RAND Journal of Economics
Vol. 47, No. 1, Spring 2016
pp. 140–165

Performance responses to competition across skill levels in rank-order tournaments: field evidence and implications for tournament design

Kevin J. Boudreau*

Karim R. Lakhani**

and

Michael Menietti***

Tournaments are widely used in the economy to organize production and innovation. We study individual data on 2775 contestants in 755 software algorithm development contests with random assignment. The performance response to added contestants varies nonmonotonically across contestants of different abilities, precisely conforming to theoretical predictions. Most participants respond negatively, whereas the highest-skilled contestants respond positively. In counterfactual simulations, we interpret a number of tournament design policies (number of competitors, prize allocation and structure, number of divisions, open entry) and assess their effectiveness in shaping optimal tournament outcomes for a designer.

1. Introduction

■ Contests and tournaments[1] have long received wide interest from economists since the seminal work of Lazear (1981). Although explicit tournaments appear rare in the labor market, the competition for promotion among executives, academics, and others can be modelled as a tournament. In addition, tournaments have been used to induce technological advances and innovations

* London Business School and Harvard Business School; kboudreau@london.edu.
** Harvard University and Crowd Innovation Laboratory at Harvard Institute for Quantitative Social Science; k@hbs.edu.
*** Crowd Innovation Laboratory at Harvard Institute for Quantitative Social Science; mmenietti@fas.harvard.edu.
We are grateful to members of the TopCoder executive team for their considerable attention, support, and resources in the carrying out of this project, including Jack Hughes, Rob Hughes, Andy LaMora, Mike Lydon, Ira Heffan, Mike Morris, and Narinder Singh. For helpful comments, we thank seminar participants at Duke University, Georgia Tech (REER conference), Harvard Business School, and London Business School. Constance Helfat (Dartmouth) provided significant stimulating input to this article. The authors would also like to acknowledge financial support from the London Business School Research and Materials Development Grant, the Harvard Business School Division of Research and Faculty Development, and the NASA Tournament Laboratory. All errors are our own.

[1] In this article, we use the terms contests and tournaments interchangeably to denote rank-order based, relative performance evaluation incentive schemes.

throughout history. Well-known "grand challenges" (Kay, 2011; Nicholas, 2011; Brunt, Lerner, and Nicholas, 2012) include the X-prize for private space flight, Defense Advanced Research Projects Agency (DARPA) challenges to develop autonomous vehicle technologies, and the Netflix contest to improve the company's movie recommendation algorithm (Murray et al., 2012). Since the turn of the century, online contest platforms (e.g., InnoCentive, Kagel, and TopCoder) that continuously host numerous contests have emerged to solve research and development challenges for commercial companies, nonprofit organizations, and government agencies.[2] These platforms have greatly expanded the use of explicit tournaments for compensation in the labor market.

In this article, we examine the performance response of competitors to the total number of competitors in a contest. We build on the theoretical framework of rank-order contests advanced by Moldovanu and Sela (2001, 2006) to clarify arguments for a heterogeneous response, in terms of effort, across competitors of different ability levels. The framework predicts that as the number of competitors increases, competitors with the lowest ability have little response, competitors with intermediate ability respond negatively, and competitors with the highest ability respond positively.

To see the intuition behind the heterogeneous response, consider how an existing competitor responds to added competition. The optimal response of a competitor depends on ability level. For low ability competitors, the probability of winning is already quite low and adding competitors might then have little effect on the likelihood of winning and optimal effort level. By contrast, a competitor of moderate ability will more likely have his probability of winning and optimal effort level diminished by added competition. However, for a competitor of high ability who might not have required high effort to win against the bulk of competitors, adding greater numbers of competitors can increase the likelihood of facing a close competitor–thus raising the optimal effort level to "stay ahead of the pack." This is an effort-inducing rivalry or racing effect (cf., Harris & Vickers, 1987).[3]

Our main contribution is to estimate the relationship between performance and competition across the distribution of ability. We study a field context, algorithm programming contests run by TopCoder, Inc. We use data on 755 cash-prize contests between 2005 and 2007, in which varying numbers of randomly assigned individuals competed to solve software algorithm problems. The response to varying numbers of competitors is estimated using a nonparametric, kernel estimator. We find the specific, heterogeneous relationship predicted by theory. We then estimate a parameterized version of the Moldovanu and Sela (2001) model, revealing results consistent with the nonparametric model and affirming the usefulness of this framework. Next, we consider a series of counterfactual contest design questions based on the structural estimates. We examine the performance and cost implications of several design dimensions: the number of competitors, the number of skill divisions, the distribution of prizes, and open entry to tournaments. A range of contest design policies allows statistically and economically significant manipulation of tournament outcomes. Given the widespread use of tournaments in the economy and potentially different objectives of tournament sponsors, these policies provide useful "levers" for tournament designers. For example, sales managers may run contests with the goal of maximizing total sales (Casas-Arce and Martínez-Jerez, 2009), whereas those managing a research and development tournament may only be concerned with attracting the best possible solution (Fullerton and McAfee, 1999).

The article proceeds as follows. In Section 2, we discuss the related literature on tournaments and all-pay auctions. We develop predictions based on the Moldovanu and Sela (2001) model of tournaments in Section 3. Section 4 describes the empirical context and data set. Section 5

[2] The US government recently passed legislation giving prize-based procurement authority to all federal agencies (Bershteyn and Roekel, 2011).

[3] Analogous arguments regarding countervailing effects of competition on innovation incentives have been made using different setups and distinct mechanisms in areas such as market competition (Aghion et al., 2005) and patent races (Schmalensee, Armstrong, and Willig, 1989).

describes a nonparametric estimator for the bid function and presents the estimate, evaluating its fit. Section 6 describes a semiparametric estimator of the bid function incorporating the Moldovanu and Sela (2001) model. We compare its performance to that of the nonparametric estimator. Section 7 extrapolates the semiparametric results to counter-factual environments. Section 8 concludes.

2. Related literature

■ Within the theoretical literature, a number of important tournament design questions have been examined, including when contests are efficient relative to alternative incentive schemes (Lazear, 1981), questions around the number and abilities of competitors (Fullerton and McAfee, 1999), and questions around prize size and structure (Moldovanu and Sela, 2001, 2006).[4]

The theoretical literature on research and development and innovation contests generally points to smaller contests as producing higher incentives, where even just two competitors have been argued to produce the highest incentives (Fullerton and McAfee, 1999; Che and Gale, 2003). Absent any form of competition, competitors 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 as the chance of winning falls. The broader theoretical literature on contests and tournaments has also considered how the related issue of composition and heterogeneity of competitors impacts contest performance. This research line has shown that an increase in homogeneity among competitors increases aggregate effort (Konrad, 2009). Moldovanu and Sela (2001, 2006) establish a number of results on the preferred structure of competition for contest designers interested in maximizing aggregate 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, optimal design is a “hard” problem in that no solution works over all environments; the particular context needs to be considered.

The empirical tournament literature examining these core questions of design remains underdeveloped, in large part because of data limitations. Theoretical models make econometric demands that are rarely satisfied by existing data sources. Empirical work to date has nonetheless made considerable progress by establishing that higher prizes lead to higher performance (Ehrenberg and Bognanno, 1990a, 1990b; Orszag, 1994) and that competing with markedly superior opponents or “superstars” decreases performance (Brown, 2011; Tanaka and Ishino, 2012). Knoeber and Thurman (1994) test more nuanced hypotheses on tournament behavior using a unique data set on compensation for poultry producers. With the exception of Orszag (1994), these studies use a reduced-form estimation strategy.

A number of empirical articles have looked specifically at the effects of varying the number of participants in a tournament, with the basic finding that *average* performance declines as the number of competitors increases. For example, Garcia and Tor (2009) show that average test scores decreased with the number of test takers and that average test times decreased when subjects believed they were taking the test with fewer other participants. Closer to questions of heterogeneity, Casas-Arce and Martínez-Jerez (2009) find that in a dynamic multiperiod contest, leaders decrease their effort when their lead in a given period extends. Boudreau, Lacetera, and Lakhani (2011) study an order-statistic mechanism whereby adding competitors increases the observed range and dispersion of solutions through an increase in the number of “draws.” That research shows that this effect increases with increased problem-solution uncertainty. It uses data from the same context studied in this article, but their theory calls for studying data at the contest level. Here, we study responses at the level of individual contestants.

The laboratory experimental literature on tournaments is comparatively well developed. To our knowledge, the earliest such work is Bull, Schotter, and Weigelt (1987), who find that average effort in tournaments generally followed the theoretical prediction, but that the variance of efforts

[4] Szymanski (2003) evokes the core issues of contest and tournament design with vivid examples from sports.

was large. This pattern has largely been sustained (Dechenaux, Kovenock, and Sheremeta, 2012). Later experiments have explored some of the theoretical features of tournaments in the labor market—notably, the ability of tournaments to filter out common shocks to productivity among competitors (Holmstrom and Milgrom, 1991). For example, in a sales group facing a bad economic climate, where each individual faces the prospect of reduced sales, rank-based compensation filters out this effect, whereas commission-based compensation does not. Wu and Roe (2006) and Wu, Roe, and Sporleder (2006) demonstrate this effect in laboratory experiments. Schotter and Weigelt (1992) examines how affirmative action policies impact labor market tournaments. Here, policies are implemented as adjustments to the cost of effort by subjects. The authors' findings agree with the theoretical predictions, whereby reducing asymmetries in the cost of effort increases efforts by subjects. A range of extensions beyond core questions of design have also been studied with both theory and experimental results, including the design of multistage tournaments (Fu and Lu, 2012), implications of sabotage and "office politics" among competitors (Carpenter, Matthews, and Schirm, 2010), and implications of self-selection into open tournaments (Dohmen et al., 2011). List et al. (2014) study the interaction of the number of competitors and uncertainty in tournaments with stochastic effort. The authors test their theory in laboratory and field settings; finding strong agreement with theoretical predictions.

The strategic environment of tournaments is quite similar to that of all-pay auctions. The literatures often differ more by application than the underlying theoretical environment. Tournaments are typically used to model labor market issues, whereas all-pay auctions are typically used to model rent-seeking environments. Data on sporting tournaments has often been used in this area of study, but empirical work on all-pay auctions is rare. Recent work on "penny auctions" is a notable exception. These are a variation on an increasing-clock auction where bidders must pay to stay in an auction, and these payments are final. Augenblick (2014) examines behavior in an online penny auction site and finds substantial overbidding, resulting in high profits for auctioneers. Experimental work with all-pay auctions is more extensive. Most experimental work also reports significant overbidding, where the total spent by players far exceeds the value of the prize (Millner and Pratt, 1989; Davis and Reilly, 1998; Dechenaux, Kovenock, and Sheremeta, 2012).

It should be noted that the predominant approaches to conceptualizing tournament design are somewhat isomorphic to formalizations in the auction literature, with the effort exerted in a contest treated as equivalent to a bid in the auction context. Early empirical work with field data on auctions used reduced-form approaches to assess the implications of theory. For example, Hendricks and Porter (1988), Hendricks, Porter, and Wilson (1994), and Porter (1995) use data from auctions of oil and gas leases to test auction theory. Following seminal work in the mid-nineties, and given the wide use and availability of appropriate data sets, there is now an extensive literature on the structural modelling of auctions (Laffont, Ossard, and Vuong, 1995; Laffont and Vuong, 1993, 1996) with a focus on first-price auctions. Typically, models of first-price auctions depend on parameters unobserved by the econometrician, for example, the private values of the object being auctioned. Structural estimation procedures produce estimates of these private values from the observable data and, with a fully specified auction model, the optimality of design choices can be evaluated. Such structural methods have been used to evaluate the welfare implications of secret reserve prices in timber auctions (Elyakime et al., 1994), the revenue-maximizing reserve prices on eBay (Bajari and Hortaçsu, 2003a), and the presence of collusive bidding in construction contracts (Bajari and Hortaçsu, 2003b). Our work applies a similar semiparametric approach pioneered in the auctions literature to the context of innovation tournaments.

3. Theoretical model

■ Here, we develop hypotheses about how increasing competition in a contest will impact competitor performance. We build on the simple and tractable model of one-shot tournaments developed in Moldovanu and Sela (2001), which incorporates multiple prizes and competitors of heterogeneous abilities.[5]

Consider n-competitors 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 cumulative distribution F_A. The distribution is continuously differentiable, with full support, and density function 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 increasing the quality of a submission.

The quality of a player's 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 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, $b_i \in \mathbf{R}_+$. The cost of bidding, $c(b_i, a_i)$, is linear in the size of bid and decreasing in the ability of a player according to a function $\delta(a)$, that is, $c(b_i, a_i) = \delta(a)b_i$. 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 $\pi_i(b_i; a_i) = \sum_{j=1}^{p} \Pr\{r_i(b_i) = j\} V_j - \delta(a_i)b_i$. This simple characterization of the contest implies an expected payoff that is 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 by mapping abilities to bid quality levels $b : [0, m] \to \mathbf{R}_+$. 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 equilibrium bid function, as in the following proposition. The proof for the case of more than two prizes can be found in Moldovanu and Sela (2001) following equation C4. The case of two prizes is considered in Proposition 1 of Moldovanu and Sela (2001), whose proof can be found in their Appendix A.

Proposition 1. Let $X = \{F_A, \delta, \mathbf{V}, n\}$ be a tournament. Then, the unique, symmetric, equilibrium bid function, where $P_{j,n}(z)$ is the probability of ranking j^{th} in ability among n competitors, is

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

Therefore, bid quality generated relates intuitively to prize sizes, ability, the probability of competitors' bids, and effort cost.

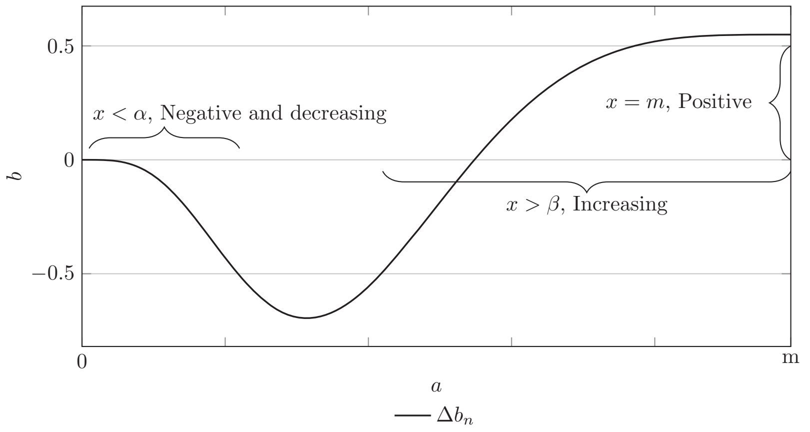
□ **Comparative statics.** Our chief interest, and where we depart from past work, is in examining comparative static implications of the theory. From the equilibrium bid expression (equation (1)), we develop predictions regarding the relationship between the number of competitors and bids across the 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. Proposition 2 is similar to Moldovanu and Sela (2006), which establishes the results on the sign of the effect of competition.

[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.

FIGURE 1

PREDICTED NONMONOTONIC RESPONSE TO COMPETITION



Note: The figure illustrates the response to competition implied by Proposition 2. The change in bid quality and expected performance caused by a change in the number of competitors from n to $n + 1$ plotted by ability. The level of ability α indicates the upper bound for guaranteed negative responses to competition. The level of ability β indicates the lower bound for guaranteed increasing (less negative/more positive) responses to competition. This illustration was produced using: (a) $F(a) = \beta(2, 5)$ over $[0, 0.99]$, (b) $\delta(a)\gamma(b) = (1 - a)b$, (c) $\mathbf{V} = (3, 2, 1)$, and (d) $n \in \{5, 6\}$.

Proposition 2. Let $X_n = \{F_A, \delta, \mathbf{V}, n\}$ and $X_{n+k} = \{F_A, \delta, \mathbf{V}, n + k\}$ be tournaments, differing in their number of competitors by $k > 0$, with bid functions b_n, b_{n+k}, respectively. Let $\Delta_k b_n = b_{n+k} - b_n$ be the difference in bid quality to k additional competitors. The cost of bid quality b is $c(b, a) = \delta(a)b$. Then, on the interval $[0, m]$:

(i) $\Delta_k b_n(0) = 0$,
(ii) $\Delta_k b_n$ decreases on $[0, \alpha]$ where $\alpha = F^{-1}[(\frac{(n-1)!(n+k-p-1)!}{(n-p-1)!(n+k-1)!})^{1/k}]$,
(iii) $\Delta_k b_n$ reaches its minimum in (α, β) where $\beta = F^{-1}[(\frac{n-1}{n+k-1})^{1/k}]$,
(iv) $\Delta_k b_n$ increases on $[\beta, m]$, and
(v) $\Delta_k b_n(m) > 0$.

Proof. See Appendix A.

Therefore, provided there are linear and heterogeneous costs of improving the bid quality by competitors, we predict that the response to increased 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 in Hypothesis 1.

Hypothesis 1. The theoretical model of Section 3 implies the following features of the response to competition:

(a) The response to competition is zero at the origin among lowest skilled competitors.
(b) The response to competition decreases and becomes negative as ability increases, up to a point, α.
(c) The response to competition becomes more positive (less negative) at ability levels above β, and continues to increase with ability level.
(d) The response to competition finally increases to the point of becoming absolutely positive.
(e) The response continues to increase with higher levels of ability until reaching the upper bound of ability, m.

4. Empirical context: algorithm contests at TopCoder, Inc.

■ Data for our study comes from TopCoder, Inc., the leading platform for delivering software solutions through contests. It has routinely delivered sophisticated software projects for Fortune 1000 companies and government agencies since 2001. More than 600,000 solvers have signed up as members of the platform, with tens of thousands from around the world regularly participating in contests.

TopCoder runs contests of a number of types. Here, we study data from its weekly "Algorithm" contests, in which competitors provide computer program solutions to algorithmic computer science problems. TopCoder endeavors to engage and retain members by designing interesting and challenging contest problems. Our interviews with the platform designers and contest organizers indicate that the problems are designed to test a contestant's ability to take problems and convert them into working computer solutions. The problems presented during the contests are representative of the types of computational challenges observed in fields as diverse as computational biology, imaging and graphics, engineering, and finance.

Beyond offering interesting problems, these contests allow the company to determine the skill level of each contest participant, many of whom participate in dozens of such contests over months or years. All participants receive a public, numeric skill rating that reveals their position within the overall skill distribution of programmers on the platform. 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 the history of ranks in past contests.

Within the contests, participants create software solutions to three problems over the course of a 75-minute period. Each problem is preassigned a point value that distinguishes it as "easy," "medium," or "hard." The most common distribution of point values is 250, 500, and 1000. The points received in a contest are the sum of points received for each problem. Solution quality is evaluated through the use of automated tests for each problem, resulting in an quantitative score for each submission. The score is a function of the correctness and speed with which individual solutions are completed after a problem is "opened" by a contestant. In each event, registered competitors, typically numbering several hundred, are assigned to virtual contest "rooms" with an average of 21 (4.56) independent contests held simultaneously during each event. The number in each room is capped at 20 competitors, and typically ranges from 16 to 20 competitors.

Contests occur in two broad divisions, I and II, based on the Elo skill rating obtained from prior participation. Division I consists of competitors who rank above a predetermined rating score; Division II includes newcomers (who have yet to establish a score) and those who rank below the cutoff for a Division I rating.

□ **Information environment.** For competitors, algorithm contests take place through an online, web interface. Competitors can log in to the contest up to 2 hours prior to the start time. Until the start time, competitors wait in the contest "arena," where they can chat with others. Those who have logged in are listed by their TopCoder user names. Prior to the start of a given event, coders do not know the identity or number of other competitors they will face, the number of independent rooms into which the contest will be divided, or the problems they will encounter. For those events featuring cash prizes, the existence of a prize is known prior to event registration. The prize pool per event is approximately $5000. First- and second-place finishers in each independent contest room receive a prize. The first-place prize is higher than the second-place, with precise levels varying across events but fully disclosed at the start of the contest. At the start time, all competitors are randomly placed into competition rooms. Competitors are aware that the room assignments are randomized.

Once placed in a competition room, the 75-minute event clock starts to tick and the contest begins. The three problems are not revealed, but their point value is displayed. Problems may now be individually accessed through the interface. Viewing a problem starts the clock used to measure the time a contestant takes to complete a solution. The point value for a correct solution falls with the time a problem is open.

The interface for the competition puts a great deal of information "at the fingertips" of the competitors; the user names of all competitors in the room are listed and color coded by skill rating. This gives competitors a sense of their relative ability within their specific contest room. Further information is available by clicking on a contestant's user name in the interface. This will open the contestant's TopCoder profile, which provides his or her exact TopCoder skill rating and detailed performance history.

The interface effectively provides a sense of the skill distribution of the contest's competitors. However, the precise skills of a competitor in any given contest will fluctuate in relation to the particular problem being posed, depending on the problem's design and the range and nature of skills it draws upon. Therefore, according to interviews with TopCoder executives and participants, the skill distribution provides only an expected indication of the true distribution of skills—something that may vary considerably from problem to problem. Consistent with the earlier formal characterization in Section 3, competitors do not know the specific true abilities of the competitors they face; rather, they may form beliefs regarding the distribution of skills. A dynamic scoreboard is also provided, indicating which competitors have submitted solutions to which problems, along with provisional scores. The correctness of solutions is not assessed until the end of the contest, so the scores displayed are not final.

□ **Payoffs.** The earlier formal characterization in Section 3 presumed payoffs are descending in rank order. We focus in our empirical analysis on the roughly 20% of contests that have cash prizes. Coders that win a prize receive an average of $110.

Our in-person interviews with TopCoder executives and dozens of competitors uniformly indicate that other sources of payoffs are relevant in this context. Consistent with recent theorizing, TopCoder participants value the symbolic rewards available in competitions (Moldovanu, Sela, and Shi, 2007; Auriol and Renault, 2008; Besley and Ghatak, 2008; Frey and Neckermann, 2008). Competing and doing well provides a credible quality signal, most obvious in the skill rating that reflects relative performance, often used to assess skills by large information technology firms and organizations (e.g., Microsoft, Google, Facebook, NASA) to screen for talent. TopCoder participants value the status, esteem, and peer recognition derived from the rank outcomes, as has been discussed in a range of tournament-like settings in the literature (Azmat and Iriberri, 2010; Blanes i Vidal and Nossol, 2011; Kosfeld and Neckermann, 2011; Brunt, Lerner, and Nicholas, 2012; Delfgaauw et al., 2013).

The existence of nonmonetary payoffs that can be realized over time may raise concerns about using the one-shot, Moldovanu and Sela (2001) model. Incorporating nonmonetary payoffs along with monetary payoffs in the one-shot model is not problematic as long as the overall payoffs remain descending in rank order, that is, $V_1 > V_2 > \ldots > V_p$. Any delay in receiving the payoff, for example, a job market signal, can be taken into account with a present-value calculation. A more serious challenge to our one-shot characterization of these contests would be the future strategic implications of current performance; in particular, of reducing current performance for some later benefit. However, we see no plausible benefit to reducing performance in a contest to obtain future benefits. There are no awards or opportunities for low-rated or low-performing TopCoder members that might entice competitors to "throw" a contest. Thus, payoffs can be broadly understood to decline with rank order in the structure earlier envisioned, that is, $V_1 > V_2 > \ldots > V_p$.

□ **Sample.** Given our econometric approach (Section 5), our interest here is to study a short panel within a 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 represents a period of stable commercial

TABLE 1 Summary Statistics of Estimation Variables

Notation	Theoretical Counterpart	Variable Description	Mean	Standard Deviation	Minimum	Maximum
$Score$	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

growth of the platform, after its initial establishment and a 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 focus our analysis on the higher skilled Division I, rather than the lower skilled Division II, to avoid missing skills measures, avoid cases where individuals are simply participating out of curiosity, assure that the skill measures are somewhat stable and meaningful, and avoid variation in performance that might simply relate to becoming accustomed to the TopCoder format. All together, we have a total of 755 independent contests (rooms) across 32 events in our sample, in which 2775 individual competitors participated. This forms an unbalanced panel of 14,042 observations of competitors within particular contests.

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

□ **Random assignment and sources of variation of key variables.** TopCoder's procedure for setting competition room sizes and assigning competitors to rooms provides quasi-experimental variation in our data set. TopCoder does not set a specific contest room size but chooses the minimum number of rooms such that not more than 20 competitors will be in any room. Hence, the number of competitors in a room is not related to the total number of competitors in an event, but roughly the total number modulus 20.

Although the total number of competitors is likely related to contest characteristics, the total number modulus 20 is not. The modulus operation redistributes fluctuations in a roughly uniform way across 0 to 19.[6] Thus, the number of competitors assigned to a room is plausibly random. Statistical tests support this argument, with the total number of competitors modulus 20 being almost perfectly uniform, and room size being uncorrelated with observed contest characteristics.

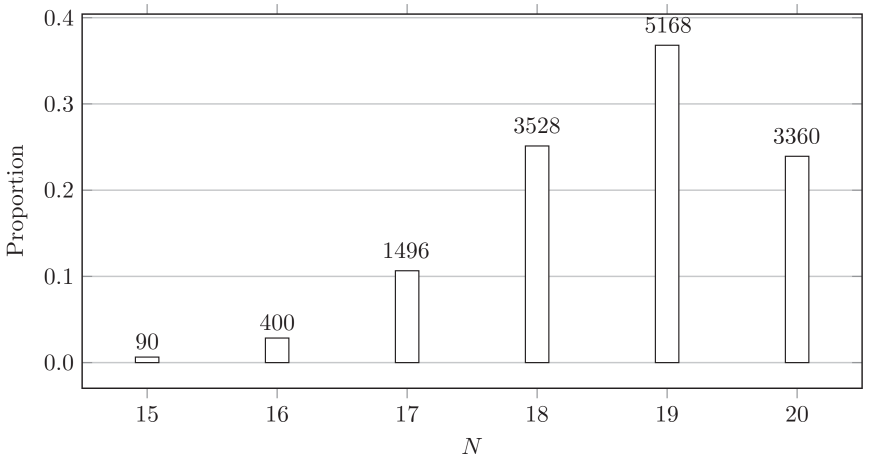
Another source of variation in numbers of competitors is created by dropouts, accounting for 5%–7% percent lower participation over the population of contests. Between the time that registration takes place and the event begins, contests typically experience some degree of drop out. Dropouts are those competitors who signed up and were assigned to a competition room but failed to check in and participate in the contest. These individuals open no problems and take no actions during the competition. Although dropouts occur randomly across independent contest

[6] Note, for example, the Linear Congruent (Pseudo-Random Number) Generator is the most common algorithm underlying random number generators, and relies on the modulus operation.

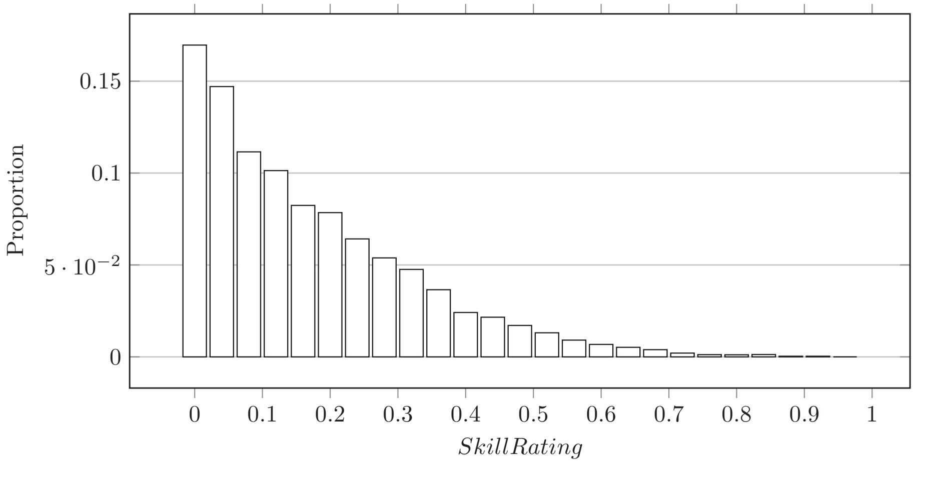
FIGURE 2

DISTRIBUTION OF CONTEST COMPETITORS

(a) Distribution of the number of competitors in contest rooms. The frequency appears at the end of each bar.



(b) Histogram of the distribution of competitor skill ratings.



rooms, a concern with dropouts is if they somehow systematically affect the skill distribution in the contest room (i.e., create a correlation between the skill distribution and the number of competitors). We find no evidence this is the case, as measures of the skill distribution in contest rooms (mean, median, variance, and skew) show no significant correlations to the number of competitors.

Once the number of competitors per room is set, competitors are randomly sorted into competition rooms. TopCoder's random assignment of competitors to rooms ensures that within a given contest, the number of competitors per room is unrelated to competitors' ability and individual characteristics. Hence, our estimates are identified through the exogenous change in the number of competitors in a room. This number varies 33.3% across the entire data set (with the bulk of the data set experiencing a 17.6% variation in the number of competitors). The precise distribution of numbers of competitors is shown above in Figure 2. 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.

5. Flexible, nonparametric estimates

■ Our main analysis begins with nonparametric estimates to test our theoretical predictions of nonmonotonic, quasi-convex responses to competition across the skill distribution (Section 3). After demonstrating results consistent with our theoretical characterization, we then shift to investigating counterfactual simulations of alternative contest design policies with a semiparametric model.

Following our earlier characterization (Section 3), the bid function or expected performance of competitor i in contest t, b_{it}, is a function of: the number of competing competitors, n_{it}; competitor ability, a_{it}; and the distribution of abilities in the field of competitors, $F_{A(it)}$. Here, we measure $\{b, n, a, F_A\}$ with empirical variables $\{Score, N, SkillRating, \overline{SkillRating}\}$.[7]

The conditional mean *Score*, the empirical estimate of the bid function, $b(n, a, F_A)$, can be summarized in the following expression, where $g(\cdot)$ is the empirical function summarizing the relationship among key variables, and ϵ_{it} is an additive zero-mean error term:

$$Score_{it} = g(N_{it}, SkillRating_{it}, \overline{SkillRating}_{it}) + \epsilon_{it}.$$

However, our interest is not so much in the conditional mean *Score*, itself. It is in how competitors' performance responds to added competitors as a function of ability level. In terms of the earlier theoretical discussion, this means an interest in estimating $\Delta_k b_n(a)$, rather than just the bid function, b_n. In terms of our empirical function, this is $g(N + k, \cdot) - g(N, \cdot)$, rather than just $g(\cdot)$. We represent the difference by the expression,

$$\Delta_{n,n+k} g(SkillRating_{it} | \overline{SkillRating}_{it}) = \frac{g(n + k, SkillRating_{it}, \overline{SkillRating}_{it}) - g(n, SkillRating_{it}, \overline{SkillRating}_{it})}{k} + \delta_{it}, \quad (2)$$

where n is a baseline number of competitors, k is an incremental addition to the number of competitors, and the error term is redefined appropriately as δ. To estimate this difference, we execute two steps: (a) we first estimate $g(\cdot)$ using a nonparametric estimator, and then (b) difference the estimated function in the N dimension, dividing by the change in N.

□ **The nonparametric estimator.** We estimate the function $g(\cdot)$ using a nonparametric kernel technique. The approach has two major components: (a) a kernel to assign weights to data points and (b) an estimator to generate $\hat{g}$ using the weighted data.

Our choice of kernel is dictated by two aspects of our data set and model: (a) g is a function of both continuous and discrete variables, and (b) the density of data across the domain varies.

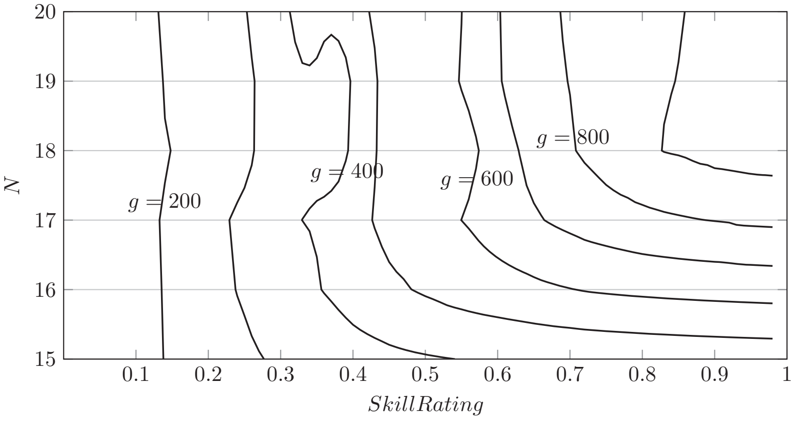
As g is a function of discrete (N) and continuous ($SkillRating$, $\overline{SkillRating}$) variables, we employ a kernel designed for mixed data proposed in Racine and Li (2004).[8] The kernel employs a second-order Epanechnikov kernel to weigh the continuous dimensions of a data point and weighs the discrete dimensions using their ordinal distance, d, by λ^d, where λ is a kernel parameter. The final kernel weights for a data point are the product of these discrete and continuous weights. A "nearest-neighbor" adaptive approach was used to improve the estimate in the presence of widely varying data density. Instead of a fixed bandwidth kernel in which every estimate uses all data points lying in a region of fixed size, the adaptive approach uses a fixed number of data points, m, for every estimate and adjusts the size of the region to match.

We employ a local linear estimator to generate $\hat{g}$ from the kernel-weighted data. The local linear estimates are simply the weighted least-squares estimate, where the weights are provided by the kernel. The local linear estimator is considered more robust along boundaries of the estimation domain, where we are particularly interested in observing effects, as its bias is not a function of the data density (Li and Racine, 2007).

[7] Results do not substantially change when including higher moments of the skill distribution beyond the mean.
[8] See Li and Racine (2007) for a textbook treatment of nonparametric estimation of mixed data.

FIGURE 3

NONPARAMETRIC BID FUNCTION, $g(\cdot)$



Note: Mean skill rating in the room, $\overline{SkillRating}$, is held at its mean value.

Our estimate, $\hat{g}$, is then a function of the data set and the two kernel parameters: (a) m the number of points to include in the estimation and (b) λ the weighting of the ordinal distance along the N dimension. These parameters are selected to minimize the mean squared-integrated error of the estimate, cross validating with a leave-one-out estimator. $\hat{g}$ is calculated at every point in the data set and the difference between the actual data point and $\hat{g}$ is used to calculate the squared error. However, the estimator "leaves out" the data point under consideration when producing the estimate. The values selected are: $m = 68$ and $\lambda = 0.58$.

Drawing inference from $\hat{g}$ about g requires that the estimator is consistent and asymptotically normal. This will be the case if the underlying function g is sufficiently smooth and the data-generating process allows the estimator to converge quickly enough as sample sizes grow toward infinity. The proof of this result can be found in Li and Racine (2007). The sufficient conditions are:

(i) Observations, $(Score, N, SkillRating, \overline{SkillRating})_{it}$, are i.i.d.
(ii) The fourth moment of the errors exists $\mathrm{E}\epsilon_{it}^4 < \infty$.
(iii) The true relationship, g, and the error distribution are three times differentiable.
(iv) $\mathrm{E}(\epsilon_{it}^2 | N_{it}, SkillRating_{it}, \overline{SkillRating}_{it})$ is continuous.
(v) The true, optimal, m satisfies $m \in [n^{\omega}, n^{1-\omega}]$ for some $\omega \in (0, \frac{1}{2})$.

The natural experiment conditions discussed previously in the fifth subsection of Section 4 give us confidence that the observations are i.i.d. We must assume that the smoothness conditions hold; however, they are not onerous. Although the location of the optimal m cannot be proven, the selected value is well inside the interval.

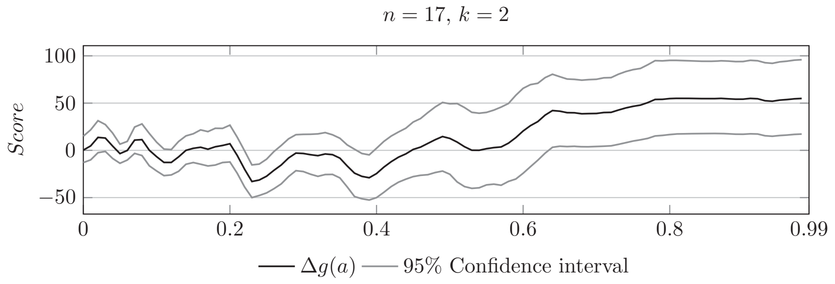
□ **Nonparametric results.** Our interest is in the response to competition as a function of the number of competitors, N, and contestant ability, $SkillRating$. Hence, we calculate $\hat{g}$ for each N in our data set $\{15, 16, \ldots, 20\}$ and along the $SkillRating$ dimension $\{0.01, 0.02, \ldots, 0.99\}$. Figure 3 plots contours of the estimated bid surface.

The estimated bid surface conforms to the Moldovanu and Sela (2001) model. Bids increase with $SkillRating$, increase with N at high ratings, and decrease or hold constant with N at lower ratings. The fitted model also explains a large fraction of variation, with $R^2 = 0.3328$.

In order to draw statistical inferences related to our main research questions, we estimate equation (2), the response to competition. Here, we focus on the interval at plus and minus one

FIGURE 4

FLEXIBLE NONPARAMETRIC ESTIMATION OF PERFORMANCE RESPONSE TO ADDED COMPETITORS



Note: The figure presents the estimated effect of increasing the number of competitors from $N = 17$ to $N = 19$, across *SkillRating*, based on a nonparametric estimator with 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 those theorized in Hypothesis 1.

standard deviation of the mean in numbers of competitors, $N = 17$ and $N = 19$. In relation to the discussion in Section 5, this would imply $n = 17$ and $k = 2$. The interval from $N = 17$ to $N = 19$ crosses the densest portion of the data set and uses representative levels and differences in N; insofar as the mean and standard deviation summarize the data.

Figure 4 presents our estimates of the expected response of *Score* to increases in N, over varying levels of *SkillRating*; along with 95% confidence intervals.[9] Despite the estimate being produced in a flexible manner with a minimum of constraints, the patterns summarized below in items 1 to 5 conform precisely to the earlier theorized Hypothesis 1.[10]

Observation 1. As seen in Figure 4, the response of *Score* to increases in N has the following features:

(i) The response of the lowest skilled competitors is indistinguishable from zero.
(ii) Proceeding rightward to those of intermediate levels of skill, the response to competition becomes increasingly negative.
(iii) Increasing beyond some intermediate level of skill, the response to competition increases (becomes less negative).
(iv) The increase continues until a skill level is reached where the response to competition becomes positive.
(v) The response continues to increase with added skill and the response is most positive at the maximum ability level.

Although we estimated the relationship described in Proposition 2 with data taken across the "middle" of the data set, that is, at standard deviations about the mean, in principle the relationship should hold at any level and change in number of competitors. Therefore, as a robustness test, we reestimate the nonparametric relationship on every possible combination of N and k within our data in addition to the preferred estimate, in Figure 4. With five levels of N, there are 15 possible estimates, that is, $(N = 15, k = 1), (N = 15, k = 2), \ldots (N = 19, k = 1)$. For those intervals over which we find significant results (12 of the 15 possible intervals), we find evidence

[9] 20,000 Pseudo-data sets, of the same size as the original data set, are created by sampling with replacement from the original data. New bid surface estimates are generated on each pseudo-data set. These bid surfaces provide the sample to form the confidence intervals of the estimator.

[10] Note that these results also conform to past findings of a negative average response to added numbers of competitors, as the bulk of competitors appear in the part of the ability domain where the response to competition is negative. Roughly 5% of observations occur in the part of the *SkillRating* domain in which the response is positive.

consistent with a nonmonotonic response to competition, with positive response at higher skills and negative response at lower skills. Specifically, we find such statistically significant results for all differences beginning at $N \in \{15, 16, 17\}$ for all applicable k. Differences beginning at $N \in \{18, 19\}$ (3 of the 15) are substantially and significantly close to zero. In Section 6, we estimate a parameterized version of the Moldovanu and Sela (2001) model providing a point estimate of the model's predicted bid surface. The estimate imposes the model's restrictions simultaneously across the bid surface, and all intervals, allowing comparisons to the nonparametric estimate.

6. Semiparametric estimates

■ In order to analyze more precise predictions of the theory and its implication for welfare (Section 7), we fit a fully parameterized version of the model of Section 3 to the data set, using a semiparametric approach. We combine nonparametric estimates of the ability distribution with a parametric version of the Moldovanu and Sela (2001) model using nonlinear least-squares. Recall, from Section 3, equation (1), the expected performance or bid function takes the following form:

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

For estimation purposes, it is useful to manipulate the prizes algebraically so that we estimate the largest prize V_1 and a series of marginal percentage prizes δv_j which relate differences in payoff magnitudes between ranks, to the absolute magnitude of the largest prize.

$$b(a) = V_1 \sum_{j=1}^{p} \Delta v_j \int_0^a \frac{1}{\delta(z)} \left\{ \frac{\partial P_{j,n}}{\partial a}(z) - \frac{\partial P_{j+1,n}}{\partial a}(z) \right\} dz,$$

$$\text{where } \Delta v_j = \frac{\Delta V_j}{V_1}.$$

In this form, the magnitude of the prizes, V_1, can be easily separated from the distribution of value among the prizes.

In estimating this function, a contestant's ability (a), the number of competitors (n), and the bid ($b(a)$) are modelled by the same variables as in the preceding subsection. The distribution of abilities is estimated by a kernel estimate, $F(z; \mathbf{a}) = K(z; \mathbf{a})$, where $\mathbf{a}$ is the vector of abilities in a contest. Then, $\frac{\partial P_{j,n}}{\partial a}(z)$ is directly calculated for each individual and contest in the data set. The model assumes linear costs in the bids.[11] The marginal cost for a contestant of given ability, δ is set to $\delta(a) = (1 - a)^2$. What then remains is to estimate the series of rank-order payoffs, $\mathbf{V}$, in order to fully specify the bid function and semiparametric model. Given the mix of cash and noncash-based rank-order payoffs enjoyed by competitors (discussed in the second subsection of Section 4), we allow for a relatively large number of rank-order payoffs beyond just the two top ranks corresponding to cash prize winners. We allow for rank-order payoffs to those finishing at least 10th, or $p = 10$.[12]

The prize components V_1 and $\{\Delta v_j\}_{j=1}^{10}$ are estimated from the data using nonlinear least-squares. We deviate from the model somewhat by allowing the bid function to have a nonzero intercept, α. A nonzero intercept implies that individuals are willing to exert some effort even if they will not win a prize. Nonlinear least-squares estimates solve the problem[13]:

[11] We also estimated models with nonlinear costs: power function and cubic splines. However, the added complexity did not improve model fit, and it greatly increased estimation complexity.

[12] The estimates are not sensitive to allowing for more prizes.

[13] In order to conform to the theoretical assumption that $\Delta v_j > 0$, we constrain the optimization to $\Delta v_j \geq 10^{-8}$.

$$\min_{V_1,\, \Delta v_j} \sum_{i=1}^{n} \left[b_i - V_1 \sum_{j=1}^{p} \Delta v_j \int_0^a \frac{1}{(1-z)^2} \left\{ \frac{\partial P_{j,n}}{\partial a}(z) - \frac{\partial P_{j+1,n}}{\partial a}(z) \right\} dz - \alpha \right]^2$$

$$\text{s.t. } V_1 > 0,$$
$$\Delta v_j > 0 \forall j \in \{1, 2, \ldots, 10\},$$
$$\sum_{j=1}^{10} \Delta v_j = 1.$$

As the parameter estimates depend on nonparametric estimates of the ability distribution, the consistency and asymptotic normality of the estimator is not immediate. However, the estimator belongs to the Andrews (1994) MINPIN class of estimators (They are estimators that MINimize a criterion function that may depend on a Preliminary Infinite dimensional Nuisance parameter estimator), whose consistency and asymptotic normality can be shown by application of his theoretical result; a version of the theorem can be found in Li and Racine (2007). Roughly speaking, we require that (a) the nonparametric estimation of the ability distribution be consistent and asymptotically normal, (b) the nonlinear least-squares estimator be consistent, and (c) that the first-order conditions of the minimization problem be *stochastically equicontinuous* at the true values of V_1, Δv_j, and α. The consistency and asymptotic normality of the ability distribution is assured if

(i) the true distribution of ability, F_A, is twice continuously differentiable and $\frac{\partial^2 F}{\partial x^2}$ is Hölder-continuous, and
(ii) $\lim_{n\to\infty} h \leq Cn^{\frac{1}{8}}$, where C is a constant and h is the selected bandwidth.

As with the earlier nonparametric estimate, we must assume that the smoothness conditions hold; however, they are not onerous. The bandwidth is selected by the rule-of-thumb $h = \hat{\sigma}\frac{4}{3}n^{-\frac{1}{5}}$, which satisfies the limit condition.

The consistency of the nonlinear least-squares estimator is straightforward, as it is a constrained ordinary least-squares estimate. Note, that our parameters V_1 and Δv_j enter the objective function linearly. Hence, we require that the weighting matrix whose elements are composed of $\int_0^a \frac{1}{(1-z)^2} \left\{ \frac{\partial P_{j,n}}{\partial a}(z) - \frac{\partial P_{j+1,n}}{\partial a}(z) \right\} dz$ must be full rank, and that bidding errors be conditionally exogenous.

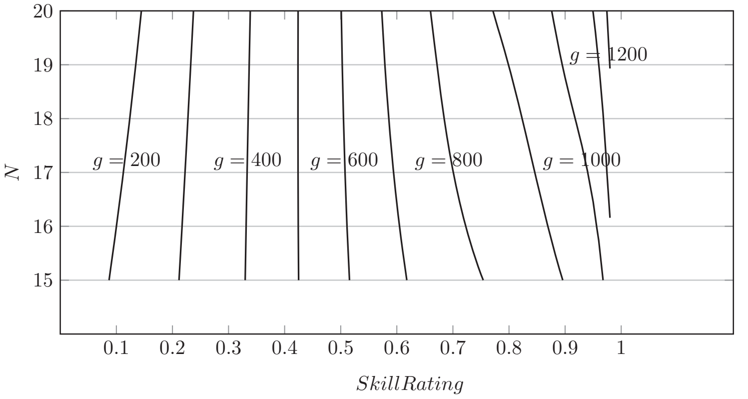
Establishing stochastic equicontinuity is somewhat involved. Roughly speaking, given the true parameter values, the weighting matrix must converge to its true value. This holds because errors in the ability distribution propagate "nicely" through the weighting matrix, so that as the ability distribution converges, the weighting matrix also converges.

□ **Semiparametric results.** Figure 5 plots contours of the estimated bid function in the n and ability domain, using the distribution of abilities pooled across all contests. Remarkably, the estimated model has an $R^2 = 0.3326$; within 1% of the nonparametric model. The 95% confidence intervals of the parametric and the nonparametric estimates overlap over 67% of the domain; 92% if weighted by the number of observations at each n.[14] Figure 6 provides a more detailed view of the bid functions across abilities at $n = 17$ and $n = 19$. Even though the estimate is constrained by the model and our parametric assumptions, it is able to fit the data as well as the best-fitting smooth function. Therefore, the fitted structural model following the Moldovanu and Sela (2001) framework outlined in Section 3 performs remarkably well in describing the data.

[14] 20,000 Pseudo-data sets, of the same size as the original data set, are created by sampling with replacement from the original data. New bid surface estimates are generated on each pseudo-data set. These bid surfaces provide the sample to form the confidence intervals of the estimator.

FIGURE 5

PARAMETRIC BID FUNCTION, $b(\cdot)$



Note: Mean skill rating in the room, $\overline{SkillRating}$, is held at its mean value.

FIGURE 6

SEMIPARAMETRIC AND NONPARAMETRIC ESTIMATION OF EXPECTED SCORE

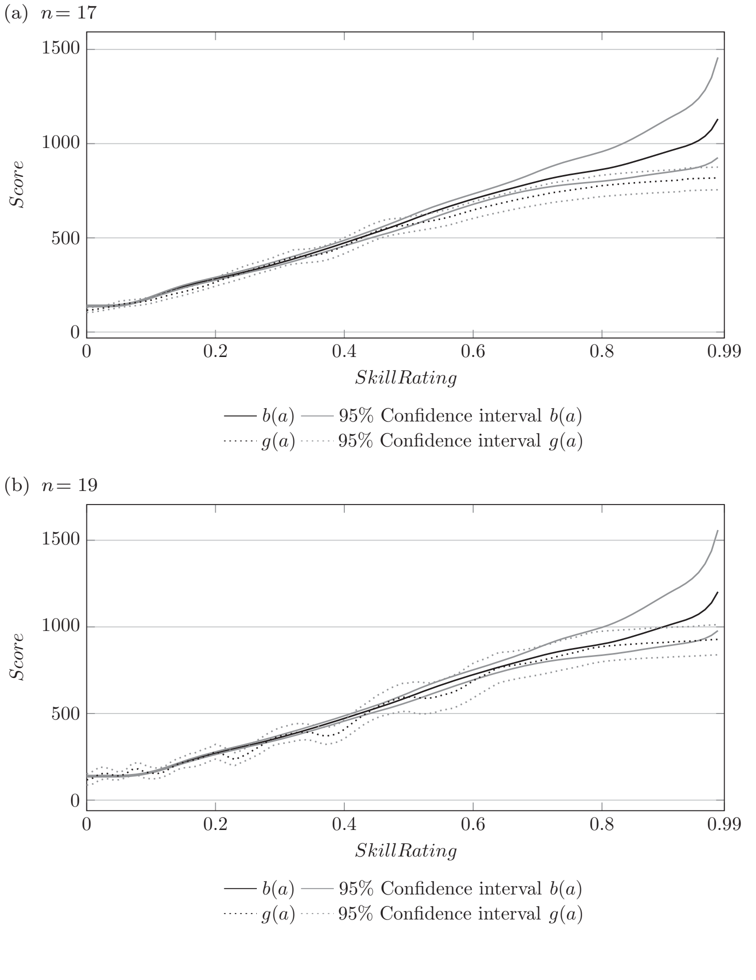


TABLE 2 Nonlinear Regression Estimates of Model Parameters

Parameter	Estimate	95% Confidence Interval
α Intercept	138.093	[134.443, 142.344]
V_1	281.003	[270.949, 293.040]
ΔV_1	0.172	[0.075, 0.323]
ΔV_2	0.230	[0.000, 0.377]
ΔV_3	0.000	[0.000, 0.000]
ΔV_4	0.000	[0.000, 0.281]
ΔV_5	0.228	[0.000, 0.287]
ΔV_6	0.000	[0.000, 0.222]
ΔV_7	0.000	[0.000, 0.000]
ΔV_8	0.000	[0.000, 0.000]
ΔV_9	0.000	[0.000, 0.000]
ΔV_{10}	0.370	[0.322, 0.396]

The estimates for the prize components with bootstrapped 95% confidence intervals are shown in Table 2. The intercept is significantly positive with a mean value of $\alpha = 138$. The maximum prize is estimated to be $\hat{V}_1 = 281$. The estimated prize distribution is significantly positive at 4 positions: 1, 2, 5, and 10. The share of the maximum prize at these positions is: 0.17, 0.23, 0.23, and 0.37, respectively. A positive intercept suggests that competitors are willing to expend some effort without a prize incentive. The presence of significant prizes beyond the first two ranks suggests that nonmonetary incentives are significant.

7. Interpretation of policies and counterfactuals

■ The high fidelity of the semiparametric model with nonparametric estimates suggests this model can aid in interpreting policies and contest design parameters, issues relevant to theorists and tournament organizers. The policies of TopCoder might be of particular interest given the success of their platform—attracting more than a half million contestants and servicing a large roster of clients.

□ **Setting contest size.** A first policy we consider is regulating the number of competitors who face one another. The number of competitors in each contest in the data set varies among the high teens and does not exceed 20. This follows TopCoder's policy of creating new contest "rooms" when there are sufficient competitors registering, rather than 20 competitors per room.

As Proposition 2 implies, increasing the number of competitors will increase maximum scores and decrease the scores of lower-ability competitors. The semiparametric estimates allow us to make predictions about the *magnitude* of these changes in order to make optimal decisions. To examine the impact on scores as the level of competition changes (holding the ability distribution constant), Figure 7 plots the maximum and mean scores for $n \in \{14, \ldots, 25\}$. Note that the near linearity of effects over the plotted interval is itself a finding of the simulation; N does not enter the bid function linearly. As the room size increases from 14 to 25, the mean score falls by 41 points ([−42, −39]), whereas the maximum score increases by 386 points ([284, 552]).

Given these results, it appears that maximum scores are about 10 times as responsive to competition as the mean score. The value of this trade-off depends on the goal of the designer. However, the relative magnitude of the changes suggests that the number of competitors is a reasonable tool for adjusting the performance of the highest-ability competitors.

The number of competitors in a contest room is not only interesting for its impact on scores. Current TopCoder policy sets prizes at the room level. Hence, adjusting room size impacts not only the number of competitors, but also total expenditures by TopCoder. Assuming total attendance is not affected by the change, larger contest rooms means average expenditures per competitor will fall, because larger rooms will accommodate the same number of competitors with fewer rooms

FIGURE 7

MEAN AND MAXIMUM SCORES ACROSS ROOM SIZES

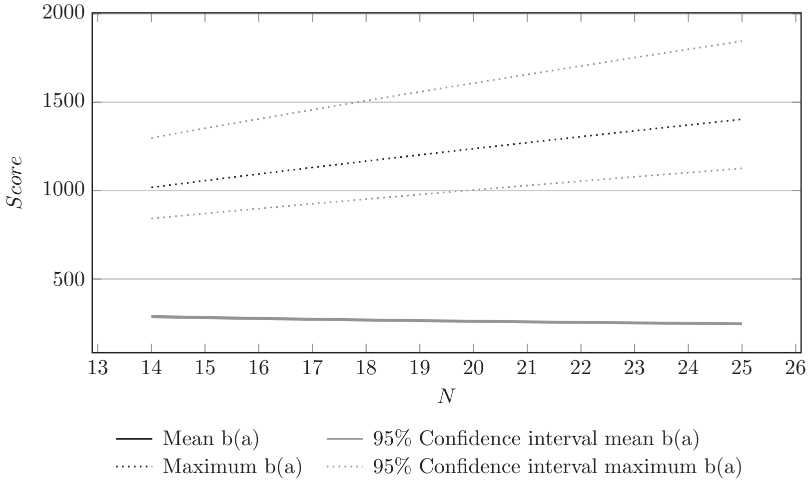
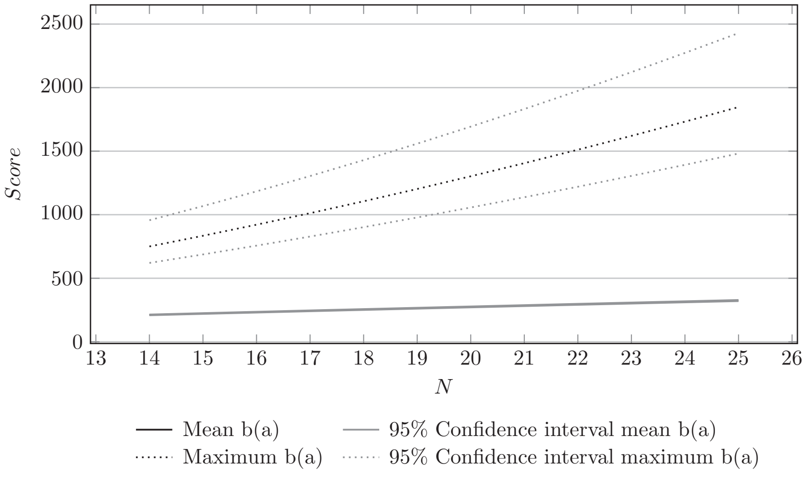


FIGURE 8

MEAN AND MAXIMUM SCORES ACROSS DIFFERENT ROOM SIZES WITH PROPORTIONAL CHANGES IN PRIZE AMOUNTS



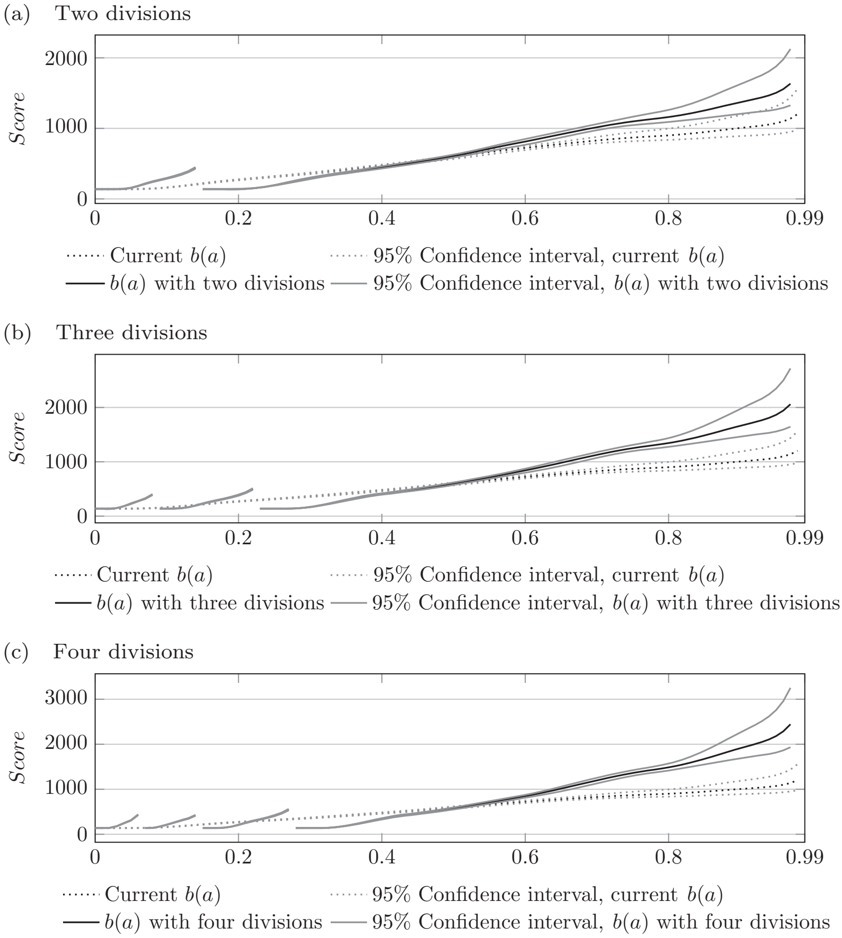
and prizes. For example, the change from 14 to 25 competitors studied here is accompanied by a 56% decrease in prize expenditures.

Repeating the earlier analysis of room size, Figure 8 plots the maximum and mean scores for $n \in \{14, \ldots, 25\}$ with the total prize amount adjusted proportionally to room size. As the room size increases from 14 to 25, the mean score increases by 113 points ([111, 115]), rather than falling 41 points as in the earlier analysis. The expected max score increases by 1096 points ([860, 1480]), rather than just 386 points. In essence, the savings from increasing competition can be used to blunt or reverse the negative impact on the mean score.

□ **Dividing competitors by ability into separate contest divisions.** Another design variable available for manipulation is segmentation of tournaments by ability. TopCoder's policy is to divide its body of competitors into two pools of competitors, roughly equal in size, according to

FIGURE 9

DISTRIBUTION OF SCORES UNDER A LARGER NUMBER OF DIVISIONS



an ability level cutoff.[15] The impact of segmentation is to increase homogeneity in competitor ability and change the quantiles of the ability distribution each competitor faces. *A priori*, it is not clear that this policy is optimal. If two divisions are superior to one, why not offer three or four divisions? We consider these choices by simulating the division of our observed competitors into additional divisions by ability, with each division containing equal numbers of competitors.

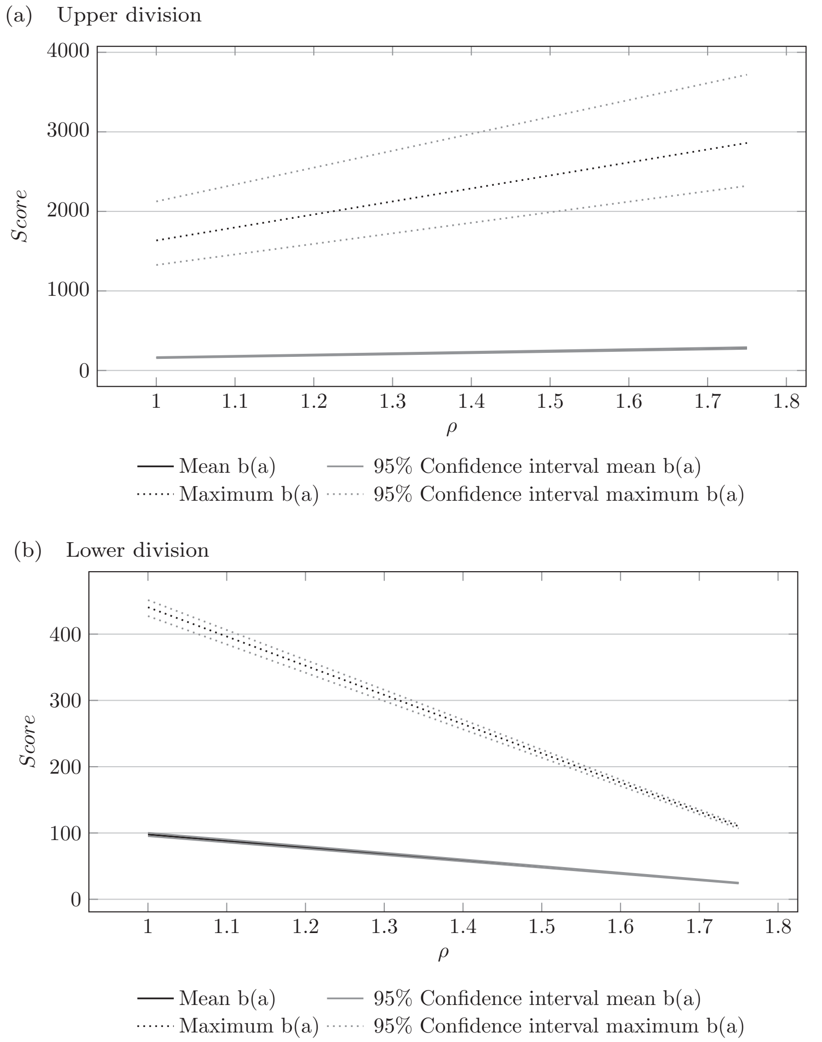
Figure 9 plots effects of having a single upper division, versus dividing competitors into two, three, or four divisions. As can be observed in the figure, increasing segmentation boosts performance near the top of newly created divisions, given the greater potential to win a prize. However, performance drops around newly created bottom ends, as there is now less competition from below.

Increasing subdivision of competitors leads the lower ability divisions to show moderate changes in scores, with the largest change a 268-point increase for the highest-ability competitors in the lowest quartile division (when moving from one to four divisions). The high-ability division shows larger score increases among the highest ability competitors of about 1500 points in the highest quartile (when moving from one to four divisions). The moderate changes in performance for those of moderate ability suggests that TopCoder's current divisions may not be far from

[15] Our analysis of the data here focused on the highly skilled division as discussed in Section 4.

FIGURE 10

MEAN AND MAXIMUM SCORES ACROSS DIFFERENT SHARES OF TOTAL PRIZES

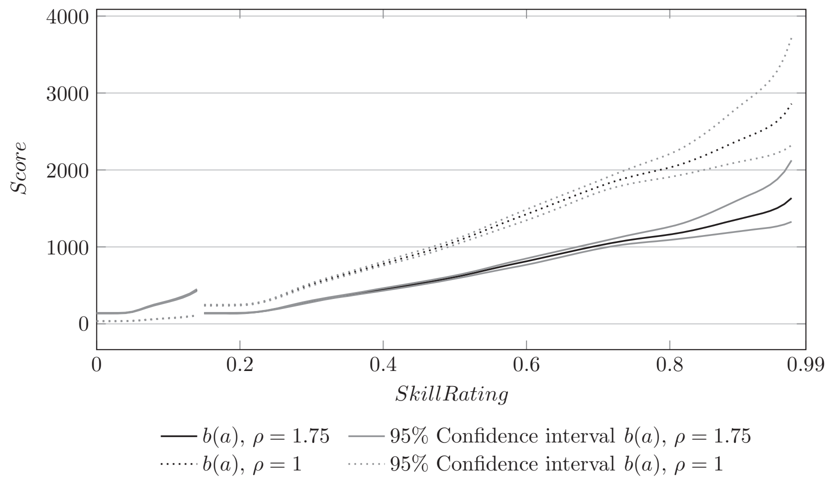


optimal if their primary concern is to assure the engagement of a range of moderate-ability coders as well as engaging top competitors.

Another consideration is that adding divisions allows for further manipulation of the prize distribution. For example, rather than maintaining the same prize level across divisions, we might choose to allocate a greater share of the prize to higher-ability divisions. We simulate two divisions, equal in numbers of competitors, and vary the share of prize expenditure allocated to each division. Figure 10 shows the mean and max scores in the two hypothetical divisions, as the share of prizes varies from equal proportions, to the higher-ability (upper) division receiving seven times the prize money of the lower-ability (lower) division. As can be seen, mean scores in both divisions are not highly responsive to varying levels of prizes. In the upper division, mean scores increase by about 121 points ([119, 123]); in the lower division, mean scores fall by about 73 points ([–74,–72]). The maximum scores are more responsive to the allocation of the prize pool. In the upper division, maximum scores increase by about 1226 points ([994, 1603]). In the lower division, maximum scores fall considerably less, by about 330 points ([–338, –321]). In a

FIGURE 11

SCORES IN UPPER AND LOWER DIVISIONS AT CURRENT PRIZE LEVELS AND AT $\rho = 1.75$



sense, finer divisions and tilting the prize pool toward higher-ability divisions provide an increase in maximum scores "for free." By contrast, generating an increase in expected maximum score by 1226 points with only prize amounts would require an increase in prizes of more than 66%.

□ **"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. There may be any number of reasons for the company to pursue this inclusive approach, what is clear from the results is that there would appear to be little downside to open admissions based on the issues studied here. First, the number of competitors in a given independent contest is fixed by capping the number of participants in any one independent contest room. This departs from many historical cases of contests in which competitors are not cordoned into separate independent contests, but are instead invited to compete in some overall challenge (Brunt, Lerner, and Nicholas, 2012). A possible worry, of course, is that the platform becomes flooded with low-quality participants, which could alter the distribution of abilities of participants in ways that might lessen rivalry among the most able competitors, in addition to other possible problems. However, to the extent this could plausibly become a problem (and in our interactions with the company we found no suggestion that it was), the two-division structure would likely deal with this contingency in a simple fashion. The creation of an upper skill division with a minimum skill threshold effectively fixes the distribution of abilities, F_A, in that division, a virtual form of certifying higher-quality contestants.

8. Summary and conclusions

■ This article analyzes how the number of competitors in a tournament affects the performance of individual contestants. We argue that the incentive response and performance of contestants should be a nonlinear function of the contestants' ability, under relatively general conditions describing a one-shot tournament. The response should be near zero for the lowest-ability competitors, become increasingly negative over moderate-ability competitors, then become less negative, and eventually turn positive, for the high-ability competitors. Therefore, although aggregate and average patterns of performance and effort may decline with increased competition, performance may in fact increase among the highest skilled contestants. These patterns essentially emerge as a result of two countervailing effects. First, increased competition reduces the likelihood of winning a prize, which reduces strategic incentives to exert high effort. Second, added rivalry is likely to induce higher performance given that failing to exert effort will reduce the likelihood of winning a prize.

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 the number of competitors.

Our main contribution is in studying fine-grained evidence on individual competitor outcomes from 755 software algorithm development contests. We exploit natural experimental variation (due to the random assignment procedure employed by the contest sponsor) to identify causal effects. Equally important, this context offers a rare opportunity to observe precise measures of individual competitor ability and performance outcomes based on objective observational measures. The performance response to competition by ability 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 the least-skilled contestants are negligibly affected by rising competition. In addition, with higher levels of ability, the response becomes progressively more negative until, toward the range of the highest skilled contestants, the relationship becomes more positive. The response to competition finally turns positive for contestants of the highest ability—creating something of an asymmetric U-shaped curve (with the right-hand side higher than the left). Therefore, the flexibly estimated relationship conforms to the particular theoretical predictions of the shape.

In order to precisely evaluate how well the Moldovanu and Sela (2001) model can match competitor behavior, we fit a semiparametric version of the Moldovanu and Sela (2001) model to the data. The model fits the data as well as the nonparametric estimate with R^2's within 1% of each other. We use the model estimated to simulate counterfactual situations related to several key contest design policies: setting contest size, dividing competitors by ability, and opening entry to all. Each provides a useful "lever" for tournament organizers to manipulate the performance of competitors, particularly when changes allow the reallocation of prize money. The policies investigated here have a much broader applicability beyond our particular empirical context, as tournament-style compensation is pervasive in the economy and is utilized by online platforms that routinely engage thousands of workers. These platforms solve innovation challenges in a range of settings, including scientific problems (Jeppesen and Lakhani, 2010), algorithm development (Boudreau, Lacetera, and Lakhani, 2011; Lakhani et al., 2013), graphic design (Wooten and Ulrich, 2011), logo development, translation services (Liu et al., forthcoming), and new product development (Poetz and Schreier, 2012). Moreover, industrial firms like Astra Zeneca, General Electric, Procter and Gamble, and Siemens, among others, have set up their own custom platforms and are offering tournament-style compensation to the general public to aide in solving their own internal research and development problems (Boudreau and Lakhani, 2013).

Appendix A

In this section, we prove Proposition 2, repeated here for convenience.

Proposition. Let $X_n = \{F_A, \delta, \mathbf{V}, n\}$ and $X_{n+k} = \{F_A, \delta, \mathbf{V}, n+k\}$ be tournaments, differing in their number of competitors by $k > 0$, with bid functions b_n, b_{n+k}, respectively. Let $\Delta_k b_n = b_{n+k} - b_n$ be the difference in bid quality to k additional competitors. The cost of bid quality b is $c(b, a) = \delta(a)b$. Then, on the interval $[0, m]$:

(i) $\Delta_k b_n(0) = 0$,
(ii) $\Delta_k b_n$ decreases on $[0, \alpha]$ where $\alpha = F^{-1}[(\frac{(n-1)!(n+k-p-1)!}{(n-p-1)!(n+k-1)!})^{1/k}]$,
(iii) $\Delta_k b_n$ reaches its minimum in (α, β) where $\beta = F^{-1}[(\frac{n-1}{n+k-1})^{1/k}]$,
(iv) $\Delta_k b_n$ increases on $[\beta, m]$, and
(v) $\Delta_k b_n(m) > 0$.

Proof. First, Properties 1 and 5 are established. Properties 2 to 4 follow after algebraic manipulation of the definition.

Property 1 follows immediately from the definition. Property 5 follows from the properties of order statistics. Note that $b(a)$ is a linear combination of expectations

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

$$b(a) = \sum_{j=1}^{p} \Delta V_j \mathrm{E}_{F_{n-1}^{n-j}} \left(\frac{1}{\delta(z)} \middle| z < a \right).$$

Hence,

$$\Delta_k b(a) = \sum_{j=1}^{p} \Delta V_j \left[\mathrm{E}_{F_{n+k-1}^{n+k-j}} \left(\frac{1}{\delta(z)} \middle| z < a \right) - \mathrm{E}_{F_{n-1}^{n-j}} \left(\frac{1}{\delta(z)} \middle| z < a \right) \right].$$

Then, at $a = m$,

$$\Delta_k b(m) = \sum_{j=1}^{p} \Delta V_j \left[\mathrm{E}_{F_{n+k-1}^{n+k-j}} \left(\frac{1}{\delta(z)} \right) - \mathrm{E}_{F_{n-1}^{n-j}} \left(\frac{1}{\delta(z)} \right) \right].$$

$F_{n+k-1}^{n+k-j} \leq F_{n-1}^{n-j}$ by Moldovanu and Sela (2006). Further, by assumption, $\frac{1}{\delta(z)}$ is strictly increasing. Hence, by Moldovanu and Sela (2006),

$$\mathrm{E}_{F_{n+k-1}^{n+k-j}} \left(\frac{1}{\delta(z)} \right) > \mathrm{E}_{F_{n-1}^{n-j}} \left(\frac{1}{\delta(z)} \right),$$

$$\Rightarrow \sum_{j=1}^{p} \Delta V_j \left[\mathrm{E}_{F_{n+k-1}^{n+k-j}} \left(\frac{1}{\delta(z)} \right) - \mathrm{E}_{F_{n-1}^{n-j}} \left(\frac{1}{\delta(z)} \right) \right] > 0.$$

Thus, Properties 5 holds.

In order to establish Properties 2 to 4, we need to consider the definitions of the order statistic distributions. It is well known that

$$f_m^i(a) = \frac{m!}{(i-1)!(m-i)!} F(a)^{i-1} (1 - F(a))^{m-i} f(a).$$

Dropping arguments for clarity,

$$\Delta_k b(a) = \sum_{j=1}^{p} \Delta V_j \int_0^a \frac{1}{\delta} \left[f_{n+k-1}^{n+k-j} - f_{n-1}^{n-j} \right] dz$$

$$\Delta_k b(a) = \sum_{j=1}^{p} \frac{(n+k-1)! \Delta V_j}{(n+k-j-1)!(j-1)!} \int_0^a \frac{1}{\delta} F^{n-j-1} (1-F)^{j-1} f \left[F^k - \frac{(n-1)!(n+k-j-1)!}{(n-j-1)!(n+k-1)!} \right] dz.$$

It follows that

$$\frac{\partial \Delta_k b}{\partial a}(a) = \frac{(n+k-1)! f(a)}{\delta(a)} \sum_{j=1}^{p} \frac{F(a)^{n-j-1} (1-F(a))^{j-1} \Delta V_j}{(n+k-j-1)!(j-1)!} \left[F(a)^k - \frac{(n-1)!(n+k-j-1)!}{(n-j-1)!(n+k-1)!} \right].$$

The bracketed term determines the sign of the derivative. Note that $F(a)^k$ is strictly increasing.

We now show that $\frac{(n-1)!(n+k-j-1)!}{(n-j-1)!(n+k-1)!}$ is decreasing in j. Comparing the j^{th} and $(j+1)^{th}$ terms,

$$\begin{aligned} \frac{(n-1)!(n+k-j-1)!}{(n-j-1)!(n+k-1)!} - \frac{(n-1)!(n+k-j-2)!}{(n-j-2)!(n+k-1)!} &> 0 \\ \Leftrightarrow \frac{(n+k-j-1)!}{(n-j-1)!} - \frac{(n+k-j-2)!}{(n-j-2)!} &> 0 \\ \Leftrightarrow (n+k-j-1) - (n-j-1) &> 0 \\ \Leftrightarrow k &> 0. \end{aligned}$$

As $k \geq 1$ by definition, $\frac{(n-1)!(n+k-j-1)!}{(n-j-1)!(n+k-1)!}$ is decreasing in j.

Hence,

$$a > F^{-1} \left[\left(\frac{n-1}{n+k-1} \right)^{1/k} \right] \Rightarrow \frac{\partial \Delta_k b}{\partial a}(a) > 0$$

and

$$a < F^{-1}\left[\left(\frac{(n-1)!(n+k-p-1)!}{(n-p-1)!(n+k-1)!}\right)^{1/k}\right] \Rightarrow \frac{\partial \Delta_k b}{\partial a}(a) < 0.$$

Therefore, Properties 2 and 4 hold. These together imply Property 3, and the proposition holds. □

Appendix B

Here, we describe the main elements of the TopCoder rating system. The formula for a coder's rating is:

$$\text{NewRating} = \frac{\text{OldRating} + \text{Weight} \cdot \text{PerfAs}}{1 + \text{Weight}}.$$

A coder's rating is updated at the end of each contest to produce NewRating.[16] OldRating is the coder's precontest rating. If a coder has never competed in a TopCoder algorithm contest, TopCoder assigns a value of 1200 to OldRating.

Rearranging terms, based on the formula for PerfAs below, yields:

$$\text{NewRating} = \text{OldRating} + \left(\frac{\text{Weight}}{1 + \text{Weight}}\right) \cdot \text{CF} \cdot (\text{APerf} - \text{EPerf}).$$

PerfAs is the provisional rating assigned to each coder at the end of a contest.

$$\text{PerfAs} = \text{OldRating} + \text{CF} \cdot (\text{APerf} - \text{EPerf}).$$

APerf is the coder's rank-order performance in the contest, calculated as a value in an inverse standard normal distribution that adjusts for the number of coders per contest:

$$\text{APerf} = -\Phi\left[\frac{\text{ARank} - 0.5}{\text{NumCoders}}\right],$$

where ARank is the coder's rank in a contest, based on total points per coder, and NumCoders is the number of coders in the contest.

EPerf is the predicted value of APerf, based on the coder's precontest rating relative to the precontest ratings of other competitors:

$$\text{EPerf} = -\Phi\left[\frac{\text{ERank} - 0.5}{\text{NumCoders}}\right],$$

where $\text{ERank} = 0.5 + \sum_i WP_i$. WP_i, or Win Probability, is the probability that the coder will have a higher score than another coder i in the contest. Each Win Probability is calculated based on the precontest ratings of coders that entered the contest, adjusted for a measure of the spread of each coder's prior contest ratings, termed Volatility. Coders that have never competed before receive an initial value of 300 for Volatility.

In the formula for PerfAs, CF denotes a "Competition Factor" for each contest. CF captures the spread of the precontest ratings of coders in the contest, based on precontest Volatilities of the competitors and a measure of the difference between the average precontest rating of competitors and individual coder precontest ratings. A greater spread of precontest ratings results in a higher competition factor, leading to a higher weight on the difference between a coder's actual and anticipated performance. Intuitively, changes in rank-order performance in a contest where coders have similar abilities, as measured by precontest ratings, are more likely to reflect random factors rather than skill, and therefore receive lower weight in calculating the new rating. Finally, in the formula for NewRating, Weight for each coder is an inverse function of the number of times that the coder has been rated previously. More experienced coders have less weight attached to the difference between their current rank-order performance, APerf, and their predicted rank-order performance, as reflected in EPerf. In addition, a coder's NewRating cannot exceed his or her OldRating by more than a set value, termed Cap, which is an inverse function of the number of times that a coder has been rated. The values of Weight and Cap insure that the ratings of more experienced coders change less over time than do the ratings of less experienced coders.

References

Aghion, P., Bloom, N., Blundell, R., Griffith, R., and Howitt, P. "Competition and Innovation: An Inverted-u Relationship." *Quarterly Journal of Economics*, Vol. 120 (2005), pp. 701–728.

Andrews, D.W. "Asymptotics for Semiparametric Econometric Models via Stochastic Equicontinuity." *Econometrica*, Vol. 62 (1994), pp. 43–72.

[16] On the TopCoder website, in the explanation of the rating system, the variable Rating is sometimes used in place of what we term OldRating. We use NewRating and OldRating for clarity.

AUGENBLICK, N. "Consumer and Producer Behavior in the Market for Penny Auctions: A Theoretical and Empirical Analysis." Personal Working Paper, 2014.

AURIOL, E. AND RENAULT, R. "Status and Incentives." *RAND Journal of Economics*, Vol. 39 (2008), pp. 305–326.

AZMAT, G. AND IRIBERRI, N. "The Importance of Relative Performance Feedback Information: Evidence from a Natural Experiment Using High School Students." *Journal of Public Economics*, Vol. 94 (2010), pp. 435–452.

BAJARI, P. AND HORTAÇSU, A. "Are Structural Estimates of Auction Models Reasonable?" Working Paper no. 9889, National Bureau of Economic Research, 2003a.

———. "The Winner's Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions." *RAND Journal of Economics*, Vol. 34 (2003b), pp. 329–355.

BERSHTEYN, B. AND ROEKEL, S.V., eds. *Memorandum for General Counsels and Chief Information Officers for Executive Departments and Agencies*. Executive Office of the President, 2011.

BESLEY, T. AND GHATAK, M. "Status Incentives." *American Economic Review*, Vol. 98 (2008), pp. 206–211.

BLANES I VIDAL, J. AND NOSSOL, M. "Tournaments without Prizes: Evidence from Personnel Records." *Management Science*, Vol. 57 (2011), pp. 1721–1736.

BOUDREAU, K.J., LACETERA, N., AND LAKHANI, K.R. "Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis." *Management Science*, Vol. 57 (2011), pp. 843–863.

BOUDREAU, K.J. AND LAKHANI, K.R. "Using the Crowd as an Innovation Partner." *Harvard Business Review*, Vol. 91 (2013), pp. 61–69.

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., AND NICHOLAS, T. "Inducement Prizes and Innovation." *Journal of Industrial Economics*, Vol. 60 (2012), pp. 657–696.

BULL, C., SCHOTTER, A., AND WEIGELT, K. "Tournaments and Piece Rates: An Experimental Study." *Journal of Political Economy*, Vol. 95 (1987), pp. 1–33.

CARPENTER, J., MATTHEWS, P.H., AND SCHIRM, J. "Tournaments and Office Politics: Evidence from a Real Effort Experiment." *American Economic Review*, Vol. 100 (2010), pp. 504–517.

CASAS-ARCE, P. AND MARTÍNEZ-JEREZ, F.A. "Relative Performance Compensation, Contests, and Dynamic Incentives." *Management Science*, Vol. 55 (2009), pp. 1306–1320.

CHE, Y.-K. AND GALE, I. "Optimal Design of Research Contests." *American Economic Review*, Vol. 93 (2003), pp. 646–671.

DAVIS, D.D. AND REILLY, R.J. "Do Too Many Cooks Always Spoil the Stew? An Experimental Analysis of Rent-Seeking and the Role of a Strategic Buyer." *Public Choice*, Vol. 95 (1998), pp. 89–115.

DECHENAUX, E., KOVENOCK, D., AND SHEREMETA, R.M. "A Survey of Experimental Research on Contests, All-Pay Auctions and Tournaments." WZB Discussion Paper no. SP II 2012-109, 2012.

DELFGAAUW, J., DUR, R., SOL, J., AND VERBEKE, W. "Tournament Incentives in the Field: Gender Differences in the Workplace." *Journal of Labor Economics*, Vol. 31 (2013), pp. 305–326.

DOHMEN, T., FALK, A., HUFFMAN, D., SUNDE, U., SCHUPP, J., AND WAGNER, G.G. "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences." *Journal of the European Economic Association*, Vol. 9 (2011), pp. 522–550.

EHRENBERG, R.G. AND BOGNANNO, M.L. "Do Tournaments Have Incentive Effects?" *Journal of Political Economy*, Vol. 98 (1990a), pp. 1307–1324.

———. "The Incentive Effects of Tournaments Revisited: Evidence from the European PGA Tour." *Industrial and Labor Relations Review*, Vol. 43 (1990b), pp. 74–88.

ELYAKIME, B., LAFFONT, J.-J., LOISEL, P., AND VUONG, Q. "First-Price Sealed-Bid Auctions with Secret Reservation Prices." *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, Vol. 34t (1994), pp. 115–141.

FREY, B.S. AND NECKERMANN, S. "Awards." *Journal of Psychology*, Vol. 216 (2008), pp. 198–208.

FU, Q. AND LU, J. "The Optimal Multi-Stage Contest." *Economic Theory*, Vol. 51 (2012), pp. 351–382.

FULLERTON, R.L. AND MCAFEE, R.P. "Auctioning Entry into Tournaments." *Journal of Political Economy*, Vol. 107 (1999), pp. 573–605.

GARCIA, S.M. AND TOR, A. "The N-Effect: More Competitors, Less Competition." *Psychological Science*, Vol. 20 (2009), pp. 871–877.

HARRIS, C. AND VICKERS, J. "Racing with Uncertainty." *Review of Economic Studies*, Vol. 54 (1987), pp. 1–21.

HENDRICKS, K. AND PORTER, R.H. "An Empirical Study of an Auction with Asymmetric Information." *American Economic Review*, Vol. 78 (1988), pp. 865–883.

HENDRICKS, K., PORTER, R.H., AND WILSON, C.A. "Auctions for Oil and Gas Leases with an Informed Bidder and a Random Reservation Price." *Econometrica*, Vol. 62 (1994), pp. 1415–1444.

HOLMSTROM, B. AND MILGROM, P. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization*, Vol. 7 (1991), pp. 24–52.

JEPPESEN, L.B. AND LAKHANI, K.R. "Marginality and Problem-Solving Effectiveness in Broadcast Search." *Organization Science*, Vol. 21 (2010), pp. 1016–1033.

KAY, L. "The Effect of Inducement Prizes on Innovation: Evidence from the Ansari Xprize and the Northrop Grumman Lunar Lander Challenge." *R&D Management*, Vol. 41 (2011), pp. 360–377.

KNOEBER, C.R. AND 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: Oxford University Press, 2009.

KOSFELD, M. AND NECKERMANN, S. "Getting More Work for Nothing? Symbolic Awards and Worker Performance." *American Economic Journal: Microeconomics*, Vol. 3 (2011), pp. 86–99.

LAFFONT, J.-J., OSSARD, H., AND VUONG, Q. "Econometrics of First-Price Auctions." *Econometrica*, Vol. 63 (1995), pp. 953–980.

LAFFONT, J.-J. AND VUONG, Q. "Structural Econometric Analysis of Descending Auctions." *European Economic Review*, Vol. 37 (1993), pp. 329–341.

———. "Structural Analysis of Auction Data." *American Economic Review*, Vol. 86 (1996), pp. 414–420.

LAKHANI, K.R., BOUDREAU, K.J., LOH, P.-R., BACKSTROM, L., BALDWIN, C., LONSTEIN, E., LYDON, M., MACCORMACK, A., ARNAOUT, R.A., AND GUINAN, E.C. "Prize-Based Contests Can Provide Solutions to Computational Biology Problems." *Nature Biotechnology*, Vol. 31 (2013), pp. 108–111.

LAZEAR, E.P. "Agency, Earnings Profiles, Productivity, and Hours Restrictions." *American Economic Review*, Vol. 71 (1981), pp. 606–620.

LI, Q. AND RACINE, J.S. *Nonparametric Econometrics: Theory and Practice*. Princeton, N.J.: Princeton University Press, 2007.

LIST, J., SOEST, D.V., STOOP, J., AND ZHOU, H. "On the Role of Group Size in Tournaments: Theory and Evidence from Lab and Field Experiments." Working Paper no. 20008, National Bureau of Economic Research, 2014.

LIU, T.X., YANG, J., ADAMIC, L., AND CHEN, Y. "Crowdsourcing with All-Pay Auctions: A Field Experiment on Taskcn." *Management Science*, (forthcoming).

MAAS, H.L.V.D. AND WAGENMAKERS, E.-J. "Psychometric Analysis of Chess." *American Journal of Psychology*, Vol. 118 (2005), pp. 29–60.

MILLNER, E.L. AND PRATT, M. D. "An Experimental Investigation of Efficient Rent-Seeking." *Public Choice*, Vol. 62 (1989), pp. 139–151.

MOLDOVANU, B. AND SELA, A. "The Optimal Allocation of Prizes in Contests." *American Economic Review*, Vol. 91 (2001), pp. 542–558.

——— "Contest Architecture." *Journal of Economic Theory*, Vol. 126 (2006), pp. 70–96.

MOLDOVANU, B., SELA, A., AND SHI, X. "Contests for Status." *Journal of Political Economy*, Vol. 115 (2007), pp. 338–363.

MURRAY, F., STERN, S., CAMPBELL, G., AND MACCORMACK, A. "Grand Innovation Prizes: A Theoretical, Normative, and Empirical Evaluation." *Research Policy*, Vol. 41 (2012), pp. 1779–1792.

NICHOLAS, T. "Cheaper Patents." *Research Policy*, Vol. 40 (2011), pp. 325–339.

ORSZAG, J.M. "A New Look at Incentive Effects and Golf Tournaments." *Economics Letters*, Vol. 46 (1994), pp. 77–88.

POETZ, M.K. AND SCHREIER, M. "The Value of Crowdsourcing: Can Users Really Compete with Professionals in Generating New Product Ideas?" *Journal of Product Innovation and Management*, Vol. 29 (2012), pp. 245–256.

PORTER, R.H. "The Role of Information in U.S. Offshore Oil and Gas Lease Auction." *Econometrica*, Vol. 63 (1995), pp. 1–27.

RACINE, J. AND LI, Q. "Nonparametric Estimation of Regression Functions with Both Categorical and Continuous Data." *Journal of Econometrics*, Vol. 119 (2004), pp. 99–130.

SCHMALENSEE, R., ARMSTRONG, M., AND WILLIG, R.D., eds. "The Timing of Innovation: Research, Development, and Diffusion." In *Handbook of Industrial Organization*. Amsterdam: North Holland, 1989.

SCHOTTER, A. AND WEIGELT, K. "Asymmetric Tournaments, Equal Opportunity Laws, and Affirmative Action: Some Experimental Results." *Quarterly Journal of Economics*, Vol. 107 (1992), pp. 511–539.

SZYMANSKI, S. "The Economic Design of Sporting Contests." *Journal of Economic Literature*, Vol. 41 (2003), pp. 1137–1187.

TANAKA, R. AND ISHINO, K. "Testing the Incentive Effects in Tournaments with a Superstar." *Journal of the Japanese and International Economies*, Vol. 26 (2012), pp. 393–404.

WOOTEN, J.O. AND ULRICH, K.T. "Idea Generation and the Role of Feedback: Evidence from Field Experiments with Innovation Tournaments." Working Paper no. 1838733, SSRN, 2011.

WU, S. AND ROE, B. "Tournaments, Fairness, and Risk." *American Journal of Agricultural Economics*, Vol. 88 (2006), pp. 561–573.

WU, S., ROE, B., AND SPORLEDER, T. "Mixed Tournaments, Common Shocks, and Disincentives: An Experimental Study." MPRA Paper no. 21, Munich Personal RePEc Archive, 2006.