response to added contestants becomes positive and stays positive until the upper bound of abilities, m. However, less can be said about the "single-valley" property. Instead of decreasing monotonically until some ability level, β, the response to competition could plausibly increase and decrease over subregions, but remain negative. Similarly, after the skill level at which the response to added contestants turns positive, α, the response need not necessarily increase monotonically, but instead could decrease over sub-regions, but remain positive. See the Appendix for a proof.

3 Empirical Context and Data

■ **TopCoder Software Algorithm Contests.** Data for our study comes from TopCoder, Inc., a web-based platform that delivers outsourced software and algorithmic solutions for firms, non-profit organizations, and government entities. It is the leading contest platform for delivering custom enterprise-scale software solutions through a contest format, regularly delivering sophisticated outsourced software projects for Fortune 1000 companies and government agencies since 2001. Roughly half a million solvers have signed up as members to the platform and tens of thousands regularly participate. The contests and work in each case is carried out online, allowing participation from most countries around the world.

TopCoder runs contests of a number of types. Here we study data from its regular weekly "Algorithm" contests, in which contestants provide computer program solutions to computer science algorithm problems over the course of 75 minutes. These problems are designed by TopCoder as a means of engaging and sustaining interest in its population of members with interesting and challenging problems. These contests also allow skill levels of contestants to be determined, as contestants typically participate in dozens of such contests over the course of many months or years. TopCoder uses an Elo-based system of measuring skills (Maas and Wagenmakers, 2005) as is standard in a range of contexts from chess grandmaster tournaments to US College Bowl systems to the National Scrabble Association and the European Go Federation. The system essentially predicts future rank based on history of ranks in past contests to that point. Typical contestants participate in dozens of individual contests.

Within the contests, participants are to provide solutions to three problems over the course of each 75 minute contest. Precise quantitative scores are generated automatically by the platform according to the correctness and speed with which individual solutions are completed after a problem is "opened" by a contestant. The most common distribution of point values is roughly 250, 500, and 1000 for the three problems, distinguishing the problems as "easy", "medium", and "hard". The points received in a contest are the sum points received for each problem. In each event, registered contestants, typically numbering several hundred, are assigned to virtual contest "rooms," not exceeding 20 contestants and typically ranging from 16 to 20 contestants, leading to roughly 51 (10.8) independent contests held at a time, each week.

Prior to the start of a given event, a coder does not know the identity or number of other contestants, the precise number of independent rooms into which it will be divided, or the problems they will encounter. For those events featuring cash prizes, this is known prior to registering for the event. The prize pool per contest is roughly \$5,000 (\$5,000.36 on average, ranging from \$4,969.00 to \$5,032.00). The cash prize pool is divided up evenly across the individual independent contest rooms. First and second place contests both receive prizes in each independent contest room. First place receives a higher prize than second place, with precise levels varying across events.

■ **Sample.** Given our econometric approach (Section 4), our interest here is to study a short panel within a most stable period of TopCoder's history during which the assignment of contests to rooms was based on a randomized assignment procedure. Here we study data from Algorithm contests offering cash prizes between 2005 and 2007. This period represents a period of stable commercial growth of the platform, after its initial establishment and period of experimentation with its business model. This period also precedes a period of expansion into new business (and contest) lines and the financial crisis of 2008. We also focus on just the top division of contests, where each competitor has a skill rating. (TopCoder divides participation of developer members into two divisions according to skill rating. When individuals initially join and do not have a skill rating, they join the lower division.) This implies a total of 774 independent contests (rooms) across 33 dates in our sample, in which 2,796 individual contestants participated–forming an unbalanced panel of 14,391observations of contestants within particular contests.

■ **Data and Variables.** Our analysis exploits observational data drawn directly from TopCoder's database over the sample of interest. Summary statistics of these variables appear in Table 1 below. Related to the bid or expected performance of individuals (b), we observe the precise quantitative measure of performance, total points received (*Score*). Related to individual ability (a), we observe TopCoder's Elo-based skill rating (*SkillRating*). For simplicity, we re-scale TopCoder's skill rating on a unit scale from minimum to maximum skilled. Of course, we are also interested in the number of contestants (n) and distribution of skills in a given contest (F_A). Here we directly observe the actual number of contestants (N). As regards the distribution of abilities, we observe all ability levels in the room and can thus construct summary statistics reflecting the skills distribution.

<Table 1>

■ **Random Assignment and Sources of Variation of Key Variables** The details of models estimated in the analysis are provided within the analysis section itself (Section 4). However here we wish to simply review essential features of the data that are central to our estimation approaches, particularly as regards the number of contestants and skills distributions. As a starting point, it should be stressed that features of the institutional context–including the "rules of the game," the technical platform, and the nature of tasks–are unchanging across the sample.

Our two primary variables for which we require exogenous variation in order to estimate relationships are the number of contestants in a given contest room and the skills of the individual contestant. As regards skills, we directly observe individuals' skills and can directly exploit random assignments to different rooms (inasmuch as we have dealt with any possible variation in skills distributions across rooms, as above). Variation in numbers of competitors also comes from exogenous sources. TopCoder pursues a policy of capping the number of contestants in each independent room contest at 20. This means creating some number of independent conference rooms and then randomly assigning participants to those separate conference rooms. The mean number of total registrants in these data is 949 with considerable variation – a standard deviation of 194 about this mean. As such, a first source of variation in numbers of contestants in each room is first determined simply by the imperfect divisibility of the total integer number of contestants into a fixed integer number of rooms. Therefore, while overall numbers of participants on a given day may be subject to trending and differences over certain days, the question of imperfect integer divisibility should be less subject to any such trending.[8] Another source of variation in numbers of contestants is created by dropouts. Between the time that a contest is announced and registration takes place (and before details of the contest are revealed), contests typically experience some degree of drop out. Random assignment becomes relevant here too, as this leads drop outs to also be distributed randomly across rooms.

4 Analysis

Our analysis proceeds first with flexible non-parametric estimates to test our theoretical predictions (Section 2). We then shift to estimating the structural model, allowing us to compare the constrained structural interpretation with the flexible empirical analysis, providing deeper insight on the basis of estimated structural parameters and allowing us to consider counterfactual simulations of alternative contest design policies.

[8]The remainder values when dividing total number of participants on a given day by 20 is almost perfectly uniform in distribution on $\{0, \ldots, 19\}$ providing there is no indication of non-random features of the data generation process, including systematic links to contest characteristics.

■ Flexible Nonparametric Estimates

Following our earlier characterization (Section 2), the bid function or expected performance of competitor i in contest t, b_{it}, is a function of: the number of competing contestants, n_{it}, competitor ability, a_{it}, and the distribution of abilities in the field of contestants, $F_{A(it)}$. Here we measure $\{b, n, a, F_A\}$ with empirical variables $\left\{Score, N, SkillRating, \overline{SkillRating}\right\}$. We refer to the empirical expected performance function as $g(Score, N, SkillRating)$, the empirical counterpart to the theoretical expression, $b\,(n, a, F_A)$.

In regards to the distribution of abilities $\left(F_{A(it)}\right)$, in principle we can largely rely on this variable to remain relatively constant across contests given the random assignment procedure. However, to the extent there is variation to control for, we introduce a measure of mean skill rating within a contest, $\overline{SkillRating}$ as a control variable.[9] and panel controls for time periods and trending as controls to provide greater assurance. (See discussion in Section 3.) Therefore, an unconstrained flexible empirical estimate of the bid function, or, in empirical terms,the conditional mean *Score* can be summarized in the following expression, where again $g(-)$ is the empirical function summarizing the relationship among key variables and ϵ_{it} is an additive zero-mean error term: $Score_{it} = g(N_{it}, SkillRating_{it}, \Theta_{it}) + \epsilon_{it}$.

Note, however, that our interest is not so much in the conditional mean performance, *Score*, but rather in the way in which contestants' performance responses to added numbers of contestants vary with ability level. In terms of the earlier theoretical discussion, this means an interest in estimating $\Delta b_n\,(a)$ rather than just the bid function, b. In terms of our empirical function, this is $\Delta g_N\,(Skill)$ rather than just $g\,(-)$. An added consideration is that the earlier theory suggests the expected performance function should be nonlinear (Section 2) and therefore so should $\Delta g_N\,(Skill)$ be nonlinear. To estimate $\Delta g_N\,(Skill)$ in a most flexible and revealing way, we execute two steps: we first estimate the conditional mean

[9] Results do not substantially change when including this variable, nor when including higher moments of the skills distribution. Also note, the empirical skills distribution is similar to an exponential distribution for which the mean is a sufficient statistic. Results presented here are also robust to time trends, year dummies, month dummies and day-of-the-week dummies.

performance (bid) function $g\,(-)$ for different numbers of contestants using a nonparametric estimator; and then difference in estimated bid functions and divide by the change in numbers of contestants, as the response to varying competition at different ability levels (evaluated with control variables set to their mean), as in the following expression, redefining the error term appropriately as δ:

$$\Delta g_N(SkillRating_{it} \mid \bar{\Theta}_{it}) =$$

$$\frac{g(m+\Delta, SkillRating_{it}, \overline{SkillRating}_{it}) - g(m, SkillRating_{it}, \overline{SkillRating}_{it})}{\Delta} + \delta_{it}, \qquad (2)$$

where m is some baseline number of contestants and Δ is an incremental addition to the number of contestants. To estimate conditional mean performance functions for each value of the discrete variable, N, we estimate the function $g(SkillRating, \overline{SkillRating} \mid N)$ with a Nadarya-Watson estimator using an Epanechnikov kernel and adaptive bandwidth (Pagan and Ullah, 1999). (The approach assumes a degree of smoothness and regularity in the estimated function, in the sense of being Lipschitz continuous in contestant ability and in the distribution of abilities of all contestants.) A "nearest-neighbor" adaptive algorithm was used in these estimates in which the bandwidth of the kernel adjusts at each estimation point to ensure 250 data points are included in the kernel. The number of data points was selected through cross-validation to minimize the integrated square error of the estimate. Our estimates iterate through different numbers of contestants in the room, estimating at plus and minus one standard deviation of the mean in numbers of contestants, i.e., $N = 17$ and $N = 19$. (This implies $m = 17$ and $\Delta = 2$ in expression (2)). The second step in estimating the response to competition across the skills distribution is to take the difference in estimates at different levels of N. The slope is estimated here by differencing estimates at $N = 19$ and $N = 17$ and dividing by two. Confidence intervals for the bid function are generated by bootstrapping repeated estimates on subsets of the data over the two-step procedure.

Figure 2 graphically presents our mean estimates of the slope response of *Score* with N over varying levels of *SkillRating*, along with 95% confidence intervals. Despite the estimate being produced in a flexible manner with a minimum of constraints, the patterns summarized below in observations 1, 2, 3, 4, and 5 conform precisely to the earlier theorized hypotheses of Section 2 (i, ii, iii, iv, v):

1. The response of the lowest-skilled contestants is indistinguishable from zero.
2. Proceeding rightward to those of intermediate levels of skill, the response to competition becomes increasingly negative.
3. Increasing beyond some intermediate level of skill, the response to competition increases (becomes less negative).
4. The increase continues until a skill level is reached where the response to competition becomes positive.
5. The response continues to increase with added skill and the response is most positive at the maximum ability level.

<Figure 2>

Apart from precisely conforming with theoretical predictions, the fitted model also explains a large fraction of variation. For example, in estimating the mean performance or bid functions, our nonparametric estimates reduce the sum of squared errors over a constant model by about 46%. Also note that these results conform with usual notions of a negative aggregate response to added numbers of contestants, as far and away the bulk of contestants appear in the part of the ability domain for which the response to competition is negative. Fewer than 5% of observations occur in the part of the *SkillRating* domain in which the response is positive.

■ Structural Maximum-Likelihood Estimates

In order to analyze more precise predictions of the theory, we fit a fully parameterized version of the model of Section 2 to the dataset, using maximum likelihood. Recall, from Section 2 equation (1) the expected performance or bid function takes the following form:

$$b(a) = \gamma^{-1}\left(\sum_{j=1}^{p} V_j \int_0^a \frac{1}{1-z}\frac{\partial P_{j,n}}{\partial a}(z)dz\right),$$

In estimating this function, a contestant's ability (a), the number of contestants (n), and the bid ($b(a)$) are modeled by the same variables as in the preceding subsection. The probability of ranking j^{th} in a room, $P_{j,n}$, is estimated directly from the actual patterns in the data set. The distribution of abilities is estimated by a kernel density estimate $F(z;\mathbf{a}) = K(z;\mathbf{a})$ where $\mathbf{a}$ is the vector of abilities in a competition room. Then $\frac{\partial P_{j,n}}{\partial a}(z)$ is directly calculated for each individual and contest in the dataset. The number of prizes is fixed at 2, $p = 2$, as this is constant in our data. The remaining model components ($\mathbf{V}$ and c) need to be estimated from maximum likelihood estimates from the data. We allow costs to take the form $(1-a)\gamma(x;\alpha,c) = (1-a)x^c$.[10] Additionally, we diverge from the theoretical model to allow a non-zero intercept of α. Therefore, given $\mathbf{V}$, c, and α, the structural equation is:

$$b_i = \left(\sum_{j=1}^{2} V_j \int_0^{a_i} \frac{1}{1-z}\frac{\partial P_{j,n}}{\partial a}(z)dz\right)^{\frac{1}{c}} + \alpha + \epsilon_i, \tag{3}$$

where we assume that $\epsilon_i \sim N(0,\sigma^2)$.

The maximum likelihood estimates of V, c and α solve the problem:

$$\underset{V,c,\alpha}{argmax}\, Pr\left\{b_i - \left(\sum_{j=1}^{p} V_j \int_0^{a_i} \frac{1}{1-z}\frac{\partial P_{j,n}}{\partial a}(z)dz\right) - \alpha\right\}$$

[10] Note that estimating $\gamma(x,c) = x^c$ is equivalent to also estimating $\gamma(x,d,c) = dx^c$ in our setup as d merely scales the V, which we also estimate.

$$\text{s.t. } V_1 \geq V_2 \, and \, c \geq 1$$

This problem is equivalent to the problem of minimizing the sum of squared errors over the same parameter space. The maximum likelihood estimates for the cost parameters are as follows:

$$c = 3.37, \qquad \alpha = 178.74.$$

The maximum likelihood estimates for the first and second prize values are as follows:

$$\begin{pmatrix} V_1 & V_2 \end{pmatrix} = \begin{pmatrix} 223.66 & 9.17 \end{pmatrix} \times 10^7.$$

(Note: Though the values for the prizes may seem large, the scale is determined by the scale of scores awarded in the contest and the choice of cost function, γ, so the absolute level has little meaning.) The estimated model explains about 29% more of the absolute variation than a constant model. (Recall, the fully flexible nonparametric estimate reduces the sum of squared errors by 46%.) Figure 3 depicts the bid function and estimated ability distribution for these typical values. We then estimate the marginal response to competition across skill levels by averaging the change from 17 to 18 and from 18 to 19 contestants, in order to provide a direct comparison with the earlier non-parametric estimates.

<Figure 3>

In order to evaluate the maximum likelihood structural estimate in comparison to the nonparametric bid function, Figure 3 also shows both the estimates superimposed on one another, along with confidence intervals for the non-parametric estimate. The nonparametric estimate falls within the 95% confidence interval over 93% of the domain. In other words, the theoretical prediction of the response to competition is not significantly different from the best case smooth and unconstrained fit of the actual response to competition.

■ Interpretation of Policies and Counterfactuals

The high fidelity of the structural model with unconstrained nonparametric estimates suggests this model and its structural parameters can interpret policies and contest design in these data. As noted earlier, Moldovanu and Sela (2006) demonstrates that even considering only the strategic incentives of contestants (ignoring any psychological or other non-economic influences), optimal design does not have a simple solution when costs are convex. Given the particular success of TopCoder in designing contests–attracting roughly half a million contestants and servicing a large roster of clients with technology developed in a regular stream of contests with high participation and performance–the policies of TopCoder might be judged of particular interest. Here we consider key contest design policies they have implemented by examining counterfactual experiments.

☐ **Capping Contest Size** The number of contestants in each contest in the data set varies in the high teens and does not exceed 20. This follows TopCoder's policy of creating new contest "rooms" when there are sufficient contestants registering, rather than 20 contestants per room. Structural estimates of our theoretical model allows us to easily simulate the impact of deviating from this policy–simply by varying the n parameter in structural model. Figure 4 plots the difference in bid functions from $n = 19$ and $n = 24$, reflecting both a current typical scenario (19), as well as heightened competition. Increasing the number of competitors in the room to 24 is projected to increase the scores of the highest ability contestants significantly– up to 189 points. While the scores of moderate ability contestants fall up to 110 points. (Simulating with lower levels of competition produces opposite patterns.)

<Figure 4>

Given these results, if the goal were simply to maximize the highest overall score in these periodic contests, adding a greater number of contestants should better achieve this goal. TopCoder virtually achieves this goal with its annual "TopCoder Open" tournament, in which a large number of the strongest contestants are invited to compete. Further, in TopCoder's contests geared at solving software and algorithmic challenges to general commercial products for paying customers (outside of our data set), the company has a very different policy where it places no constraints whatsoever on the number of contestants who might enter and compete for a given project. A different way of interpreting the results of the simulation with an artificially high (24) number of contestants, however, is that while the peak score might be boosted on account of strategic incentives by 189 points and lower scores might appear to fall by a lower amount of 110, in fact the weighted average effect is highly negative, simply because far and away the greatest mass of contestants resides within the part of the curve that is negatively affected by added competition. Thus, if there is some interest in the wider cross-section performing with high effort, an added boost in competition may be undesirable in dampening incentives for a great many contestants. In fact, the true objective of these conferences is to stimulate and maintain the interest of a large fraction of the roughly half a million members who have signed up to the TopCoder platform. Therefore, the policy of not exceeding 20 contestants would appear to support some kind of tradeoff between stimulating high effort and high-flying performance among right-tail contestants, while attempting to avoid dampening the incentives of lower-ranked contestants (who constitute the vast majority of the TopCoder membership).

□ **Prize Allocation and Structure** An implication of capping the size of independent contest rooms, as discussed above, is that this might also imply the magnitude of the prize. For example, larger rooms might imply fewer rooms and greater allocation of cash prizes to each room, if in fact the prize pool is fixed. We can simulate this added effect of capping contest size by repeating the comparison of 19 versus 24 contestants, as in the earlier analysis. However, here we consider the effects of proportionally increasing the amount of the prize (i.e., increasing independent contest room size by 26%, as in this comparison, could imply $1/(1+26\%)$ fewer rooms and $1 + 26\%$ times the allocation of a fixed prize per room). As can be seen in Figure 5, this adds a disproportionate effect in boosting the highest-skilled contestants, further accentuating the earlier findings that a larger contest favors the performance of the highest-skilled contestants. The boost to performance created by a higher prize does not outweigh the reduced incentives from higher competition for contestants below the top tier.

<Figure 5>

Another regular policy of prize allocation policy pursued by TopCoder is the allocation of two prizes per independent contest room–i.e., not only does the top winner receive a prize, but so does the competitor submitting the second-best solution. If we approximate TopCoder's goal as seeking to maximize overall performance in terms of the total sum of scores, the contest design goal is as follows:

$$\max_{V \in \Re^p_+} \int_0^m b_n(z; \mathbf{V}) f(z) dz.$$

Moldovanu and Sela (2001) show that distributing the prize pool across two or more prizes can only be optimal where costs are convex, and two prizes will be preferred to a single prize if and only if

$$\int_0^m (B(a) - A(a)) \frac{\partial \gamma^{-1}}{\partial x} A(a) f(a) da > 0, \quad (4)$$

where $A(a) = \int_0^a \frac{1}{1-z} \frac{\partial P_{1,n}}{\partial a}(z) dz$ and $B(a) = \int_0^a \frac{1}{1-z} \frac{\partial P_{2,n}}{\partial a}(z) dz$ are the weights on the first and second prizes in the linear cost bid function. Note that equation (4) only depends on the distribution of abilities and cost structure and, therefore, we can determine whether the condition for optimal allocation of prizes is maintained even despite our estimates being drawn from data in which only two prizes are used. Evaluating equation (4) based on our maximum likelihood estimates for $n = 19$, $\gamma(x) = x^c$, and $F(a) = K(a; \mathbf{a})$ and using the pooled distribution of abilities, the left-hand side has a value of 0.1683. Hence, TopCoder is correct in splitting prize money over two prizes, if the goal is maximizing overall output.

□ **Distinct High- and Low- Ability Contest Divisions** One further design variable available for manipulation is segmentation of tournaments by ability. TopCoder's policy is to divide its body of contestants into two roughly equally sized pools of contestants according to a cutoff ability level. (Our analysis here focused only on the high-skilled division.)

To consider the effect of the segmentation policy, here we simulate the implications of further segmenting those within our data set (in the high-skilled division) into two additionally segmented divisions, according to skills above or below the level of 0.5. Figure 6 illustrates how each half of the ability distribution would react to such a split. The lower half of the ability distribution shows a universal improvement in performance up to 366 points, as a result of removing competition for extremely able contestants. In effect, intermediately-skilled contestants would acutely feel the dampening effects of competition within the wider division and experience a drop in competitive intensity and a greater likelihood of winning where competition is more likely to produce a stimulating rivalry effect.

<Figure 6>

The upper half of the ability distribution shows a mix of reactions: some abilities show decreased performance, others increased performance. The performances of those with abilities from 0.5-0.7 show a large drop – up to 1430 points. Without this divisional split these contestants were in the 90^{th} percentile and there was a large likelihood that any added contestants would be drawn from below them in the skills distribution. However, in this revised division they are closer to the bottom of the division and adding contestants from only among higher skilled contestants makes them now more acutely sense the dampening effects of competition. As they have little chance of winning a prize, they put in little effort. By contrast, higher ability contestants show a very large increase in performance– up to 789 points.[11] The increase in quality of contestants forces them to compete harder to win. Thus, consistent with TopCoder's active advertising of the virtues of competition in stimulating

[11]This value should be seen as only indicative of the large potential for performance increases, as TopCoder's scoring system has fixed maximal scores that, in fact, make this extrapolation not feasibly attainable.

high quality solutions, within this context it in fact appears that the very top contestants sense particularly heightened strategic incentives to supply higher levels of effort with higher competition–but particularly with higher competition among comparable rivals. This again corroborates the use of the TopCoder Open invitational tournament among the very highest skilled contestants as a stage for the most competitive rivalries. More broadly, given the objectives of engaging a wide cross section of its membership with these contests, it can be understood why the company avoids too finely-grained divisions so as not to disincentive its 90th percentile contestants.

□ **"Open" Membership to the Platform** A direct extension of the earlier two issues is to consider TopCoder's policy of open admissions to its platform–irrespective of preparation, skill, or background. While there many be any number of reasons for the company to pursue this inclusive approach, what is clear from the results and earlier points is that there is little downside to open admissions. First, the sheer number of possible contestants is made irrelevant by capping the number of participants in any one independent contest room. This represents a qualitative departure from most contests we have seen in history where contestants are not cordoned into separate independent contests. With this question of number having been dealt with, there is then only the question of skills distribution. A possible worry, of course, is that the platform becomes flooded with low quality participants and this could alter the distribution of abilities of participants in ways that might lessen rivalry among the most able contestants, in addition to other possible problems. However, to the extent this could plausibly become a problem (we found no suggestion it was in our interactions with the company and its trade partners), the divisionalization policy would likely deal with this contingency in a simple fashion. The creation of an upper skill division with a minimum threshold skill effectively fixes the distribution of abilities, F_A, in that division, a virtual form of certification. This limits any effects of low quality entrants to the lower division.

5 Conclusions

This article analyzes how the level of competition and size of a tournament affects performance as a result of how strategic interactions affect contestants' incentives to exert high levels of effort. We argue that, under relatively general conditions describing a one-shot tournament, the incentive response and performance of contestants should be a nonlinear function of the ability and skill level of contestants. The response to increased competition across increasing ability levels should initially decrease at greater skill levels and eventually become more positive (less negative) and possibly even turn positive at highest skill levels. Therefore, while aggregate and average patterns of performance and effort may decline with increased competition, performance and effort may in fact increase among the highest-skilled contestants.

The sometimes stimulating effect of competition is analogous to the longtime usual intuition that it is better to have two contestants rather than one in a tournament, as the presence of at least one more competitor of sufficient skill can generate a need to exert more effort at the margin to maximize one's expected payoff (Harris and Vickers, 1987). However, whereas much of the literature–both theoretical and empirical–has stressed that increased competition beyond a minimum level may reduce the probability of winning to a level where incentives become depressed, here we clarify this stimulating effect of rivalry may persist at least for the highest-skilled contestants. This is because the addition of greater numbers of contestants increases the likelihood that "right-tailed" contestants sense some level of sufficiently skilled contestants to experience the stimulating effect of rivalry and it is possible this stimulating effect of rivalry may outweigh the incentive dampening effects of competition. We illustrate these arguments within the analytical framework developed by Moldovanu and Sela (2001), which features a one-shot n-player tournament with the possibility of multiple prizes and contestants of heterogeneous abilities. Our arguments depend principally on examining comparative statics in relation to varying levels of competition and varying skill levels.

Our main contribution is in studying fine-grained evidence on individual competitor outcomes from 774 software algorithm development contests, where it is possible to identify causal effects by exploiting quasi-experimental variation due to the random assignment procedure employed by the contest sponsor, TopCoder. Equally important, this context offers a rare opportunity to observe precise measures of individual competitor skill and performance outcomes, based on objective observational measures. The performance response to competition by skill level is first estimated with a nonparametric kernel estimator, providing the best-fit relationship with a minimum of constraints imposed. The estimate agrees with the theoretical predictions, showing that least skilled contestants are negligibly affected by rising competition.In addition, with higher levels of skill, the response becomes progressively more negative until, towards the range of highest-skilled contestants, the relationship becomes more positive (less negative) and the response to competition finally turns positive for the very highest skilled contestants–in a sort of asymmetric-U shape (with the right hand side higher than the left). Therefore, the flexibly-estimated relationship conforms to the quite particular predictions of the shape following the theory and arguments. We also find that our maximum likelihood estimate of the structural model produces a very similar estimated response to competition across the skills distribution, further affirming our analysis and conclusions.

We use the structurally estimated model to interpret the design of the contests within our data set and to simulate counterfactuals related to several key contest design policies. These include the capping the number of entrants, the dividing of a fixed prize pool among multiple independent contests held simultaneously, the prize structure in each contest, the creation of distinct divisions of contestants divided by skill level and the policy of allowing open entry to all comers on the platform. What becomes clear in this discussion is that this wide range of instruments can, at least in part, be interpreted as a means of managing both the level of competition and the skills distribution in a way as to manage tradeoffs created by non-monotonic responses to competition.

REFERENCES

Aghion, P., Bloom, N., Blundell, R., Griffith, R., Howitt, P. “Competition and innovation: An inverted-u relationship.” *Quarterly Journal of Economics.* Vol. 120 (2005), pp. 70-728.

Bershteyn, B., Roekel, S. V., eds. *Memorandum for General Counsels and Chief Information Officers for Executive Departments and Agencies.* Executive Office of the President, 2011. URL https://cio.gov/wp-content/uploads/downloads/2012/09/Prize_Authority_in_the_America_COMPETES_Reauthorization_Act.pdf

Boudreau, K., Lacetera, N., Lakhani, K. “Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis.” *Management Science*. Vol. 57 (2011), pp. 843-863.

Brown, J. “Quitters never win: The (adverse) incentive effect of competing with superstars.” *Journal of Political Economy.* Vol. 119 (2011), pp. 982--1013.

Brunt, L., Lerner, J., Nicholas, T. “Inducement prizes and innovation.” Discussion paper SAM 25 2011, Norwegian School of Economics, 2011.

Carpenter, J., Matthews, P., Schirm, J. “Tournaments and Office Politics.” *American Economic Review*, Vol. 100 (2010), pp. 504-17.

Che, Y.-K., Gale, I. “Optimal design of research contests.” *American Economic Review.* Vol. 93 (2003), pp. 646-671.
Ehrenberg, R. G., Bognanno, M. L., 1990a. “Do tournaments have incentive effects?” *Journal of Political Economy*. Vol. 98 (1990a), pp. 1307-1324.

Dechenaux, E., Kovenock, D., & Sheremeta, R. A survey of experimental research on contests, all-pay auctions and tournaments. All-Pay Auctions and Tournaments.” (2012) Available at SSRN: http://ssrn.com/abstract=2154022.

Dohmen, T., Falk, A. Performance pay and multidimensional sorting: Productivity, preferences, and gender. *The American Economic Review*. Vol. 101 (2011), pp. 556-590.

Ehrenberg, R. G., Bognanno, M. L., 1990b. “The incentive effects of tournaments revisited: Evidence from the european PGA tour.” *Industrial and Labor Relations Review.* Vol. 43 1990b, pp. 74-88.

Fu, Q., Lu, J. “The optimal multi-stage contest.” *Economic Theory*.

Eriksson, T. “Executive compensation and tournament theory: Empirical tests on Danish data.” *Journal of Labor Economics*. Vol. 17 (1999), pp. 262-280.

Fullerton, R. L., Linster, B. G., McKee, M., Slate, S. “Using auctions to reward tournament winners: Theory and experimental investigations.” *The RAND Journal of Economics* Vol. 33 (2002), pp. 62-84.

Fullerton, R. L., McAfee, R. P. “Auctioning entry into tournaments.” *Journal of Political Economy.* Vol. 107 (1999), pp. 573-605.

Harris, C., Vickers, J. “Racing with uncertainty.” *The Review of Economic Studies.* Vol. 54 (1987), pp. 1-21.

Knoeber, C. R., Thurman, W. N. “Testing the theory of tournaments: An empirical analysis of broiler production.” *Journal of Labor Economics.* Vol. 12 (1994), pp. 155-179.

Konrad, K. A. *Strategy and Dynamics in Contests*. Oxford, UK: Oxford University Press, 2009.

Kremer, M., Williams, H. Incentivizing innovation: Adding to the tool kit. In: Innovation Policy and the Economy. University of Chicago Press, pp. 1-17, 2010.

Lazear, E., Rosen, S. “Rank-order tournaments as optimum labor contracts.” *Journal of Political Economy.* Vol. 89 (1981), pp. 841-864.

Moldovanu, B., Sela, A. “The optimal allocation of prizes in contests.” *American Economic Review.* Vol. 91 (2001), pp. 542-558.

Moldovanu, B., Sela, A., 2006. “Contest architecture.” *Journal of Economic Theory Vol.* 126 (2001), pp. 70-96.

Murray, F., Stern, S., Campbell, G., MacCormack, A. “Grand innovation prizes: A theoretical, normative, and empirical evaluation.” *Research Policy.* Vol. 41 (2012), pp. 1779-1792.

Orszag, J. “A new look at incentive effects and golf tournaments.” *Economics Letters*. 46 (1994), pp. 77-88.

Pagan, A., Ullah, A. *Nonparametric Econometrics*. Cambridge University Press, 1999.

Reinganum, J., "Chapter 14 The timing of innovation: Research, development, and diffusion", In: Richard Schmalensee and Robert Willig, Editor(s), *Handbook of Industrial Organization*, Elsevier, Vol. (1989), Pages 849-908

Scholz, F., Stephens, M. “K-sample anderson-darling tests.” *Journal of the American Statistical Association*. Vol. 82 (1987), pp. 918-924.

Szymanski, S. “The economic design of sporting contests.” *Journal of Economic literature*, Vol. 41 (2003), pp. 1137–1187.

Tanaka, R., Ishino, K. “Testing the incentive effects in tournaments with a superstar.” *Journal of The Japanese and International Economies*, in press.

Taylor, C.. “Digging for golden carrots: An analysis of research tournaments.” *American*

Economic Review. Vol. 85 (1995), pp. 872-890.

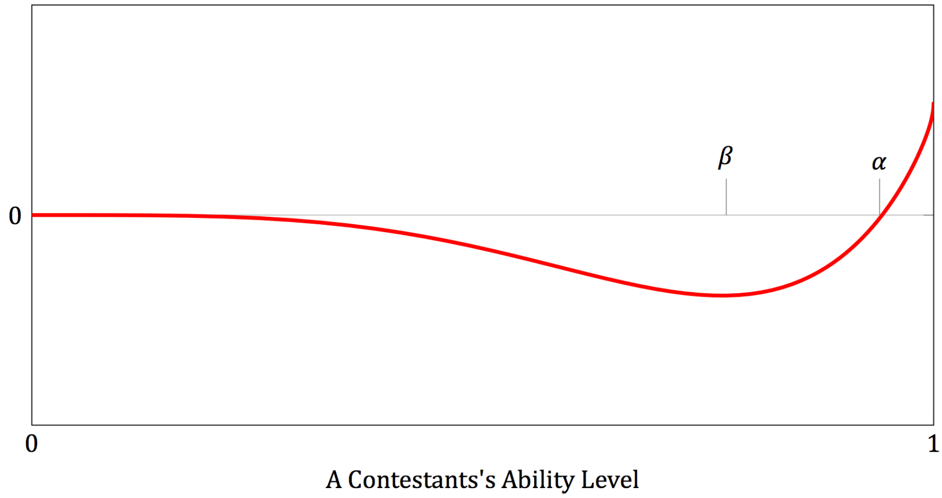
Terwiesch, C., Xu, Y. "Innovation contests, open innovation, and multiagent problem solving." *Management Science.* Vol. 54 (2008), pp. 1529-1543.

Tullock, G. "Efficient rent seeking", in Buchanan, J.M., Tullock, R., and G. Tullock (eds.) *Toward a theory of rent seeking society*, Texas A&M University Press, College Station (1980).

Wright, B., "The economics of invention incentives: Patents, prizes, and research contracts." *The American Economic Review.* Vol. 73 (1983), 691-707.

FIGURES

FIGURE 1
Predicted ``Single-Valleyed" Non-Monotonic Response to Competition



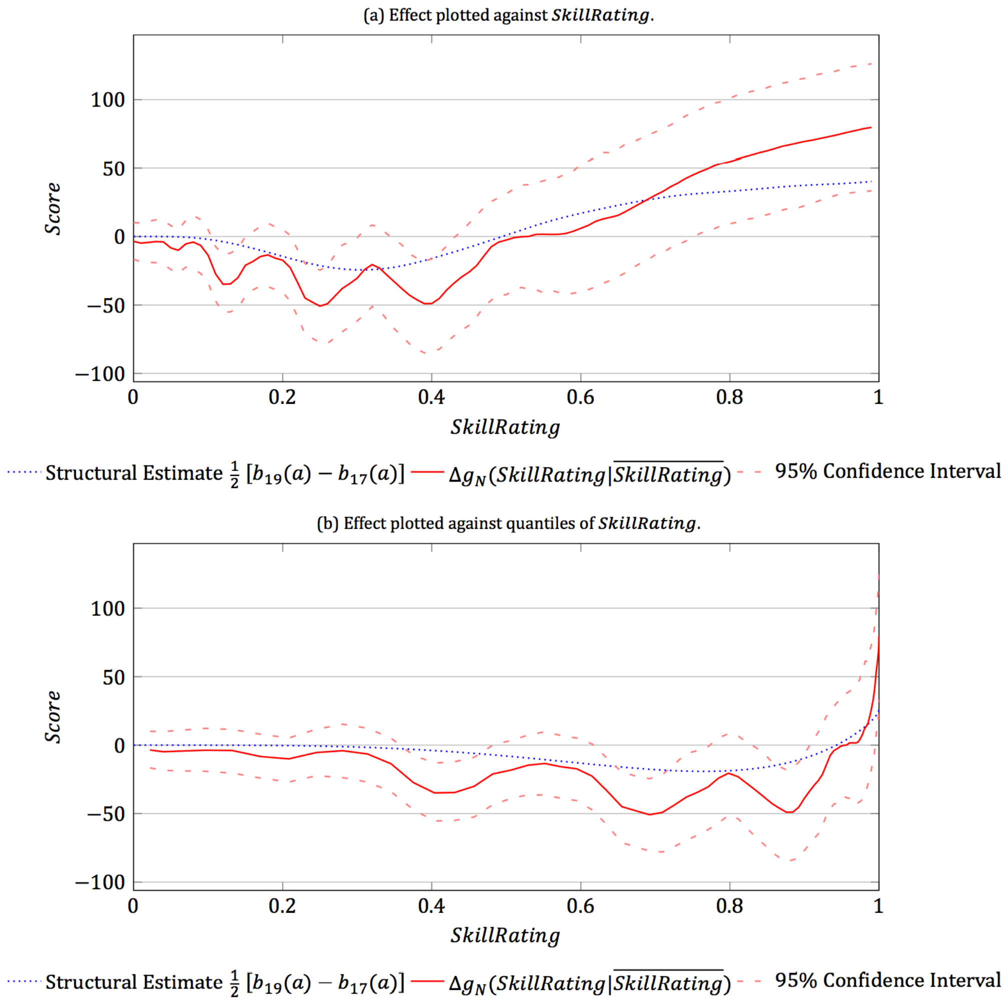
Note: Illustration of the response to competition implied by proposition 2. The change in bid quality and expected performance caused by a change in the number of contestants from n to $n+1$ plotted by ability. The level of ability β indicates the point at which increases in ability begin to result in more positive (less negative) responses to competition. The level of ability α indicates the point at which increases in ability begin to result in absolutely positive responses.

FIGURE 2

Flexible Nonparametric Estimation of Performance Response to Added Contestants

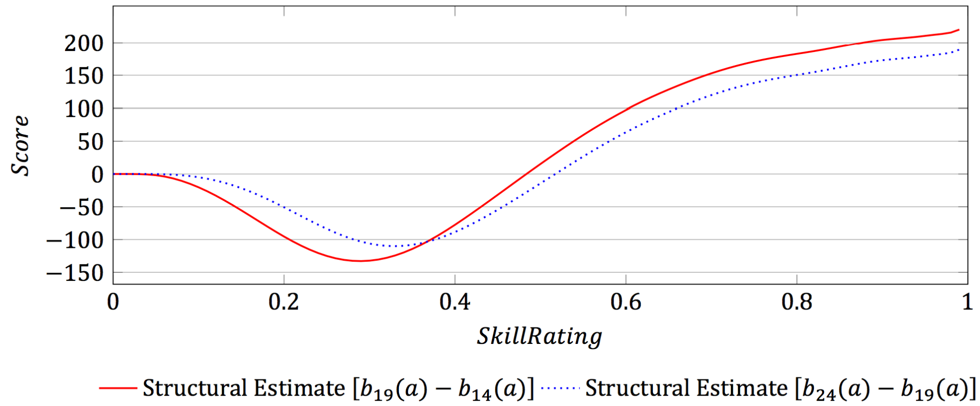
Note: The figure presents estimated effect of increasing number of contestants from *N=17* to *N=19*, across varying *SkillRating*, based on a Nadarya-Watson estimator using Epanechnikov kernel of 250 nearest-neighbour data points; bootstrapped confidence intervals. Over 95% of data points are to the left of the point at which the line crosses zero. The patterns conform to theorised hypotheses i, ii, iii, iv and v of Section 2.

FIGURE 3
Structural Estimation of Performance Response to Added Contestants



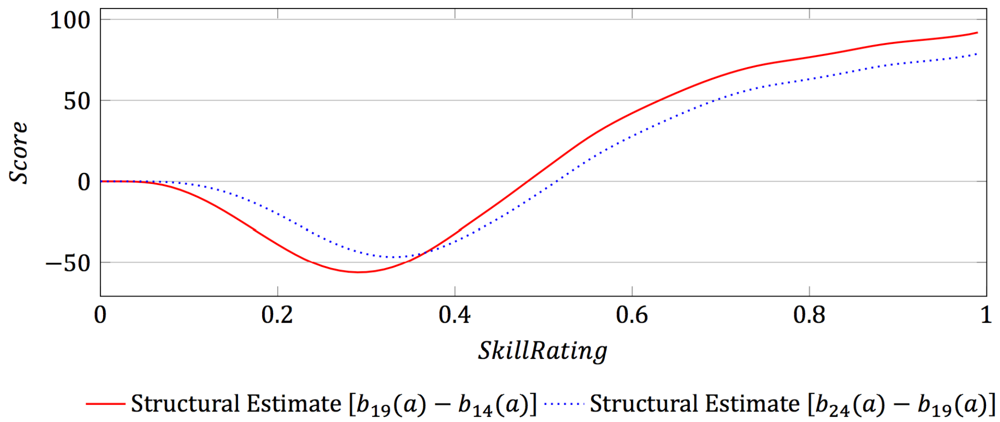
Note: The figure presents a maximum likelihood structural estimate of the effect of increasing number of contestants from *N=17* to *N=19*, across varying *SkillRating*, based on the model presented in Section 2. Results are plotted along with the earlier nonparametric estimate from in Figure 2.

FIGURE 4
Simulated Magnitude of Impact from Capping Participation



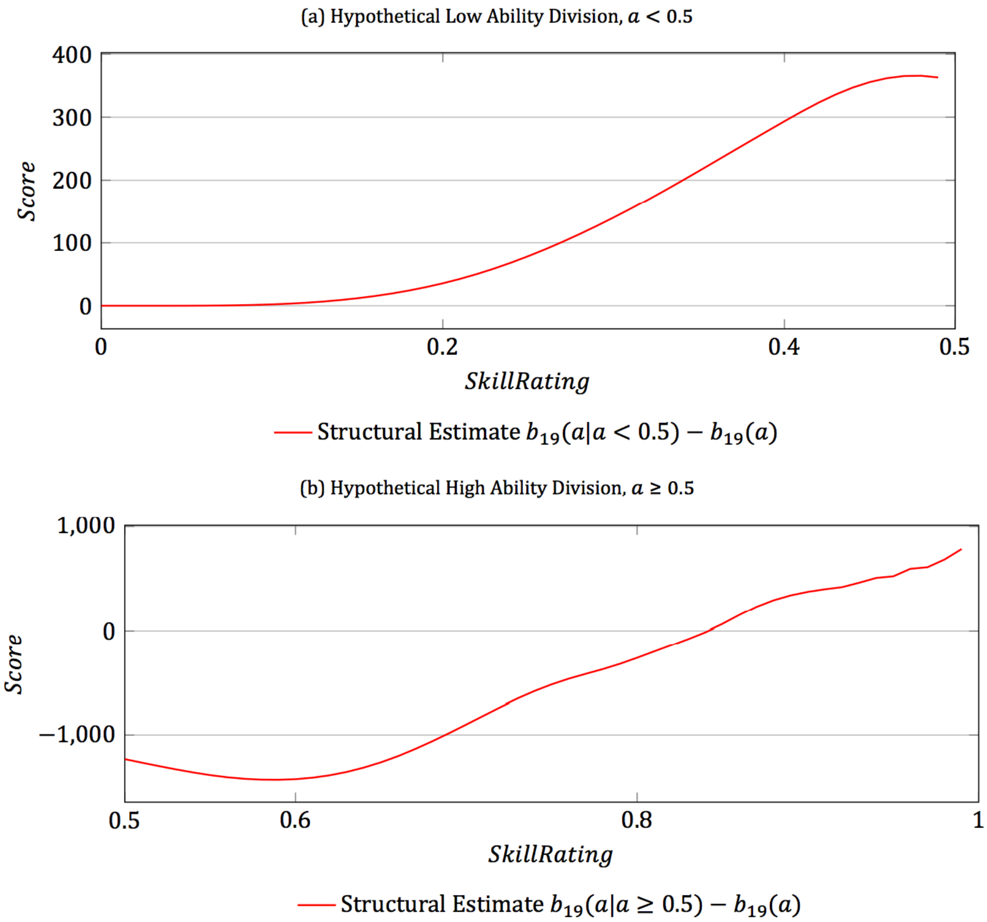
Note: Projection of performance response to large changes in the number of contestants using the structural estimate in order to assess the current policy of TopCoder to cap the number of contests at 20.

FIGURE 5
Simulated Magnitude of Impact from Capping Participation and with a Fixed Prize Pool



Note: Projection of performance response to large changes in the number of contestants using the structural estimate in order to assess the current policy of TopCoder to cap the number of contests at 20. The indirect impact of a proportionally changed prize pool is included in the response.

FIGURE 6
Simulated Magnitude of Impact from Creating Separate Divisions



Note: Bidding projections using the structural estimate. The predicted change in bid function resulting from splitting the current competitor pool into two divisions. A low ability division composed of those with abilities below 0.5 and a high ability division composed of those above 0.5.

TABLES

TABLE 1 Summary Statistics of Estimation Variables

Notation	Theoretical Counterpart	Variable Description	Mean	Std. Dev.	Min.	Max
$Score$	b	Final score in a competition.	271.09	274.22	0.00	1722.00
$SkillRating$	a	TopCoder rating.	0.20	0.16	0.01	0.99
$\overline{SkillRating}$	F_A	Mean of the TopCoder ratings in the competition room.	0.20	0.04	0.05	0.37
N	n	Number of competitors in the competition room.	18.66	1.08	15.00	20.00

Appendices

A Linear Costs

The probability of winning prize j is closely related to order statistics. Let $F_{j,n}$ be the distribution of the j^{th} *lowest* value from a sample of n random variables distributed according to F. $F_{j,n}$ can be written as

$$F_{j,n} = \sum_{k=j}^{n} \binom{n}{k} F^k (1-F)^{n-k}.$$

The pdf of $F_{j,n}$ can be written as

$$f_{j,n} = \frac{n!}{(j-1)!\,(n-j)!} F^{j-1} (1-F)^{n-j} f.$$

It follows that

$$\begin{aligned} P_{j,n} &= \binom{n-1}{j-1} (1-F)^{j-1} F^{n-j} = F_{n-j,n-1} - F_{n-j+1,n-1} \\ \frac{\partial P_{j,n}}{\partial a} &= f_{n-j,n-1} - f_{n-j+1,n-1} \end{aligned}$$

Assume costs in a Moldovanu and Sela (2001) tournament are linear, i.e., $\gamma(x) = x$. Then we can rewrite the bid function as

$$b(a) = \sum_{j=1}^{p} V_j \int_0^a \frac{1}{1-z} \frac{\partial P_{j,n}}{\partial a}(z)dz$$

$$b(a) = \sum_{j=1}^{p} V_j \int_0^a \frac{1}{1-z} f_{n-j,n-1}(z) - f_{n-j+1,n-1}(z)dz$$

$$b(a) = \int_0^a \frac{1}{1-z} \sum_{j=1}^{p} V_j (f_{n-j,n-1}(z) - f_{n-j+1,n-1}(z))dz.$$

Let $V_{p+1} = 0$. Then

$$b(a) = \int_0^a \frac{1}{1-z} \sum_{j=1}^{p} (V_j - V_{j+1}) f_{n-j,n-1}(z)dz$$

$$b(a) = \sum_{j=1}^{p} (V_j - V_{j+1}) \int_0^a \frac{1}{1-z} f_{n-j,n-1}(z)dz$$

Consider two tournaments X_n and X_{n+1} such that X_n has n competitors and X_{n+1} has $n+1$ competitors. We can write the bid functions for X_n and X_{n+1} and take the difference

$$b_{n+1}(a) - b_n(a) = \sum_{j=1}^{p} (V_j - V_{j+1}) \int_0^a \frac{1}{1-z} f_{n-j+1,n}(z)dz - \sum_{j=1}^{p} (V_j - V_{j+1}) \int_0^a \frac{1}{1-z} f_{n-j,n-1}(z)dz$$

$\Delta b_n = b_{n+1}(a) - b_n(a)$

$$\Delta b_n = \sum_{j=1}^{p} (V_j - V_{j+1}) \int_0^a \frac{1}{1-z} (f_{n-j+1,n}(z) - f_{n-j,n-1}(z))dz$$

$\frac{\partial \Delta b_n}{\partial a}$ can be computed directly.

$$\frac{\partial \Delta b_n}{\partial a} = \sum_{j=1}^{p} (V_j - V_{j+1}) \frac{1}{1-a} (f_{n-j+1,n}(a) - f_{n-j,n-1}(a))$$

$$\frac{\partial \Delta b_n}{\partial a} = \frac{1}{1-a} \sum_{j=1}^{p} (V_j - V_{j+1}) (f_{n-j+1,n}(a) - f_{n-j,n-1}(a))$$

$$\frac{\partial \Delta b_n}{\partial a} = \frac{1}{1-a} \sum_{j=1}^{p} (V_j - V_{j+1}) \frac{-n!}{(n-j)!\,(j-1)!} F(a)^{n-j-1} (1-F(a))^{j-1} f(a) \left\{ \frac{n-j}{n} - F(a) \right\}$$

$$\frac{\partial \Delta b_n}{\partial a} = \frac{1}{1-a} f(a) \frac{F(a)^{n-1}}{1-F(a)} \sum_{j=1}^{p} (V_j - V_{j+1}) \frac{-n!}{(n-j)!\,(j-1)!} \left(\frac{1-F(a)}{F(a)} \right)^j \left\{ \frac{n-j}{n} - F(a) \right\}$$

$$\frac{\partial \Delta b_n}{\partial a} = \frac{f(a)}{1-a} \frac{F(a)^{n-1}}{1-F(a)} \sum_{j=1}^{p} \frac{-n!\,(V_j - V_{j+1})}{(n-j)!\,(j-1)!} \left(\frac{1-F(a)}{F(a)} \right)^j \left\{ \frac{n-j}{n} - F(a) \right\}$$

In order to establish that $\frac{\partial \Delta b_n}{\partial a} < 0$ for all $a < \beta$, we show that for any $a, a' \in [0, m)$ such that $a < a'$, if $\frac{\partial \Delta b_n}{\partial a}(a') < 0$ then $\frac{\partial \Delta b_n}{\partial a}(a) < 0$.

$$\begin{aligned} &\frac{\partial \Delta b_n}{\partial a} \leq 0 \\ \Leftrightarrow \frac{f(a)}{1-a} \frac{F(a)^{n-1}}{1-F(a)} \sum_{j=1}^{p} \frac{-n!\,(V_j - V_{j+1})}{(n-j)!\,(j-1)!} \left(\frac{1-F(a)}{F(a)} \right)^j \left\{ \frac{n-j}{n} - F(a) \right\} &\leq 0 \\ \Leftrightarrow \sum_{j=1}^{p} \frac{n!\,(V_j - V_{j+1})}{(n-j)!\,(j-1)!} \left(\frac{1-F(a)}{F(a)} \right)^j \left\{ \frac{n-j}{n} - F(a) \right\} &\geq 0. \end{aligned}$$

Similarly,

$$\begin{aligned} &\frac{\partial \Delta b_n}{\partial a} \geq 0 \\ \Leftrightarrow \frac{f(a)}{1-a} \frac{F(a)^{n-1}}{1-F(a)} \sum_{j=1}^{p} \frac{-n!\,(V_j - V_{j+1})}{(n-j)!\,(j-1)!} \left(\frac{1-F(a)}{F(a)} \right)^j \left\{ \frac{n-j}{n} - F(a) \right\} &\geq 0 \\ \Leftrightarrow \sum_{j=1}^{p} \frac{n!\,(V_j - V_{j+1})}{(n-j)!\,(j-1)!} \left(\frac{1-F(a)}{F(a)} \right)^j \left\{ \frac{n-j}{n} - F(a) \right\} &\leq 0. \end{aligned}$$

Suppose that $\frac{\partial \Delta b_n}{\partial a}(a^\star) < 0$. Let $\delta > 0$.

$$\frac{\partial \Delta b_n}{\partial a}(a^\star - \delta) \leq 0$$

$$\Leftrightarrow \sum_{j=1}^{p} \frac{n!\,(V_j - V_{j+1})}{(n-j)!\,(j-1)!} \left(\frac{1 - F^\star + \delta}{F^\star - \delta}\right)^j \left\{\frac{n-j}{n} - F^\star + \delta\right\} \geq 0$$

$$\sum_{j=1}^{p} \frac{n!\,(V_j - V_{j+1})}{(n-j)!\,(j-1)!} \left(\frac{1 - F^\star + \delta}{F^\star - \delta}\right)^j \left\{\frac{n-j}{n} - F^\star\right\} + \sum_{j=1}^{p} \frac{n!\,(V_j - V_{j+1})}{(n-j)!\,(j-1)!} \left(\frac{1 - F^\star + \delta}{F^\star - \delta}\right)^j \delta \geq 0.$$

Since $\left(\frac{1-F^\star+\delta}{F^\star-\delta}\right)^j > \left(\frac{1-F^\star}{F^\star}\right)^j$ and $\frac{\partial \Delta b_n}{\partial a}(a^\star) < 0$ by assumption it follows that $\frac{\partial \Delta b_n}{\partial a}(a^\star - \delta) < 0$. Suppose that $\frac{\partial \Delta b_n}{\partial a}(a^\star) > 0$.

$$\frac{\partial \Delta b_n}{\partial a}(F^\star + \delta) \geq 0$$

$$\Leftrightarrow \sum_{j=1}^{p} \frac{n!\,(V_j - V_{j+1})}{(n-j)!\,(j-1)!} \left(\frac{1 - F^\star - \delta}{F^\star - \delta}\right)^j \left\{\frac{n-j}{n} - F^\star - \delta\right\} \leq 0$$

$$\sum_{j=1}^{p} \frac{n!\,(V_j - V_{j+1})}{(n-j)!\,(j-1)!} \left(\frac{1 - F^\star - \delta}{F^\star + \delta}\right)^j \left\{\frac{n-j}{n} - F^\star\right\} - \sum_{j=1}^{p} \frac{n!\,(V_j - V_{j+1})}{(n-j)!\,(j-1)!} \left(\frac{1 - F^\star - \delta}{F^\star + \delta}\right)^j \delta \leq 0$$

Since $\left(\frac{1-F^\star-\delta}{F^\star+\delta}\right)^j < \left(\frac{1-F^\star}{F^\star}\right)^j$ and $\frac{\partial \Delta b_n}{\partial a}(a^\star) > 0$ by assumption it follows that $\frac{\partial \Delta b_n}{\partial a}(a^\star + \delta) > 0$. Therefore $\frac{\partial \Delta b_n}{\partial a}$ is single-valleyed. In order to establish the second part of the proposition, consider the the form of $\frac{\partial \Delta b_n}{\partial a}$.

$$\frac{\partial \Delta b_n}{\partial a} = \frac{f(a)}{1-a} \frac{F(a)^{n-1}}{1 - F(a)} \sum_{j=1}^{p} \frac{-n!\,(V_j - V_{j+1})}{(n-j)!\,(j-1)!} \left(\frac{1 - F(a)}{F(a)}\right)^j \left\{\frac{n-j}{n} - F(a)\right\}.$$

For small enough a such that $F(a) < \frac{n-p}{n}$, $\frac{\partial \Delta b_n}{\partial a} < 0$. Similarly, for a such that $F(a) > \frac{n-1}{n}$, $\frac{\partial \Delta b_n}{\partial a} > 0$. Then the second part of the proposition follows from the single-valleyed property.

B Convex Costs

Suppose that costs are now convex, of the form $\gamma > 0$, $\gamma' > 0$, and $\gamma'' > 0$. We know from Moldovanu and Sela (2001) that the bid functions are transformations of the lienar case,

$$B_n(a) = \gamma^{-1}(b_n(a)) \qquad\qquad B_{n+1}(a) = \gamma^{-1}(b_{n+1}(a)).$$

Hence the difference of the bid functions is

$$\Delta B_n(a) = \gamma^{-1}(b_{n+1}(a)) - \gamma^{-1}(b_n(a)).$$

By assumption γ is strictly increasing, hence the γ^{-1} is strictly increasing as well. Hence γ^{-1} preserves the ordering of b_n and b_{n+1}. Therefore $sign(\Delta B_n(a)) = sign(\Delta b_n(a))$.