The Distribution of Talent across Contests

Ghazala Azmat
Marc Möller

This version: June 2013
(May 2013)

Barcelona GSE Working Paper Series

Working Paper nº 600

The Distribution of Talent across Contests*

Ghazala Azmat †
and
Marc Möller ‡

Abstract

Do the contests with the largest prizes attract the most able contestants? To what extent do contestants avoid competition? In this paper, we show, theoretically and empirically, that the distribution of abilities plays a crucial role in determining contest choice. Sorting exists only when the proportion of high-ability contestants is sufficiently small. As this proportion increases, contestants shy away from competition and sorting decreases, such that, reverse sorting becomes a possibility. We test our theoretical predictions using a large panel data set containing contest choice over three decades. We use exogenous variation in the participation of highly-able competitors to provide empirical evidence for the relationship among prizes, competition, and sorting.

JEL *classification*: D82, M52, D02
Keywords: Contests, sorting , ability, prize allocation.

1 Introduction

Competition is a defining feature of most economic and social environments. Contestants of differing ability compete for valuable but limited resources by exerting effort. In many cases, contestants choose from a variety of potential contests. For example, architects

*We thank seminar participants at Bocconi, LSE, McGill, Toronto, and Stanford GSB for valuable comments and suggestions.

†Queen Mary University of London, CEP (LSE), Universitat Pompeu Fabra, Barcelona GSE. Email: g.azmat@qmul.ac.uk

‡Department of Economics, University of Bern. Email: marc.moeller@vwi.unibe.ch

choose design competitions; pharmaceutical companies select from a range of R&D contests; athletes pick sports tournaments; and college graduates apply for positions that offer alternative promotion schemes.

Rewarding contestants according to their relative performance is motivated primarily by the desire to increase effort. Lazear and Rosen (1981) were the first to consider rank-order tournaments as a way to provide incentives. In recent years, a large theoretical literature has been developed, determining the optimal design of such tournaments.[1] A common theme in this literature is that contestants exert greater efforts when prizes are larger and more concentrated towards the highest ranks.[2] Empirical evidence for these incentive effects has been provided by Ehrenberg and Bognanno (1990), using data on golf contests, and by Eriksson (1999) and Bognanno (2001) for labor tournaments. However, while the relationship between prizes and effort seems to be well understood, little is known about their influence on contest selection.[3]

For other incentive schemes, which use absolute rather than relative performance evaluation, selection effects have been found to be as important as incentive effects. Lazaer (2000) documents a 44-percent increase in productivity for a firm switching from salaries to piece rates and attributes half of this increase to selection effects. High-ability workers find firms offering piece rates more attractive than firms offering salaries. In the context of tournaments, it remains an open question whether selection effects play a similar role.

In this paper, we show that the contests with the largest prize(s) do not necessarily attract the most-able contestants. Instead, the distribution of talent across contests depends, in a systematic way, on the overall distribution of abilities among contestants. In our model, two types of contestants (high- and low-ability), choose between two types of contests (strong and weak competition). High-ability contestants have lower (constant) marginal costs of effort than low-ability contestants have. Strong competition offers fewer, but greater prizes than weak competition does. Our main theoretical result shows that the share of high-ability contestants who choose strong competition is decreasing in the overall fraction of high-ability contestants. When high-ability contestants become sufficiently numerous, sorting is reversed, i.e. weak competition attracts a larger share of high-ability contestants than strong competition does.

At a first glance, the possibility of reverse sorting seems counterintuitive since in this

[1]For an extensive survey, see Konrad (2009).

[2]Exceptions to this rule exist when contestants are risk-averse (Krishna and Morgan, 1998) or effort costs are sufficiently convex (Moldovanu and Sela, 2001).

[3]A notable exception is the work by Damiano, Li, and Suen (2010, 2013), who study the issue of sorting with a focus on peer effects rather than on effort choices.

case contestants are attracted by contests with smaller prizes *and* stronger opposition. However, we show that in the presence of a higher number of prizes, competition is mitigated. As a consequence, the contestants' effort costs are lower under weak competition than under strong competition. This underlines the importance of incorporating effort decisions into models of contest choice.

Our results have the following implications for contest design. They show that, when contest choice is endogenous, selection effects cannot be neglected, and the optimal prize allocation depends crucially on the distribution of abilities among potential contestants. This holds true, regardless of whether the objective is to maximize aggregate output or the winner's performance. More importantly, selection effects can be diametrically opposed to incentive effects, and the positive influence of large/concentrated prizes on efforts may be more than compensated by their negative influence on the self-selection of talented contestants.

In labor tournaments and other settings, testing for selection effects is difficult, if not impossible. In this paper, we take advantage of an unusually clean opportunity to investigate the extent of sorting across contests in a sports setting. With around 20,000 observations, we examine the contest choices of professional marathon runners over three decades. The setup allows us to abstract from a number of identification problems present in other types of data. In a labor-market setting, for example, it is often difficult to disentangle firms and worker types or to observe individual performance in team settings, while in marathons, individual performance is readily available, together with complete information on contest and runner characteristics.

There are two key features that make marathons the ideal setting to test our model. First, five major marathons (Berlin, Boston, Chicago, London, and New York) offer more than 50 percent of the total available prize money and, on average, allocate a greater percentage of the prize money to the winner. This allows us to identify a runner's decision between competing in a major or a minor marathon, as a choice between strong and weak competition. Second, highly-talented East-African runners, mainly from Kenya and Ethiopia, dominate the sport of marathon running. This dominance is striking and unparalleled in other sports. For example, according to the International Association of Athletics Federations' (IAAF) Top List, the 50 fastest male marathon runners in 2012 were exclusively from Kenya or Ethiopia. This endows us with a proxy of the contestants' abilities (runners' origin), which, unlike performance measures (finishing times), is independent of effort and prize considerations.

Following Brückner and Ciccone (2010), we use exogenous variation in local economic conditions to predict the participation of East-African runners. We find that the likelihood that a high-ability runner will participate in a marathon is increasing in the race's prize budget but decreasing in the expected number of high-ability opponents. The participation of one additional East-African opponent in last year's race must be compensated by a $6250 increase in the race's prize budget to keep the race equally attractive to high-ability runners. Interestingly, the steepness of a race's prize structure is found to have a positive effect on participation when opposition is expected to be weak but a negative effect when opposition is expected to be strong. Hence, whether selection effects are in line with or opposed to incentive effects seems to depend on the overall competitiveness of the environment.

Our paper documents that sorting exists only when the proportion of high-ability participants is sufficiently small. The results constitute first evidence for tournament selection effects in a real setting.[4] In line with our main theoretical result, we find that, when the number of high-ability participants increases, potential participants become more likely to avoid competition. In particular, when the share of talented contestants increases by ten percent, the fraction of these runners who choose to participate in major races falls by 10.3 percent. These results suggest that, depending on the ability distribution and prize structure, contestants avoid one another to the extent that reverse sorting becomes a possibility.

2 Theoretical Framework

We present a simple theoretical framework to illustrate the effect of changes in the ability distribution on the level of sorting across contests. The model demonstrates that the provision of strong incentives increases participation of talented contestants, but that talent crowds out talent. The model makes precise how these two factors interact, resulting in a negative relation between the frequency of high abilities and the level of sorting and in the possible existence of reverse sorting.

[4]In an experimental setting, some recent studies have focused on the choice between tournaments and alternative incentive schemes and on selection effect between tournaments offering a single prize of differing size (Dohmen and Falk, 2011; Eriksson et al. 2009; Leuven et al. 2011).

2.1 Setup

We assume a continuum of contests and a continuum of risk-neutral players.[5] All contests allow for the same number of participants, denoted by $N \geq 3$. In order to balance the number of players with the number of available contest slots, we assume that there exists a mass 1 of players and a mass $\frac{1}{N}$ of contests.

There are two types of contests, *strong* contests and *weak* contests, $j \in \{S, W\}$. They differ on two dimensions: first, the number of prizes they offer, and second, the size of the prize. More specifically, a contest of type j offers $M_j \in \{1, 2, \ldots, N-1\}$ prizes, identical in size, $b_j > 0$.[6] Strong contests award larger ($b_S > b_W$) but fewer ($M_S < M_W$) prizes than weak contests.[7] Apart from the differences in their prize structures, strong and weak contests are assumed to be identical. For simplicity, we assume that both types exist in equal fractions.[8]

There are two types of players, low-ability players and high-ability players, $i \in \{L, H\}$. A low-ability player's constant marginal cost of effort is normalized to $c_L = 1$, while a high-ability player's marginal cost of effort is lower and equal to $c_H = c \in (0, 1)$. The crucial parameter of the model is the fraction of players who have high-ability, denoted by h. We focus on the case in which high-ability players are in the minority, $h \in (0, \frac{1}{2})$. This assumption guarantees that, if they desire, all high-ability players can enter a strong contest.

The model has two stages. In the first stage, players enter contests, and in the second stage, they compete by exerting effort. At the *entry stage*, players form expectations about their opponents' abilities based on their knowledge of the overall distribution of types and the equilibrium strategies. At the *competition stage*, players observe their opponents' abilities and then simultaneously make their effort choices.

We model competition as a perfectly discriminating contest, where prizes are awarded to the players who exert the highest levels of effort.[9] This follows an extensive literature

[5]Our model captures settings with many contests and a large number of players. In such settings, a single player's action has no effect on the optimal contest choice of the remaining players. This rules out coordination issues and guarantees the uniqueness of equilibrium. The implications of risk aversion are discussed at the end of the section.

[6]The assumption that all the contest's prizes are identical makes the model tractable. A general description of competition for the case of heterogeneous players and non-identical prizes is still missing. Cohen and Sela (2008) made a first step in this direction.

[7]In a labor tournament setting, Yun (1997) shows that first-best efforts and efficient self-selection can be achieved when workers are offered the choice between a tournament with many large prizes and a tournament with few small prizes.

[8]Our results remain qualitatively unchanged when this assumption is relaxed. For details, see the discussion at the end of this section.

[9]Alternatively, winners could be determined stochastically–i.e., in dependence of efforts and random

on contest design (see, for example, Clark and Riis (1998) and Moldovanu and Sela (2001, 2006)). In terms of payoffs, a player of type i who exerts effort $e \geq 0$ in a contest of type j will receive utility $U_i^j = b_j - c_i e$ if he wins one of the M_j prizes, and $U_i^j = -c_i e$ otherwise.

Since, at the competition stage, players can guarantee themselves a payoff of zero by exerting zero effort, at the entry stage, no player will choose not to participate in any contest at all. This means that if a fraction $q_i \in [0,1]$ of type i players enters strong contests, then the remaining fraction $1-q_i$ will enter weak contests. The players' behavior at the entry stage can, therefore, be completely described by the fractions of low-ability (q_L) and high-ability (q_H) players that enter a strong contest.

The distribution of players across contests can be characterized as exhibiting: *complete sorting* when all high-ability players enter a strong contest, $q_H = 1$; *partial sorting* when a larger number of high-ability players enter strong contests than weak contests, $q_H > \frac{1}{2}$; and reverse sorting when the opposite is the case, $q_H < \frac{1}{2}$.

An equilibrium distribution of talent (q_H, q_L) has to satisfy two conditions: an *optimality condition* and a *feasibility condition*. The optimality conditions requires that no player must be able to increase his payoff by entering another (type of) contest. This means that if players of the same type i enter both types of contests, $q_i \in (0,1)$, then these players must expect equal payoffs. In addition, if all players of type i enter the same type of contest–i.e., $q_i \in \{0,1\}$–then their expected payoff must not be higher in the other type of contest. The feasibility condition requires that the number of players who participate in a given type of contest must equal the number of available slots in contests of this type:

$$hq_H + (1-h)q_L = h(1-q_H) + (1-h)(1-q_L) = \frac{1}{2}. \tag{1}$$

Our analysis proceeds by backward induction and consists of two steps. Section 2.2 characterizes the players' effort choices and expected payoffs in a contest with a *given* set of opponents. The main result necessary for the subsequent analysis, which is the focus of our study, is that a player's expected payoff is positive (and equal to $b_j(1-c)$) if and only if the player has high-ability and the number of high-ability opponents is strictly smaller than the number of prizes M_j. In Section 2.3, we use this insight to derive our main theoretical results on the players' individual contest choice and the equilibrium distribution of talent across contests. All proofs are given in the Appendix.

factors. For a discussion of this case, see footnote 10.

2.2 Competition

In making their effort choices, players trade off a higher chance of winning against an increase in their costs of effort. In this section, we derive testable predictions on the influence of prizes and opposition on efforts. We briefly discuss the nature of the equilibrium, with a focus on the players' expected payoffs, which are crucial for the entry-stage analysis contained in the next section. A more-detailled characterization of the players' equilibrium strategies based on the results of Clark and Riis (1998) can be found in the Appendix.

In equilibrium, $M+1$ of the most-able players are active by randomizing over an interval of potential effort choices, while the remaining players choose zero effort with certainty. The equilibrium depends on the relationship between the number of high-ability participants N_H and the number of available prizes M. When $N_H \leq M$, then some of the active players will have low-ability and all players will randomize over the interval $[0, b]$. Since players must be indifferent between their potential effort choices and are guaranteed to win a prize when choosing $e = b$, expected payoffs must be $b - b = 0$ for all low-ability players and $b - cb$ for all high-ability players. For low-ability players, (expected) prize winnings are exactly offset by the (expected) costs of effort.[10] High-ability players enjoy a comparative advantage due to their lower marginal cost of effort and, therefore, obtain a positive payoff. This comparative advantage disappears when $N_H > M$. In this case, all active players have the same high-ability, randomize over $[0, \frac{b}{c}]$, and expect a payoff of zero. In the Appendix, we prove the following:

Proposition 1 *High-ability players' (expected) efforts are increasing in the size of prizes, as well as in the steepness of the contest's prize structure. Efforts are minimal when the number of high-ability players equals the number of prizes.*

Increasing the size of prizes b leads to larger marginal returns to effort. Moreover, by simultaneously increasing b and decreasing M in such a way that the contest's prize budget Mb remains unchanged, we can increase the prize structure's steepness. In the proof of Proposition 1, we show that both changes lead to a first-order (stochastic) upward shift of the distribution of effort. These results are intuitive and in line with other models in which players are assumed to be risk-neutral and to have linear costs of effort (e.g.,

[10] This is a consequence of contests being perfectly discriminating. If contests involved a random element, then the expected payoffs of low-ability players would depend on prizes, but this dependence would still be weaker than it is for high-ability players. Since sorting can be expected to be strongest when ability matters most, the absence of randomness is the most conservative assumption with respect to our finding that sorting may be reversed. For a detailed study of the relationship between a contest's prize structure and its randomness, see Azmat and Möller (2009).

Moldovanu and Sela (2001)). The influence of the allocation of prizes on the incentives to exert effort has been a major theme of the literature on contest design.

The intuition for the last part of Proposition 1 is as follows. For $N_H > M$, all $M+1$ active players have the same (high) ability, leading to strong competition. For $N_H = M$, exactly one of the active players has low-ability. This player is discouraged by the comparative advantage of his M high-ability opponents and therefore exerts low levels of effort. High-ability players anticipate this and, therefore, also exert low effort. Finally, when $N_H < M$, more than one low-ability players are active. Low-ability players are encouraged to exert effort by the presence of other (active) low-ability players, which, in turn, leads to higher efforts by the high-ability players.

Having characterized the players' effort choice at the competition stage, we are now ready to move to the paper's main focus and consider the players' contest choice at the entry stage.

2.3 Contest choice

In this section, we first derive the players' preferences over contests in dependence of the contest's prize structure and the expected opposition. In a second step we then determine the equilibrium allocation of talent across contests.

Individual preferences

The analysis in the preceding section showed that low-ability players expect the same (zero) payoff, independent of the type of contest they enter. Hence, low-ability players are indifferent between the two types of contests, and we can concentrate our analysis on the preferences of high-ability players. The expected payoff of a high-ability player does depend on the specific features of the contest he enters. In the preceding section, we demonstrated that in a contest offering M prizes of size b, a high-ability player expects a positive payoff equal to $b(1-c)$ if the number of high-ability opponents is smaller than M and a zero payoff otherwise.

At the time of entry, the number of high-ability opponents in a given contest is uncertain. Hence, from the viewpoint of the entry stage, the player's preferences will depend on the likelihood p with which an opponent has high-ability. The probability with which a high-ability player obtains a positive payoff is then identical to the probability with which he meets, at most, $M-1$ high-ability opponents. This is given by the binomial sum

$$G(M,p) = \sum_{N_H=0}^{M-1} \binom{N-1}{N_H} (p)^{N_H} (1-p)^{N-1-N_H}. \tag{2}$$

A high-ability player's expected payoff from entering the contest is

$$E[U_H] = b(1-c)G(M,p). \tag{3}$$

It depends on the contest's prize structure, represented by M and b, and the expected opposition, given by the likelihood p of meeting high- rather than low-ability opponents. The following proposition contains comparative statics, which will be the first subject of our empirical analysis.

Proposition 2 *A high-ability player's expected payoff from entering a contest is increasing in the number M and size b of its prizes, but decreasing in the probability p with which opponents have high-ability. Payoffs are increasing in the steepness of the contest's prize structure when opposition is weak ($p < \bar{p}$) but decreasing when opposition is strong ($p > \bar{p}$).*

The first part of Proposition 2 is intuitive and follows easily from (2) and (3). The last part of Proposition 2 considers the effect of a decrease in the number of prizes, accompanied by an increase in the size of the prize. As can be seen from the proof, the particular value taken by the threshold $\bar{p}$ depends on the specific changes in M and b. Intuitively, when the probability of meeting high-ability opponents is small, high-ability players prefer a steeper prize structure due to their comparative advantage over low-ability players. In contrast, when the probability of meeting high-ability opponents is large, high-ability players prefer a flatter prize structure due to their mitigating effect on competition and the resulting decrease in effort costs.

To summarize, while prizes, both in size and number, are predicted to affect a player's decision to enter a particular contest positively, the effect of (expected) opposition is negative. Moreover, opposition not only has a level effect, but also an interactive effect with the steepness of the contest's prize structure.

Equilibrium allocation

Having described the players' individual preferences, we now determine their equilibrium allocation across the two types of contests. Our analysis proceeds as follows. For a given allocation (q_H, q_L), we determine the likelihoods p_j of meeting high-ability opponents in a contest of type $j \in \{S, W\}$, which allows us to calculate the players' expected payoffs in both types of contest. We then verify whether the optimality and feasibility conditions outlined above are satisfied. The indifference of low-ability players implies that optimality needs to be checked only for high-ability players and that feasibility is guaranteed by the low-ability players' willingness to fill any slot that has remained idle.

For a given allocation (q_H, q_L), the number of high-ability players who choose a strong contest is given by hq_H. There are $\frac{1}{2N}$ strong contests, each offering N slots. The likelihood with which a slot in a strong contest is filled with a high-ability opponent can be calculated by dividing the number of high-ability players who choose a strong contest, hq_H, by the overall number of slots available in the strong contests, $\frac{1}{2}$. It is given by $p_S = 2hq_H$. Similarly, the likelihood with which a slot in a weak contest is filled by a high-ability opponent is given by $p_W = 2h(1 - q_H)$.

To check optimality for high-ability players, we need to consider the difference between their expected payoffs from entering a strong versus a weak contest. From (3) this difference is

$$\Delta \equiv (1 - c)[b_S G(M_S, p_S) - b_W G(M_W, p_W)]. \tag{4}$$

High-ability players strictly prefer a contest of type S (W) when $\Delta > 0$ $(\Delta < 0)$ and are indifferent when $\Delta = 0$. In the Appendix, we prove the following result:

Proposition 3 *There exists a unique equilibrium allocation (q_H^*, q_L^*) of abilities that depends on the overall fraction h of high abilities in the population of players. In particular, there exist critical values $\bar{h} \in (0, \frac{1}{2})$ and $\bar{\bar{h}} \in (\bar{h}, \frac{1}{2}]$ such that the following holds:*

1. *For $h \leq \bar{h}$, sorting is complete, $q_H^* = 1$. All high-ability players enter strong contests.*

2. *For $\bar{h} < h < \bar{\bar{h}}$, sorting is only partial, $q_H^* \in (\frac{1}{2}, 1)$. Strong contests attract a greater number of high-ability players than weak contests. Moreover, talent crowds out talent–i.e., q_H^* is strictly decreasing in h.*

3. *For $\bar{\bar{h}} \leq h$, sorting is reversed, $q_H^* \leq \frac{1}{2}$. Strong contests attract a smaller number of high-ability players than weak contests.*

An increase in M_S or b_S and a decrease in M_W or b_W all lead to a higher level of sorting by increasing q_H^* and $\bar{h}$.

The intuition for this result is as follows. Strong contests offer high prizes, while weak contests mitigate competition by offering many prizes of smaller value. From the viewpoint of a high-ability player, effort considerations become more important as the likelihood of meeting high-ability rivals increases, and his comparative advantage over low-ability players plays a smaller role. When high abilities become sufficiently frequent, the mitigation of competition outweighs all else, such that high-ability players prefer weak contests over strong contests, even though prizes are smaller and rivals are more-able in the former than in the latter. This contrasts with the common intuition that, in equilibrium, contest choices should be driven by a trade-off between high prizes and strong opposition versus low prizes and weak opposition. The possibility of reverse sorting, therefore, underlines the importance of including effort considerations in models of contest choice.

For the general case, we cannot rule out that $\bar{\bar{h}} = \frac{1}{2}$. To see that reverse sorting is indeed a possibility given our assumption that $h < \frac{1}{2}$, we provide an example where $\bar{\bar{h}} < \frac{1}{2}$.

Example: Reverse sorting between one-prize and two-prize contests. Consider the special case in which both types of contests have the same total prize budget B. Let strong contests award their entire budget to the player with the highest effort–i.e., $M_S = 1$ and $b = B$. Let weak contests offer two identical prizes instead–i.e., $M_W = 2$ and $b = \frac{B}{2}$. In the proof of Proposition 3, we show for the general case that Δ is strictly decreasing in q_H. This is intuitive since an increase in q_H raises the expected opposition in a strong contest while lowering the expected opposition in a weak contest. Hence, $\bar{\bar{h}} < \frac{1}{2}$ if and only if $\Delta(q_H = \frac{1}{2}) < 0$ for some $h < \frac{1}{2}$. For the special case under consideration, substitution of M and b into (4) leads to

$$\Delta(q = \frac{1}{2}) = (1-c)\frac{B}{2}(1-h)^{N-2}(1-Nh). \tag{5}$$

This shows that reverse sorting between one-prize and two-prize contests of identical budgets exists when $h > \frac{1}{N}$. For example, when contests allow for 20 participants, then sorting would be reversed already when more than 5% of the players in the population of potential participants have high-ability. ■

Let us discuss possible implications of risk aversion on the players' contest choice. From the viewpoint of a high-ability player, each contest can be understood as a lottery with two possible outcomes. A high payoff is obtained when the number of high-ability participants fails to exceed the number of prizes, and a low payoff is obtained otherwise. For $q_H > \frac{1}{2}$, the high payoff, though smaller, is more likely to be obtained in weak contests than in strong contests. Hence, weak contests constitute the less-risky lottery. Risk aversion gives high-ability players an additional incentive to choose a weak rather than a strong contest.[11] This makes our assumption of risk-neutrality the most conservative

[11] This is in line with Dohmen and Falk's (2011) experimental finding that subjects who choose a tournament rather than a fixed payment have a lower degree of risk aversion.

with respect to the possibility of reverse sorting.[12]

Finally, let us consider the effect of relaxing our assumption that both types of contests exist in equal fractions. Suppose, for example, that there exists a larger number of strong contests than weak contests. In this case, the likelihood of meeting a high-ability opponent in a strong contest is lower than $2hq_H$, and the likelihood of meeting a high-ability player in a weak contest is higher than $2h(1-q_H)$, for any given value of q_H. This makes strong contests more attractive relative to weak contests, leading to a (weak) upward shift in the equilibrium value of q_H^*. The thresholds $\bar{h}$ and $\bar{\bar{h}}$ shift to the right. The results in Proposition 3 change quantitatively but remain qualitatively unchanged.

3 Empirical Framework

The theoretical framework makes precise that sorting in contests exists only if the probability of meeting other talented contestants is sufficiently small. Testing the predictions of the model requires the observability of individual abilities and an exogenous change in the overall distribution of abilities. Our test relies on two sources of variation: the variation in the distribution of ability of contestants and the variation in prizes and prize structure across contests. In this section, we test the predictions of our model using a large panel dataset of international city marathons and professional marathon runners, which spans over three decades.

The marathon setting is an ideal one for testing our model. Marathons share many features with other contests, such as those seen in a labor-market setting. However, unlike in labor tournaments, prizes and performance are easily observed. It is often difficult, if not impossible, to know the pay structure within firms. Moreover, workers' individual performance is seldom observed; nor are there well-defined measures that are recognized across firms, even for those in the same industry or sector. While marathons are fairly homogeneous in their setup, firms often differ in dimensions other than their pay structure.[13] Finally, professional runners typically enter two marathons per year, while employment relations are established less frequently, making equilibrium behavior less likely to emerge.

Beyond these advantages of marathons over labor tournaments, two important factors

[12] Note that this discussion ignores that risk aversion may also influence the way in which players compete. It has been shown, for example, that risk aversion decreases the effort of low-ability contestants but increases the effort of high-ability contestants in single-prize contests (Fibich et al., 2006).

[13] Some marathons have faster (flatter) race courses than others, but there exist conversion factors constructed by the Association of Road Running Statisticians to make marathons comparable. We adjust all the finishing times in our dataset using these conversion factors.

make them the ideal setting to test our theory: first, the possibility to differentiate between strong and weak contests; and, second, the opportunity to identify some of the most-able contestants by exogenous measures rather than by their performance.

Regarding contest types, there are five races, known as the World Marathon Majors, which have a special status in running, comparable to the Grand Slam tournaments in tennis. For historical reasons, these marathons in Berlin, Boston, Chicago, London, and New York, offer the highest prize budgets and have the largest number of participants.[14] The major marathons award more than 50 percent of all total prize money, and compared with other races, they are about twice as likely to choose a prize allocation that is steeper than the average. Major marathons are characterized not only by large and concentrated prize budgets, but also by a high number of runners competing for each prize, identifying the World Marathon Majors as the strong contests of our theoretical model. A marathon runner, therefore, faces the trade-off that is at the heart of our setup: Participate in a major marathon, which offers large prizes but strong competition, or choose a minor marathon with smaller prizes but weak competition.

Identifying the ability distribution of contest participants is often complicated, and basing it on outcome variables, such as finishing times, is likely to be endogenous to the prize distribution. Here, we take advantage of a unique opportunity to recognize ability based on ethnic origin, which allows us to abstract from the usual identification problems. In the 1980s, a number of East-African runners began participating in long-distance running contests, and from the onset, it became apparent that these runners were very talented. Today, a surprisingly high fraction of the best marathon runners are of East-African origin. In 2009, for example, 88 of the 100 fastest (male) marathon runners were from either Kenya or Ethiopia.[15] This dominance, unparalleled in other sports, has been explained by genetic, social, nutritional, and geographical factors (Noakes, 1985). For the purpose of our analysis, this fact allows us to identify some of the most-able contestants by origin, which, unlike past performance, is independent of prize and effort considerations.

Since the 1980s, the number of East-African runners who compete internationally has increased, and marathon running, in general, has become more competitive. The change in competition can be seen in Figure 1, which depicts the ratio of the fastest race time

[14]Collectively, the group annually attracts more than five million on-course spectators, 250 million television viewers, and 150,000 participants. Its economic impact has been claimed to lie above $400 million. For more details, see http://worldmarathonmajors.com/US/about/.

[15]See Top List of the International Association of Athletic Federations (IAAF) available online at http://www.iaaf.org/statisitics/toplist/index.html.

of the year over the average time of runners finishing a race in the top 20. While in the early 1980s, the fastest runners had a comparative advantage of around six percent, this advantage decreased to less than two percent in the late 2000s. For a race won in two hours and ten minutes, this is equivalent to a reduction in the lead from eight minutes (2.6km) to two minutes (600m).

As a brief preview of our results, Figure 2 depicts the distribution of East-African runners across the two race categories, major and minor. In line with the predictions of our theoretical model, the higher the overall proportion of East-African (high-ability) participants, the lower their share in a major versus a minor marathon.

3.1 Data Description

We use data from the Association of Road Running Statisticians containing detailed race and runner information for the largest international marathons from 1986 to 2009. We restrict attention to the 35 most relevant marathons.[16] These are the races that have existed for the longest time, such that they are present in our sample for the whole period. They feature the highest participation, highest prize budgets and the fastest winning times. For each race, we observe the date, location, and the prize distribution. At the runner level, we identify the top (professional) finishers for each race. Since we are interested in the race choice of the most-able runners, we restrict attention to the first twenty finishers of each race. Since marathons award fewer than twenty prizes for each race, our data contain runners who win and runners who do not win a prize. We have information on the runners' gender, nationality, date of birth, finishing time, finishing position, and the prize awarded (if any). Tables 1 and 2, provide the main descriptive statistics for races and runners, respectively.

In Table 1, we show the descriptive statistics separately for major and minor races. From this table, we can see that there are stark differences between these race categories. Major races award around eight times as much prize money as minor races ($221,689 compared with $26,371). The prize structure of a major race is also steeper than that of a minor race (57 percent of the major races have a prize allocation that is steeper than the average, compared to only 35 percent for minor races).[17] In addition, major marathons have (overall) around three times more participants than minor marathons

[16]These are: Beijing, Berlin, Boston, California International, Chicago, Dallas, Detroit, Dublin, Frankfurt, Gold Coast, Grandma's, Hamburg, Honolulu, Houston, Italia, Kosice, London, Los Angeles, Madrid, New York, Ottawa, Paris, Reims, Richmond, San Antonio, Rome, Seoul, Stockholm, Tokyo, Turin, Twin Cities, Valencia, Venice, Vienna, Warsaw.

[17]Steepness is measured by the Herfindahl concentration index, calculated for the top three prizes.

(22,332 compared with 6,838). A majority of these runners are amateurs, but their number acts as a good signal for the level of competition in these races.

The two types of races also differ in the quality of the runners they attract. From Table 1, we can see that, on average, over all years, the fraction of high-ability runners has been considerably larger in the major races. This holds whetheror not we identify high-ability runners by origin or by (course-adjusted) finishing times. For example, 18 percent of the finishers in the major races were East-African, compared to only 14 percent in the other races. Similarly, 29 percent of runners in the major races had a finishing time within five percent of the year's best, compared with only eight percent in the minor races. As a consequence, winning times in major races are, on average, eight minutes faster which is equivalent to a 2.6km lead. Part of the difference in finishing times can be explained, in accordance with the model, by the higher effort (incentive effect) induced by the larger prizes offered in a major race. The remaining part is due to selection effects, which will be the focus of our analysis.

Table 2 shows the descriptive statistics of runners. In this table, we compare East-African runners, high-ability Non-East-African runners, and other Non-East-African runners, respectively.[18] For male runners, we see that East-African runners are comparable to high-ability Non-East-African runners on a number of dimensions, including prize money ($7,676 versus $8,284), finishing times (two hours, 14 minutes versus two hours, 12 minutes), and the number of marathons entered in a given year (1.42 versus 1.44). Compared with other runners, however, these two groups look very different. For female runners, the same patterns hold. East-African runners are comparable with the best Non-East-Africans, lending support to our identification of East-African runners as high-ability contestants; but both groups are noticeably different from other runners. The focus of the analysis will be on these high-ability runners.

3.2 Do Runners Choose Races Based on Prizes?

Based on runner-race characteristics (finishing times, prizes), how important are (expected) prize winnings in a runner's race choice? It could be the case that a runner's race choice is driven by other (unobservable) factors such as sponsors' preferences. This issue is crucial for the rest of the analysis and for determining whether our empirical setting is appropriate to test our model.

As an illustration, we use the last year of the data, 2009, to investigate a runner's

[18] High-ability Non-East-African runners are defined as the 100 fastest Non-East-African runners within their gender category, based on their fastest finishing time for a given year.

potential prize winnings, while holding the behavior of all other runners fixed. We then construct the counterfactual outcome by counting the number of races in which the runner could have obtained a higher prize than in the one he actually ran, assuming identical, course-adjusted, performance.

We find that a surprisingly high fraction of runners choose a race that maximizes their prize winnings ex post. In particular, around 40 percent of the prize winners could not have earned a higher prize in any other marathon. A further 20 percent had only one alternative race in which their prize would have been higher. This suggests that (expected) prize winnings are an important determinant of runners' behavior.

3.3 Individual contest choice

Since our focus lies on selection effects rather than incentive effects, we first test the model's predictions with respect to the runners' contest choice. The effort analysis is postponed until the end of the section.

OLS Analysis

To test Proposition 2, we investigate how a runner's expected payoff from a marathon and, hence, his probability of entering, depends on the race's characteristics. Letting P_{ijt} denote the probability with which runner i enters race j in time period t, we estimate the following equation using OLS:

$$p_{ijt} = \alpha_0 + \alpha_A A_{jt-1} + \alpha_B B_{jt} + \alpha_C C_{jt} + \alpha_{AC}(A_{jt-1} * C_{jt}) + X_i\beta + \varepsilon_{ijt}. \tag{6}$$

The variable A_{jt-1} denotes the level of expected opposition. It is measured as the proportion of high-ability participants among the race's top 20 finishers in the previous year. The variable B_{jt} denotes the marathon's total prize budget. C_{jt} is a measure of the prize structure's steepness, calculated as the Herfindahl concentration index, based on the first three prizes. We also include a vector of control variables, X_i, containing the runner's age, nationality, gender, and ranking in the previous year, as well as dummy variables indicating whether the race took place on the runner's home turf and whether the year was an Olympic year. We also control for time dummies and race fixed-effects.

According to Proposition 2, the probability with which a runner enters a race will be increasing in prize money, B_{jt}, such that $\alpha_B > 0$, and decreasing in expected opposition, A_{jt-1}, such that $\alpha_A < 0$. Moreover, we expect the effect of steepness, C_{jt}, on entry to depend on the level of expected opposition. The model predicts that steep prizes are attractive only when there are sufficiently few opponents, and unattractive otherwise.

Therefore, we expect the cofficient on the interaction term $(A_{jt-1} * C_{jt})$ to be negative $(\alpha_{AC} < 0)$.

Since Proposition 2 is concerned with the preferences of high-ability contestants, we want to restrict our attention to the race choice of the top runners. However, since A_{jt-1} itself is based on the behavior of high-ability runners, we cannot estimate 6. In order to deal with this problem, we split the high-ability runners into two groups. In particular, we restrict the participation analysis to the high-ability Non-East-African runners and choose the proportion of East-African runners in a race's previous edition as proxy for the expected opposition.[19] We showed in Table 2 that both groups of runners are comparable in their characteristics. While this should give a causal estimation of the effect of expected opposition on race participation, there may be some residual issues that create other types of endogeneities. We will deal with these explictly in the next section.

In Table 3, we present the results without the interaction between opposition and prize steepness. Column 1 and 2 presents the baseline regression without and with controls, respectively. Column 3 includes time dummies and time dummies interacted by gender to control for the changing trends in the participation of (East-African) runners in marathons. Column 4 includes race fixed-effects, which allows for race-specific features that are attractive or unattractive to runners. Races tend to take place in the same month each year, but we also control for this, as a means to account for seasonal effects. Overall, we find that an increase in expected opposition is associated with a decrease in the entry of a high ability contestant in a race, and total prize money has a positive effect on entry. The results allow us to determine the "prize" that contestants are willing to pay for a reduction in opposition. They imply that a high-ability Non-East-African runner's likelihood of participation remains unchanged when a reduction in the number of East-African opponents by one is accompanied by a \$6,250 decrease in the race's prize budget.[20]

With respect to the prize allocation, Table 3 shows that, overall, steepness has a positive effect on participation once we control for time dummies. In Table 4, we present the results with the interaction between opposition and prize steepness. Columns 1 to 4 are presented in the same way as in Table 3. Interestingly, we find that there is a differential effect of prize steepness on entry, depending on the expected level of opposition. In line with the predictions of Proposition 2, we find that an increase in the prize structure's

[19] Our results are robust with respect to changes in the cut-off point for our definition of "high-ability."

[20] A reduction in the number of East-Africans by one is equivalent to a five percentage point decrease in expected opposition since the determination of A_{jt-1} is based on the race's top 20 finishers. Keeping the likelihood of participation constant, therefore, requires a prize budget reduction by $100,000 \cdot 0.05 \cdot \frac{0.0099}{0.0080} = 6,250$ dollars.

steepness is associated with an increase in entry only if the level of opposition is sufficiently small. As expected opposition increases, the effect of steepness is negative and significant.

IV Analysis

We have shown that expected opposition is negatively related to participation. An important concern, however, is that the main variable of interest, A_{jt-1}, might be correlated with some unobservable characteristics, leading to a biased estimate of α_A. If a race becomes attractive to all high-ability runners for reasons unexplained by our set of observables, it will create a positive correlation between the entry of these runners and the error term. For example, a race may announce a special award for the achievement of a new course record, thereby raising its attractiveness for both sets of runners: East-African and high-ability Non-East-African. This would translate into an upward-biased estimate of α_A. To deal with this issue, we instrument for the participation of East-African runners, A_{jt-1}, using exogenous variation in their entry. In other words, we use exogenous variation in the participation of East-African runners that is uncorrelated with the (unobservable) race characteristics. We do this by instrumenting A_{jt-1} with rainfall, as well as commodity prices, in Kenya and Ethiopia in the previous year, $t-1$. Both variables are correlated with the number of East-African runners who compete in a given year but uncorrelated with race characteristics. Moreover, the race choice of Non-East-African runners will be unaffected by these instruments, except through the effect that they have on A_{jt-1}.

The reasoning behind the two instruments follows a growing literature, mainly in political economy, which relates rainfall and commodity prices to economic conditions in Sub-Saharan countries. It has been shown that rainfall levels positively affect income per capita (Miguel et al., 2004) and the functioning of democratic institutions (Brückner and Ciccone, 2010) in Sub-Saharan African countries. In addition, Deaton (1999) documented that commodity price downturns cause rapidly worsening economic conditions in Sub-Saharan African economies. Therefore, we expect rainfall and commodity prices to have a positive effect on the international marathon participation of East-African runners. This is intuitive, since most East-African runners rely on the support of sponsors, part of which are local businesses or regional government agencies.

We construct international commodity price indices for Kenya and Ethiopia following Deaton (1999) and Brückner and Ciccone (2010). For this purpose, we use the International Monetary Fund monthly price data for exported commodities for the period 1986 to 2009 and the countries' export shares of these commodities taken from Deaton for 1990. The rainfall data, covering the period 1986 to 2009, is taken from the NASA Global

Precipitation Climatology Project. The first-stage estimates show that rainfall and commodity prices are, indeed, strongly related to the participation of East-African runners in international marathons. In particular, with the exception of commodity prices in Ethiopia, positive rainfall shocks and commodity price upturns, increase the number of East-African runners competing internationally. The instruments are strong, with a high F-statistic.

In Table 4, Column 5, we present the results for the IV estimates. Since the instruments are annual and do not vary across races, we focus on the interaction of the instrumented expected opposition with prize steepness. As in the OLS regression, we find that steepness has a positive effect on entry, but there is a heterogeneous effect, depending on the level of (expected) opposition. As opposition increases, prize steepness becomes less attractive. The effect is stronger than in the OLS regressions, suggesting that α_A is, indeed, upward-biased when using OLS.

Robustness

In the remainder of this section, we address four important concerns: 1) the endogeneity of the total prize budget; 2) the identification of Major races as strong contests; 3) potential changes in the quality of East-African runners in years when there are few (many); and 4) the possibility that high-ability Non-East-African runners finish outside the first twenty places in their races.

First, we may be concerned that race organizers adjust the total prize budget, B_{jt}, to keep their race attractive to high-ability contestants. If entry falls, race organizers may increase prize money. As a consequence, the coefficient on B_{jt} would be biased downwards. We deal with this problem by instrumenting the value of a race's prize budget with the exchange rate of the country where the race takes place, relative to a currency basket. We expect that a move in the exchange rate is associated with an exogenous change in the value of the race's prize budget. This change should not be associated directly with race entry. In order to construct a currency basket, we use the annual Special Drawing Rights basket provided by the International Monetary Fund.[21] Table 5 shows that when we instrument for the prize budget, the coefficient is positive and significant, as previously seen. However, the coefficient is larger (0.05 compared with 0.01 in the OLS), suggesting that the OLS is, indeed, downward-biased.

Second, in our subsequent analysis of sorting, we identify the major races as the

[21]This basket contains U.S. Dollars, Euros, Japanese Yen, and Pounds Sterling. Weights assigned to each currency are adjusted annually to take account of changes in the share of each currency in world exports and international reserves.

strong contests–i.e., as those with a steep prize structure. We verify our identification by repeating the analysis in the previous sections through making a distinction between entry into major and minor races. We define the variable Major, which takes the value 1 if the race is a major race and 0 otherwise, and we use it as an alternative to the Herfindahl concentration index to measure the prize structure's steepness. We find that our main results from the previous section hold. Being a major race increases entry, but as opposition increases, major races become less attractive to enter. The results are presented in Table 6.

Third, since our instrument is constructed using the changes in the entry of East-African runners in different years, we might be concerned that in the years when there are more (fewer) East-African runners, the quality of the marginal runner is lower (higher). We check this by looking at the finishing times of East-African runners in the years when there are many (few) and find that these times are not statistically different from one another.

Finally, our analysis was restricted to the top 100 (high-ability) Non-East-African runners in a given year and focused on the top 20 finishers in each race. We might worry that, at times when there are many East-African runners competing, high-ability Non-East-Africans fail to finish within the top 20 of their race. We check that this is not an issue using information on runners outside the top 20 finishers. We find that each year, all top 100 runners are within our top 20 race finishers.

3.4 Distribution of talent

While Proposition 2 was concerned with the individual preferences of contestants, Proposition 3's focus is on the equilibrium distribution of players across contests. We now move from the determinants of individual race choice to the analysis of the aggregate distribution of runners across races using the time-series variation of our dataset.

To test Proposition 3, we analyze whether an increase in the overall number of high-ability contestants leads to a more balanced distribution of talent across contests. More specifically, we test the following equation:

$$S_t^M = \alpha_0 + \alpha_1 HA_t + \alpha_2 B_t^M + t + \varepsilon_t. \tag{7}$$

The dependent variable, S_t^M, measures the level of sorting. It denotes the proportion of East-African runners who choose to participate in a major rather than a Minor marathon in period t. For $S_t^M = 1$ sorting is complete–i.e., East-African runners participate exclusively in major marathons. The main variable of interest, HA_t, is the overall proportion of East-African runners, in period t. According to Proposition 3, sorting should be decreasing in HA_t. The variable B_t^M denotes the proportion of the total prize money that is awarded in the Major marathons. According to Proposition 3, sorting should be increasing in B_t^M. We control for both time trends and for whether the year was an Olympic year. Since marathons can be divided into spring and autumn races and runners typically choose one from each group, we consider contest choice for a given gender category, per season rather than per year to allow for a richer analysis.

Table 7 shows the estimates for equation (7). Since in our theoretical model, the number of strong contests is identical to the number of weak contests, we first restrict our analysis (columns 1 to 4) to the top ten races. These races include the five major marathons, as well as the next five most important races (Hamburg, Honolulu, Frankfurt, Paris, and Rome). In columns 5 to 8, we consider the runners' allocation across all 35 races. The results are similar for both samples.

We find that an increase in the fraction of high-ability contestants leads to a significant decrease in sorting. More specifically, as the fraction of East-African runners in the top ten races increases by one percent, the share of East-Africans who choose a major marathon decreases by 1.28 percent. The effect is even stronger, 1.32 percent, when all 35 races are considered. These results constitute evidence for the decrease in sorting, as predicted by Proposition 3. As expected, we also find evidence for a positive relation between sorting and prize budget differences. In particular, a one-percent increase in the proportion of prize money awarded by the major races, leads to an increase in the share of East-African runners entering a major race by 1.22 percent for the top ten races and by 0.52 percent for all 35 races.

It is reassuring that these effects persist when we control for time trends, gender and differential trends across gender. We see that in an Olympic year, the proportion of East-African runners entering a Major marathon increases by ten percent. This is intuitive since participation in the Olympics is restricted by country quotas. Due to the large number of talented Kenyan and Ethiopian runners, many of them are unable to run the Olympic marathon, whereas runners of comparable ability but different nationality are able to participate with a higher probability. As a result, the proportion of East-African runners in the Major races, the next-best alternative to the Olympics, is higher in Olympic years.

We check the robustness of these results by using an alternative proxy for talent. Rather than using origin, we identify a group of high-ability runners in a given season using performances. Note that, since effort and ability are hard to separate, finishing times may be related to prize money. An advantage of using origin is, therefore, that this definition of high-ability is independent of prize money considerations. We identify high-ability runners as those who have (adjusted) finishing times within one-percent of the fastest finishing time during the season.[22] We also look at those finishing within five and ten percent of the fastest time, respectively. We conjecture that changes in the overall number of high-ability runners over the years are a result of the increase in African participation. However, this measure of high-ability is less restrictive, especially if the quality and composition of the group of East-African runners are changing over time.

Table 8 shows that our main results still hold when we repeat the analysis for the alternative measure of ability based on rankings. The sorting of high-ability runners into major races is increasing in the proportion of prize money on offer, but decreasing in the overall proportion of high ability runners. Interestingly, the decrease is stronger the more able the runners under consideration. In particular, a ten-percent increase in the proportion of high-ability runners reduces sorting by 46, seven, or three percent when high-ability refers to runners within one, five, or ten percent of the fastest time, respectively. Finally, note that in contrast to our estimation based on runners' origin, the Olympic year dummy is no longer significant, which is in line with the reasoning provided above.

3.5 Effort choice

In this section, we test the theoretical predictions on effort. In accordance with our previous analysis, we restrict attention to high-ability runners. The model has three main predictions: first, effort is increasing in the prize budget; second, effort is increasing in the steepness of the races' prize structure; and third, effort depends on the level of opposition.

We test the predictions using the following specification, where we estimate the effort, E_{ijt}, of (high-ability) runner i in race j at time t, using the runner's finishing time:

$$E_{ijt} = \alpha_0 + \alpha_B B_{jt} + \alpha_C C_{jt} + \alpha_{HA} HA_{jt} + X_i\beta + \varepsilon_{ijt}. \tag{8}$$

As before, B_{jt}, measures the total prize budget of race j at time t, and C_{jt} measures race j's prize steepness at time t. The vector X_i of controls is the same as before, and we again include time and race fixed-effects. HA_{jt} measures the total number of high-ability opponents in race j at time t, which is calculated using the Top 100 runners (by gender category), irrespective of ethnicity.

[22]The identification of high-ability runners is done separately for men and women.

The findings for (8) are shown in Table 9. We see that, in line with the model's predictions, as the races' prize budget increases, so too does effort. Similarly, as the prize structure becomes steeper, effort also increases. Interestingly, we find that as the total number of high-ability opponents increases, effort also increases. Our theoretical model predicts effort to be minimized when the number of high-ability players is equal to the number of prizes. Starting from this situation, effort increases regardless of whether we increase or decrease the number of high-ability opponents. Since the number of prizes is inversely related to the prize structure's steepness, we expect opposition to lead to lower efforts (longer finishing times) when the prize structure is relatively flat (more prizes than opposition) and to higher efforts (shorter finishing times) when the prize structure is relatively steep (fewer prizes than opposition). Although not significant, the negative sign of the interaction term in Table 9 is in line with this reasoning.

4 Conclusion

How do contestants choose in which contest to compete? And how much do they value potential prize offerings relative to expected opposition? Do contestants prefer contests with few high prizes over contests with many low prizes? And how do these preferences depend on their abilities? In this paper, we have provided both theoretical and empirical insight into these questions.

We have proposed a tractable model of contest choice that incorporates the contestants' effort decisions. The model's main result shows that the allocation of talent across contests depends on its overall distribution within the population of potential contestants. The standard intuition that contestants sort according to abilities fails to hold in general. Sorting is decreasing as high abilities become more frequent, and reverse sorting has been shown to be a possibility.

Using the contest choices of professional marathon runners, we have tested our model and found that a high-ability contestant's likelihood of participating in a contest is increasing in the total prize money offered but decreasing in the probability with which he expects to meet other high-ability contestants. The results highlight the trade-off between prizes and opposition and allow us to determine the "prize" that contestants are willing to pay for a decrease in opposition and that organizers must offer to guarantee their contest's attractiveness. Moreover, using exogenous variation in the level of competition, our results provide evidence for a strong negative relation between the level of sorting and the overall frequency of highly-talented contestants.

The design of a marathon tournament is a particular setting, yet it offers an ideal testing ground for our model. The basic trade-off between prizes and opposition, which determines contest selection in our framework, is present in a broad variety of settings, including labor tournaments, procurement contests, and R&D competition. Unlike in these other settings, in marathons, ability and incentives, as well as counterfactual payments, are observable, which allows us to disentangle these factors. Contrary to common belief, a steeper prize structure does not always attract more talented participants. We have shown that sorting is highly dependent on the ability distribution and whether selection effects are in line or opposed to incentive effects ultimately depends on the environment.

Tables and Figures

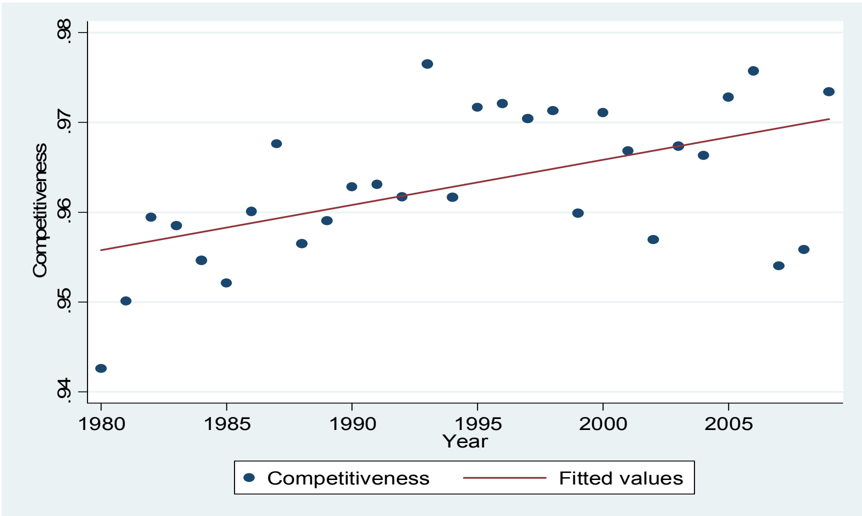

Figure 1: **Competitiveness of Marathon Running.** Competitiveness is defined as the ratio of the fastest (male) winning time of a year over the average finishing times of top 20 (male) finishers in all races. Finishing times are adjusted for racecourse differences.

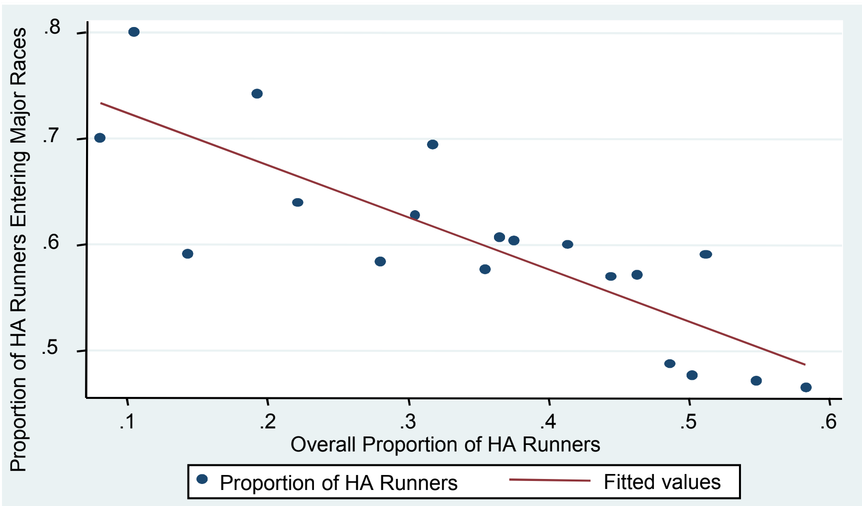

Figure 2: **Contest Choice.** "Overall proportion of HA runners" is the proportion of East-African runners among the top 20 finishers of the races in Berlin, Boston, Chicago, London, New York, Hamburg, Honolulu, Frankfurt, Paris, and Rome in a given year. "Proportion of HA Runners Entering Major Races" refers to the fraction of those East-African runners who entered a major race (Berlin, Boston, Chicago, London, New York) rather than a minor race (Hamburg, Honolulu, Frankfurt, Paris, and Rome).

	Major Races			All other Races		
Variable	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
Prize Budget ($)	238	221,689	126,466	1381	26,371	40,460
No. of Participants	236	22,332	10,143	859	6,838	6,462
Steep Prize	238	0.57	0.50	1381	0.35	0.48
Winning Time (hh:min)	238	02:17	00:09	1381	02:25	00:13
Fraction HA (Origin)	238	0.18	0.18	1381	0.14	0.22
Fraction HA (1%)	238	0.03	0.06	1381	0.00	0.02
Fraction HA (5%)	238	0.29	0.26	1381	0.08	0.17
Fraction HA(10%)	238	0.66	0.29	1381	0.36	0.36

Table 1: **Descriptive Statistics (Races)**Means and standard deviations for major and minor marathons, respectively. Major races are the Berlin, Boston, Chicago, London, and New York marathons. The sample period is 1986 to 2009. "Total Prize" is the sum of prizes awarded in a race in US dollars at 2000 prices. "Steep Prize" takes value 1 if the Herfindahl index, calculated for the top three prizes, is above its mean value. "No. of Participants" is the total number of participants, including amateurs, in a race. These data were collected separately from various sources, including ARRS and race websites. "Winning Time" is adjusted using ARRS conversion factors to ensure that times are comparable across races. "Fraction HA (Origin)" refers to the fraction of runners from East Africa. "Fraction HA (1%) (5%), (10%)" refers to the fraction of runners finishing within 1%, 5%, and 10% of the best time of the year, respectively.

	Male Runners								
	East-African			Top 100 Non-East-African			All others		
Variable	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
Age	2892	28.78	4.54	2684	30.05	4.14	4619	30.96	5.16
Prize ($)	2892	7,676	17,780	2684	8,284	16,048	4619	833	2,075
No. Races	2892	1.42	0.6	2684	1.44	0.61	4619	1.17	0.45
Fraction entering Major Race	2892	0.23	0.42	2684	0.38	0.49	4619	0.14	0.34
Finish Time	2892	2:14	0:05	2684	2:12	0:02	4619	2:20	0:05
	Female Runners								
	East-African			Top 100 Non-East-African			All others		
Variable	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
Age	646	27.69	4.44	2621	30.82	5.35	4840	32.26	6.31
Prize ($)	646	12,420	25,536	2621	10,339	18,319	4840	815	1,885
No. Races	646	1.45	0.59	2621	1.54	0.72	4840	1.19	0.46
Fraction entering Major Race	646	0.32	0.47	2621	0.43	0.49	4840	0.19	0.39
Finish Time	646	2:33	0:08	2621	2:32	0:04	4840	2:46	0:07

Table 2: **Descriptive Statistics (Runners)** Means and standard deviations (by gender category) for East-African runners, Top 100 Non-East-African runners, and all other runners, respectively. The sample period is 1986 to 2009. “No. of Races” is the number of races run in a given year. “Prize” is the prize money in US dollars at 2000 prices that a runner wins (on average) per race. “Finishing Times” have been adjusted using ARRS conversion factors to ensure that race courses are comparable.

Variable	OLS [1] enter	OLS [2] enter	OLS [3] enter	OLS [4] enter
Exp. Opposition (t-1)	-0.0293***	-0.0288***	-0.0141***	-0.0099**
	[0.002]	[0.002]	[0.004]	[0.004]
Total Prize ('00000)	0.0300***	0.0304***	0.0291***	0.0080***
	[0.001]	[0.001]	[0.001]	[0.002]
Steep Prize	-0.0044**	-0.0049**	0.0096***	0.0137***
	[0.002]	[0.002]	[0.002]	[0.002]
Female		0.0006	0.0036	0.0011
		[0.001]	[0.003]	[0.002]
Age		-0.0000**	-0.0000**	-0.0000**
		[0.000]	[0.000]	[0.000]
At Home		-0.0295***	-0.0364***	-0.0360***
		[0.006]	[0.006]	[0.007]
Nationality: US		0.0036***	0.0041***	0.0041***
		[0.001]	[0.001]	[0.001]
Rank (t-1)		-0.0001***	-0.0001***	-0.0001***
		[0.000]	[0.000]	[0.000]
Constant	0.0274***	0.0294***	0.0342***	0.0333***
	[0.001]	[0.001]	[0.004]	[0.004]
Time Fixed Effects	No	No	Yes	Yes
Race Fixed Effects	No	No	No	Yes
Observations	144,880	144,120	144,120	144,120
R-Squared	0.019	0.02	0.024	0.037

Table 3: **Probability of Entering a Race (OLS).** *,**,*** denotes significance at the 10%, 5%, and 1% level, respectively. The sample is restricted to the runners who were among the Top 100 Non-East-African runners in the previous year. The sample period is 1986 to 2009. "Exp. Opposition (t-1)" is the fraction of East-African runners among the top 20 finishers of the race in the previous year. "Total Prize ('00000)" is the total prize money of a race in US dollars at 2000 prices. "Steep Prize" is the Herfindahl index, calculated for the top three prizes. "At home'" takes the value 1 if the runner is racing in his or her home country. "Nationality" takes the value 1 if the runner is from the US and 0 otherwise. "Rank (t-1)" is the ranking of the runner in the previous year (between 1 and 100). The time fixed effects include a complete set of month and year dummies, as well as year and gender interactions.

Variable	OLS [1] enter	OLS [2] enter	OLS [3] enter	OLS [4] enter	IV-Opposition [5] enter
Exp. Opposition (t-1)	-0.0021	-0.0021	0.0138***	0.0046	
	[0.003]	[0.003]	[0.005]	[0.005]	
Total Prize ('00000)	0.0305***	0.0309***	0.0297***	0.0088***	0.0104***
	[0.001]	[0.001]	[0.001]	[0.002]	[0.001]
Steep Prize	0.0054**	0.0050**	0.0207***	0.0197***	0.0288***
	[0.002]	[0.002]	[0.002]	[0.002]	[0.004]
Exp.Opp(t-1)*Steep Prize	-0.0908***	-0.0912***	-0.0964***	-0.0499***	-0.0977***
	[0.008]	[0.008]	[0.009]	[0.009]	[0.016]
Female		-0.0002	0.0037	0.0009	0.0059
		[0.001]	[0.003]	[0.002]	[0.006]
Age		-0.0000**	-0.0000**	-0.0000**	0.0000
		[0.000]	[0.000]	[0.000]	[0.000]
At Home		-0.0294***	-0.0368***	-0.0362***	-0.0345***
		[0.006]	[0.006]	[0.007]	[0.006]
Nationality: US		0.0036***	0.0041***	0.0041***	
		[0.001]	[0.001]	[0.001]	
Rank (t-1)		-0.0001***	-0.0001***	-0.0001***	-0.0000***
		[0.000]	[0.000]	[0.000]	[0.000]
Constant	0.0249***	0.0273***	0.0327***	0.0325***	0.0335***
	[0.001]	[0.001]	[0.004]	[0.004]	[0.005]
Time Fixed Effects	No	No	Yes	Yes	Yes
Race Fixed Effects	No	No	No	Yes	Yes
Observations	144,880	144,120	144,120	144,120	144,120
R-squared	0.02	0.02	0.024	0.037	0.037
P-Value of F-test of exc. ins.					0.0000

Table 4: **Probability of Entering a Race (Instrument for Expected Opposition).** *,**,*** denotes significance at the 10%, 5%, and 1% level, respectively. Expected opposition is instrumented with the commodity price index in Kenya and Ethiopia in the previous year, as well as the (log) rainfall in Kenya and Ethiopia in the previous year. For definition of variables, see Table 3.

Variable	IV-Prize [1] enter	IV-Prize [2] enter
Exp. Opposition (t-1)	-0.0116***	0.0251**
	[0.004]	[0.012]
Total Prize ('00000)	0.0589**	0.0515**
	[0.025]	[0.023]
Steep Prize	-0.0066	0.0118***
	[0.010]	[0.004]
Exp.Opp(t-1)*Steep Prize		-0.1241***
		[0.042]
Female	0.0255*	0.0206*
	[0.013]	[0.012]
Age	0.0000	0.0000
	[0.000]	[0.000]
At Home	-0.0413***	-0.0407***
	[0.007]	[0.007]
Nationality	0.0041***	0.0041***
	[0.001]	[0.001]
Rank (t-1)	-0.0001***	-0.0001***
	[0.000]	[0.000]
Constant	-0.047	-0.027
	[0.081]	[0.075]
Time Fixed Effects	Yes	Yes
Race Fixed Effects	Yes	Yes
Observations	144,120	144,120
R-squared	0.021	0.026
P-Value of F-test of exc. ins.	0.0000	0.0000

Table 5: **Probability of Entering a Race (Instrument for Prizes).** *,**,*** denotes significance at the 10%, 5%, and 1% level, respectively. Total prize money is instrumented with the exchange rate of the country of the race relative to the Special Drawing Rights currency basket provided by the IMF. For definition of variables, see Table 3.

Variable	OLS [1] enter	OLS [2] enter	OLS [3] enter	OLS [4] enter	OLS [5] enter	OLS [6] enter	IV-Opposition [7] enter
Expected Opposition (t-1)	-0.0189***	-0.0183***	-0.0175***	-0.0070***	-0.0080***	-0.0075*	
	[0.002]	[0.003]	[0.004]	[0.002]	[0.003]	[0.004]	
Total Prize ('00000)	0.0147***	0.0148***	0.0141***	0.0206***	0.0208***	0.0206***	0.0249***
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.001]
Major Race	0.0504***	0.0511***	0.0535***	0.0682***	0.0691***	0.0671***	0.0719***
	[0.004]	[0.004]	[0.004]	[0.004]	[0.004]	[0.005]	[0.003]
Major Race*Exp. Opposition (t-1)				-0.1636***	-0.1651***	-0.1512***	-0.2237***
				[0.015]	[0.015]	[0.015]	[0.010]
Female		0.0006	0.0017		-0.0013*	0.0011	-0.0032
		[0.001]	[0.002]		[0.001]	[0.002]	[0.006]
Age		-0.0000**	-0.0000**		-0.0000**	-0.0000**	0.0000
		[0.000]	[0.000]		[0.000]	[0.000]	[0.000]
At Home		-0.0306***	-0.0316***		-0.0321***	-0.0349***	-0.0284***
		[0.006]	[0.006]		[0.006]	[0.006]	[0.006]
Nationality: US		0.0039***	0.0040***		0.0038***	0.0041***	
		[0.001]	[0.001]		[0.001]	[0.001]	
Rank (t-1)		-0.0001***	-0.0001***		-0.0001***	-0.0001***	-0.0000***
		[0.000]	[0.000]		[0.000]	[0.000]	[0.000]
Constant	0.0254***	0.0272***	0.0351***	0.0220***	0.0249***	0.0307***	0.0349***
	[0.001]	[0.001]	[0.004]	[0.001]	[0.001]	[0.004]	[0.005]
Time Fixed Effects	No	No	Yes	No	No	Yes	Yes
Observations	144,880	144,120	144,120	144,880	144,120	144,120	144,120
R-squared	0.023	0.024	0.027	0.026	0.027	0.029	0.03
P-Value of F-test of exc. ins.							0.0000

Table 6: **Probability to Enter a Race (Strong versus Weak Contests).** *,**,*** denotes significance at the 10%, 5%, and 1% level, respectively. "Major Race" takes value 1 if the race is a Berlin, Boston, Chicago, London, or New York marathon. Expected opposition is instrumented with the commodity price index in Kenya and Ethiopia in the previous year, as well as the (log) rainfall in Kenya and Ethiopia in the previous year. For definition of variables, see Table 3.

	Top 10 Races				All Races			
Variable	Sorting [1]	Sorting [2]	Sorting [3]	Sorting [4]	Sorting [5]	Sorting [6]	Sorting [7]	Sorting [8]
Proportion of HA (Origin)	-0.7742***	-0.3551**	-1.0272**	-1.2758**	-0.6214***	-0.5321***	-1.2260**	-1.3164***
	[0.187]	[0.171]	[0.501]	[0.494]	[0.131]	[0.125]	[0.471]	[0.465]
Proportion of Prize		1.1128***	1.1749***	1.2193***		0.4822***	0.4887***	0.5204***
		[0.190]	[0.195]	[0.189]		[0.139]	[0.138]	[0.137]
Female	-0.0894*	-0.0734*	-0.2516	-0.2575*	-0.0090	-0.0297	0.0303	0.0139
	[0.050]	[0.042]	[0.153]	[0.148]	[0.035]	[0.033]	[0.113]	[0.111]
Trend			0.0125	0.02			0.0250*	0.0271*
			[0.017]	[0.016]			[0.015]	[0.015]
Trend*Female			0.0014	-0.0008			-0.0099	-0.0102
			[0.008]	[0.008]			[0.006]	[0.006]
Olympic Year				0.0967**				0.0598*
				[0.039]				[0.031]
Constant	0.8727***	-0.2134	-0.1688	-0.2579	0.4619***	0.1331	-0.0424	-0.0743
	[0.097]	[0.202]	[0.280]	[0.273]	[0.065]	[0.113]	[0.175]	[0.173]
Observations	79	79	79	79	79	79	79	79
R-squared	0.19	0.448	0.471	0.513	0.275	0.375	0.399	0.429

Table 7: **Sorting of High-Ability Runners (Origin).** *,**,*** denotes significance at the 10%, 5%, and 1% level, respectively. High-ability runners are defined as those who originate from Kenya or Ethiopia. Top 10 Races include the Major races (Berlin, Boston, Chicago, London, and New York) as well as Hamburg, Honolulu, Frankfurt, Paris, Rome. The dependent variable, “Sorting”, is the proportion of high-ability runners who enter a major rather than a minor race. “Proportion of HA” is the overall fraction of high-ability runners in the population of runners. Both variables are calculated separately for each race season (spring, autumn). “Proportion of Prize” is the proportion of the overall prize money awarded in the major races. “Trend” is a linear trend for the sample period 1986 to 2009. “Olympic Year” takes value 1 in years 1988, 1992, 1996, 2000, 2004, and 2008 and 0 in all other years.

VARIABLES	Top 10 Races Sorting [1]	Sorting [2]	Sorting [3]	All Races Sorting [4]	Sorting [5]	Sorting [6]
Proportion of HA (1%)	-1.9589***			-4.6357**		
	[0.707]			[2.148]		
Proportion of HA (5%)		-0.2751*			-0.7163***	
		[0.159]			[0.214]	
Proportion of HA (10%)			-0.1194			-0.3075***
			[0.146]			[0.110]
Proportion of Prize	0.3263*	1.0318***	1.1413***	1.2664***	0.7091***	0.4475***
	[0.176]	[0.126]	[0.119]	[0.286]	[0.140]	[0.082]
Female	-0.1602	-0.0097	-0.0432	0.1364	-0.0995	-0.1608*
	[0.128]	[0.105]	[0.122]	[0.221]	[0.112]	[0.083]
Trend	-0.0193**	-0.0139**	-0.0036	0.0171	-0.0170**	-0.0166***
	[0.008]	[0.007]	[0.008]	[0.014]	[0.007]	[0.005]
Trend*Female	0.0102*	0.0027	-0.0007	-0.0086	0.0060	0.0063**
	[0.006]	[0.005]	[0.005]	[0.010]	[0.005]	[0.003]
Olympic Year	0.0233	-0.0231	0.0071	-0.0634	-0.0008	0.0137
	[0.035]	[0.025]	[0.023]	[0.061]	[0.028]	[0.016]
Constant	1.0270***	0.1000	-0.1355	-0.3148	0.4351*	0.5788**
	[0.259]	[0.242]	[0.336]	[0.380]	[0.240]	[0.220]
Observations	79	79	79	79	79	79
R-squared	0.314	0.719	0.692	0.364	0.603	0.622

Table 8: **Sorting of High-Ability Runners (Performance).** *,**,*** denotes significance at the 10%, 5%, and 1% level, respectively. High-ability runners are defined as those with an (adjusted) finishing time within 1% (5%, 10%) of the race seasons's fastest time in their gender category. The dependent variable, "Sorting", is the proportion of high-ability runners who enter a major rather than a minor race. "Proportion of HA 1% (5%, 10%)", is the overall fraction of high-ability runners in the population of runners. Both variables are calculated separately for each race season (spring, autumn). For definition of other variables, see Table 7.

Variable	[1] Finish Time	[2] Finish Time	[3] Finish Time	[4] Finish Time
Steep Prize	-74.89**	-64.25	-44.83*	-32.50
	[33.913]	[60.663]	[26.386]	[48.065]
Total Prize ('00000)	-44.69***	-44.69***	-25.55***	-25.56***
	[6.843]	[6.845]	[5.290]	[5.292]
Opposition	-6.98***	-6.38**	-5.11***	-4.41*
	[1.136]	[2.822]	[0.892]	[2.346]
Opposition*Steep Prize		-1.50		-1.74
		[6.014]		[5.090]
Female	1,133.95***	1,134.01***	1,138.92***	1,139.00***
	[13.837]	[13.801]	[10.788]	[10.759]
Age	-0.08	-0.08	-0.06**	-0.06**
	[0.053]	[0.053]	[0.026]	[0.026]
At Home	17.58	17.56	-2.45	-2.47
	[14.251]	[14.254]	[10.909]	[10.913]
Nationality	48.63**	48.62**	34.08**	34.08**
	[24.642]	[24.646]	[16.173]	[16.162]
Rank (t-1)			4.38***	4.38***
			[0.124]	[0.124]
Constant	8,218.35***	8,211.02***	7,879.72***	7,871.22***
	[78.076]	[88.285]	[45.089]	[55.986]
Time Fixed Effects	Yes	Yes	Yes	Yes
Race Fixed Effects	Yes	Yes	Yes	Yes
Observations	4,452	4,452	4,452	4,452
R-squared	0.90	0.90	0.932	0.932

Table 9: **Individual Effort (Finishing Time).** *,**,*** denotes significance at the 10%, 5%, and 1% level, respectively. The dependent variable, "Finish Time", is the total number of seconds to finish a race. For definition of other variables, see Table 3.

Appendix

Equilibrium effort distributions

In our setting, players value a prize identically but differ in their marginal cost of effort. Since players are risk-neutral and effort costs are linear, the model is equivalent to a multi-unit all-pay auction, where the value of a unit to a player of type i is given by $v_i = b/c_i$.[23] The following analysis is, therefore, a direct adaptation of Clark and Riis (1998).

Suppose that N_H high-ability players and $N_L = N - N_H$ low-ability players have entered a contest offering M prizes of size b. Due to the deterministic nature of the contest, the equilibrium is necessarily in mixed strategies.[24] Moreover, in equilibrium, more than M players must be active–i.e., provide effort with positive probability. If the number of active players fails to exceed the number of prizes, each player can secure a prize by providing an arbitrarily small amount of effort. Finally, if a player is inactive, then there cannot exist an active player with a lower ability. This is because by imitating the strategy of the less-able player, the inactive player could secure a strictly positive payoff. If all players were different, these arguments would imply that in equilibrium the set of active players would consist of the $M+1$ most-able players. Since, in our setting with two types, some players have identical abilities, there also exist equilibria with more than $M+1$ active players. All equilibria are payoff-equivalent if *all* players are identical (Siegel, 2009) or competition is for a *single* prize (Baye et al., 1996). We select the equilibrium that is the only one to survive if we introduce arbitrarily small ability differences among players of the same type. In particular, we focus on the equilibrium in which exactly $M+1$ of the most-able players are active.

In the following, denote by $F_i(e)$ the cumulative effort distribution of an active player with type i. The description of equilibrium depends on the number of high-ability participants, N_H, and we will consider the cases where $N_H > M$, $N_H = M$, and $N_H < M$ separately.

When $N_H > M$, only high-ability players will be active. Since payoffs would become negative, no player will ever choose an effort higher than $\frac{b}{c}$. Instead, each player will choose

[23] In an M-unit all-pay auction, a bidder who bids x and values the object at v_i obtains the utility $v_i - x$ if his bid is among the M highest bids. Otherwise, his utility is $-x$. To match the auction with a perfectly discriminating contest, identify bids with efforts and multiply utilities by c_i.

[24] Suppose, some player i follows a pure strategy by choosing an effort level $e > 0$. In equilibrium, no other player would choose an effort in some small interval below e because by choosing an effort slightly higher, players can ensure that they win over i. But this implies that choosing e cannot be optimal for player i since he could reduce his effort without decreasing his chance of winning a prize.

an effort in $[0, \frac{b}{c}]$. Since, in equilibrium, each player has to be indifferent with respect to his potential effort choices, and payoffs are zero at both extremes of the support, it has to hold that $[1-(1-F_H(e))^M]b - ce = 0$ for all $e \in [0, \frac{b}{c}]$. Here $[1-(1-F_H(e))^M]$ is the probability that the player wins a prize, which happens unless all the other active players exceed his effort choice. Rearranging leads to the equilibrium distribution

$$F_H(e) = 1 - (1 - \frac{c}{b}e)^{\frac{1}{M}}. \tag{9}$$

In equilibrium, each of the $M+1$ active high-ability players expects a payoff of zero, making it optimal for all other players to remain inactive.

When $N_H = M$, all high-ability players and one of the low-ability players will be active. For the same reason as above, the support of F_L is given by $[0, b]$, and for all $e \in [0, b]$, it has to hold that $[1-(1-F_H(e))^M]b - e = 0$. It follows that

$$F_H(e) = 1 - (1 - \frac{1}{b}e)^{\frac{1}{M}}. \tag{10}$$

High-ability players are able to secure a prize by bidding b leading the payoff $b - cb$. Their indifference implies that $[1-(1-F_H(e))^{M-1}(1-F_L(e))]b - ce = b(1-c)$ for all $e \in [0, b]$. Substitution of F^H and rearranging gives

$$F_L(e) = 1 - c(1 - \frac{e}{b})^{\frac{1}{M}}. \tag{11}$$

High-ability players enjoy a comparative advantage over the low-ability contender due to their lower marginal cost of effort. Therefore, in equilibrium, high-ability players obtain a positive payoff. This explains why it cannot be the case that a low-ability player is active while a high-ability player is inactive.

Finally, consider the case $N_H < M$. As in the case where $N_H = M$, the maximal effort a player chooses with positive probability is the same for both types of players and equal to b. However, it can no longer be the case that a low-ability player chooses zero effort with positive probability. If this were the case, then another low-ability player would obtain a positive profit from choosing zero effort and would not be indifferent between the efforts in $[0, b]$. But if $F_L(0) = 0$ for all low-ability players, then zero effort can no longer be in the support of F_H. This is because high-ability players obtain the payoff $b(1-c) > 0$ from choosing $e = b$, but their payoff would be zero for $e = 0$. In fact, high-ability players will randomize over an interval $[\underline{e}, b]$ where the lower support $\underline{e} > 0$ is to be determined. The indifference of high-ability players on $[\underline{e}, b]$ implies that

$$[1-(1-F_H(e))^{N_H-1}(1-F_L(e))^{M+1-N_H}]b - ce = b(1-c). \tag{12}$$

Similarly, the indifference of low-ability players on $[\underline{e}, b]$ implies that

$$[1-(1-F_H(e))^{N_H}(1-F_L(e))^{M-N_H}]b-e=0. \tag{13}$$

Solving these two equalities simultaneously shows that for all $e \in [\underline{e}, b]$:

$$F_L(e)=1-c^{\frac{N_H}{M}}(1-\frac{e}{b})^{\frac{1}{M}} \quad \text{and} \quad F_H(e)=1-c^{\frac{N_H}{M}-1}(1-\frac{e}{b})^{\frac{1}{M}}. \tag{14}$$

The lower support of F_H can be determined from $F_H(\underline{e})=0$ to be $\underline{e}=b(1-c^{M-N_H})$. Finally, on $[0, \underline{e})$, F_L can be determined from $[1-(1-F_L(e))^{M-N_H}]b-e=0$, leading to

$$F_L(e)=1-(1-\frac{e}{b})^{\frac{1}{M-N_H}}. \tag{15}$$

This completes the characterization of the equilibrium effort distributions.

Proof of Proposition 1

The comparative statics with respect to the number and size of prizes follows from the fact that the cumulative effort distributions F_H derived in the previous section depend negatively on M and b. Hence, the effort distribution with more and/or higher prizes first-order stochastically dominates the effort distribution with fewer and/or lower prizes. To consider the effect of steepness, note that for a fixed prize budget, we can substitute M by $\frac{1}{b}$ in F_H. The resulting functions depend negatively on b, i.e. distributions that correspond to larger steepness (higher b) first-order stochastically dominate distributions that correspond to lower steepness.

It remains to show that (expected) efforts are minimized when $N_H = M$. Comparing the cumulative distribution functions for the cases $N_H > M$ and $N_H = M$, it holds that $F_H^{N_H>M}(e) < F_H^{N_H=M}(e)$ for all $e \in [0, b]$. Hence, the distribution of effort with $N_H > M$ first-order stochastically dominates the distribution of effort with $N_H = M$. In particular, expected effort is higher for $N_H > M$ than for $N_H = M$. To see that expected efforts are also higher for $N_H < M$, note that a high-ability player's likelihood of winning a prize is decreasing in N_H for all $N_H \leq M$. This follows from the fact that high-ability players have a greater chance of winning than low-ability players do. Since expected payoffs are constant in this range, it must be the case that expected effort is decreasing in N_H. ■

Proof of Proposition 2

It is immediate that $E[U_H]$ is increasing in b and M, but decreasing in p. To prove the last claim of Proposition 2, increase the steepness of the contest's prize structure by letting

$\tilde{M} < M$ and $\tilde{b} > b$, and consider

$$\begin{aligned} \frac{E[U_H] - E[\tilde{U}_H]}{1-c} &= bG(M,p) - \tilde{b}G(\tilde{M},p) \qquad (16) \\ &= b \sum_{N_H=0}^{M-1} \binom{N-1}{N_H} p^{N_H} (1-p)^{N-1-N_H} - \tilde{b} \sum_{N_H=0}^{\tilde{M}-1} \binom{N-1}{N_H} p^{N_H} (1-p)^{N-1-N_H} \\ &= b\text{Prob}(\tilde{M} \leq N_H \leq M-1) - (\tilde{b} - b)\text{Prob}(N_H \leq \tilde{M} - 1). \end{aligned}$$

The first term represents the advantage of the flatter prize structure. When the number of high-ability opponents N_H turns out to be between $\tilde{M}$ and $M - 1$, the flatter prize structure guarantees a positive payoff, b, whereas payoffs are zero for the steeper prize structure. The second term represents the advantage of the steeper prize structure. When the number of high-ability opponents is smaller or equal to $\tilde{M} - 1$, payoffs are positive for both prize structures, but the steeper prize structure offers an extra payoff $\tilde{b} - b > 0$. Note that the likelihood ratio $\text{Prob}(N_H \leq \tilde{M} - 1)/\text{Prob}(\tilde{M} \leq N_H \leq M - 1)$ is strictly decreasing in p. It converges to 0 for $p \to 1$ and to ∞ for $p \to 0$. Hence, there exists a $\bar{p} \in (0, 1)$ such that $E[U_H] - E[\tilde{U}_H] \geq 0$ if and only if $p > \bar{p}$. The steeper prize structure $(\tilde{M}, \tilde{b})$ guarantees a higher payoff if and only if the likelihood p with which opponents have high-ability is smaller than $\bar{p}$. The threshold $\bar{p}$ is decreasing in $M - \tilde{M}$ and increasing in $\tilde{b} - b$. ■

Proof of Proposition 3

In a contest in which an opponent has high-ability with probability p, let

$$E_p[N_H|N_H \leq M-1] = \sum_{m=0}^{M-1} \binom{N-1}{N_H} p^{N_H} (1-p)^{N-1-N_H} N_H \qquad (17)$$

denote the expected number of high-ability opponents conditional on their number being, at most, $M - 1$. Let

$$E_p[N_H] = p(N-1) \qquad (18)$$

denote the (unconditional) expected number of high-ability opponents. The equilibrium is determined by equation (4) with $p_S = 2hq_H$ and $p_W = 2h(1 - q_H)$. It follows from

$$\begin{aligned}\frac{dG(M,p)}{dp} &= \sum_{N_H=0}^{M-1} \tbinom{N-1}{N_H}[N_H p^{N_H-1}(1-p)^{N-1-N_H} - (N-1-N_H)p^{N_H}(1-p)^{N-2-N_H}] \\ &= \sum_{N_H=0}^{M-1} \tbinom{N-1}{N_H} p^{N_H-1}(1-p)^{N-2-N_H}[N_H - (N-1)p] \qquad (19) \\ &= \frac{G(M,p)}{p(1-p)}\{E_p[N_H|N_H \leq M-1] - E_p[N_H]\} < 0 \qquad (20)\end{aligned}$$

that

$$\frac{d\Delta}{dq_H} = 2h\left[b_S \frac{dG(M_S,p_S)}{dp} + b_W \frac{dG(M_W,p_W)}{dp}\right] < 0. \qquad (21)$$

The higher the fraction of high-ability players who choose contests of type S, the less willing are high-ability players to enter such contests. The fact that $b_S > b_W$ implies that

$$\Delta(q_H = 0) = b_S - b_W G(M_W, 2h) > 0. \qquad (22)$$

Hence, there cannot exist an equilibrium in which $q_H^* = 0$. Moreover,

$$\Delta(q_H = 1) = b_S G(M_S, 2h) - b_W. \qquad (23)$$

Note that $\Delta(q_H = 1)$ is strictly decreasing in h with $\Delta(q_H = 1) \to -b_W < 0$ for $h \to \frac{1}{2}$ and $\Delta(q_H = 1) \to b_S - b_W > 0$ for $h \to 0$. Hence, there exists a unique $\bar{h} \in (0, \frac{1}{2})$ such that $\Delta(q_H = 1) \geq 0$ if and only if $h \leq \bar{h}$. Therefore, an equilibrium in which $q_H^* = 1$ exists if and only if $h \leq \bar{h}$. Moreover, the equation $\Delta(q_H^*) = 0$ has a solution $q_H^* \in (0,1)$ if and only if $h > \bar{h}$. This solution and, hence, the equilibrium is unique. To determine how q_H^* depends on h for $h > \bar{h}$, consider

$$\begin{aligned}h\frac{d\Delta}{dh} &= \left[b_S p_S \frac{dG(M_S,p_S)}{dp} - b_W p_W \frac{dG(M_W,p_W)}{dp}\right] \qquad (24) \\ &= \frac{b_S G(M_S,p_S)}{1-p_S}\{E_{p_S}[N_H|N_H \leq M_S - 1] - E_{p_S}[N_H]\} \qquad (25) \\ &- \frac{b_W G(M_W,p_W)}{1-p_W}\{E_{p_W}[N_H|N_H \leq M_W - 1] - E_{p_W}[N_H]\}.\end{aligned}$$

For p_S and p_W such that $\Delta = 0$, we can substitute $b_S = b_W \frac{G(M_W,p_W)}{G(M_S,p_S)}$ to get

$$\begin{aligned}\frac{h}{b_W G(M_W,p_W)}\frac{d\Delta}{dh} &= \frac{1}{1-p_S}\{E_{p_S}[N_H|N_H \leq M_S - 1] - E_{p_S}[N_H]\} \qquad (26) \\ &- \frac{1}{1-p_W}\{E_{p_W}[N_H|N_H \leq M_W - 1] - E_{p_W}[N_H]\}.\end{aligned}$$

It is one of the properties of the binomial distribution that the difference between the unconditional and the tail conditional means increases more strongly than linearly in the underlying probability p (Johnson et al., 1992). Thus, the first term is strictly decreasing in p_S. Since, for $q_H^* \geq \frac{1}{2}$, it holds that $p_S \geq p_W$, we can find an upper bound for the first term by setting $p_S = p_W$ to get

$$\frac{h(1-p_W)}{b_W G(M_W, p_W)} \frac{d\Delta}{dh} \leq E_{p_W}[N_H | N_H \leq M_S - 1] - E_{p_W}[N_H | N_H \leq M_W - 1] < 0. \quad (27)$$

The last inequality follows from $M_S < M_W$. Thus, we have shown that at any equilibrium such that $q_H^* \geq \frac{1}{2}$ and, hence, $p_S^* \geq p_W^*$, it holds that $\frac{d\Delta}{dh}|_{q_H = q_H^*} < 0$. Together with $\frac{d\Delta}{dq_H} < 0$, this implies that q_H^* is strictly decreasing in h as long as $q_H^* \geq \frac{1}{2}$. This also means that once q_H^* has crossed $\frac{1}{2}$ from above, it will stay below $\frac{1}{2}$ for all higher values of h. In other words, there exists a $\bar{\bar{h}} \in (\bar{h}, \frac{1}{2}]$ such that $q_H^* \leq \frac{1}{2}$ for all $h \geq \bar{\bar{h}}$. ■

References

[1] Azmat, G., Möller, M. (2009) "Competition Amongst Contests." *RAND Journal of Economics* **40**, 743-768.

[2] Baye, M., Kovenock, D., De Vries, C., G. (1996) "The All-Pay Auction with Complete Information." *Economic Theory* **8**, 291-305.

[3] Bognanno, M., L. (2001) "Corporate Tournaments" *Journal of Labor Economics* **19**(2), 290-315.

[4] Brückner, M., Ciccione, A. (2010) "Rain and the Democratic Window of Opportunity." *Econometrica*, forthcoming.

[5] Clark, D., J., Riis, C. (1998) "Competition over more than one Prize." *American Economic Review* **88**, 276-289.

[6] Cohen, C., Sela, A. (2008) "Allocation of Prizes in Asymmetric All-Pay Auctions." *European Journal of Political Economy* **24**, 123-132.

[7] Damiano, E., Hao, L., and Suen, W. (2010) "First in Village or Second in Rome?" *International Economic Review* **51**(1), 263-288.

[8] Damiano, E., Hao, L., and Suen, W. (2013) "Competing for Talents." *Journal of Economic Theory* forthcoming.

[9] Deaton, A., (1999) "Commodity Prices and Growth in Africa." *Journal of Economic Perspectives* **13**(3), 23-40.

[10] Dohmen, T., Falk, A. (2011) "Performance Pay and Multidimensional Sorting: Productivity, Preferences, and Gender." *American Economic Review* **101**, 556-590.

[11] Ehrenberg, R., G, Bognanno, M., L. (1990) "Do Tounaments have Incentive Effects?" *Journal of Political Economy* **98** (6), 1307-1324.

[12] Eriksson, T. (1999) "Executive Compensation and Tournament Theory: Empirical Tests on Danish Data." *Journal of Labor Economics* **17**(2), 262-280.

[13] Eriksson, T., Teyssier, S., Villeval, M. (2009) "Self-Selection and the Efficiency of Tournaments." *Economic Inquiry* **47**(3), 530-548.

[14] Fibich, G., Gavious, A., Sela, A. (2006) “All-Pay Auctions with Risk Averse Players.” *International Journal of Game Theory* **34**, 583-599.

[15] Johnson, N. L., Kotz S., Kemp, A. (1992) *Univariate Discrete Distributions*, John Wiley & Sons.

[16] Konrad, K. A. (2009) *Strategy and Dynamics in Contests*, Oxford University Press.

[17] Krishna, V., Morgan, J. “The Winner-Take-All Principle in Small Tournaments.” *Advances in Applied Microeconomics*, Vol. 7 (1998), pp. 61-74.

[18] Lazear, E., P. (2000) “Performance Pay and Productivity.” *American Economic Review* **90**(5), 1346-1361.

[19] Lazear, E., P., Rosen, S. (1981) “Rank-Order Tournaments as Optimal Labor Contracts.” *Journal of Political Economy*, **89**, pp. 841-864.

[20] Leuven, E., Oosterbeek, H., Sonnemans, J., Van der Klaauw, B. (2011) “Incentives versus Sorting in Tournaments: Evidence from a Field Experiment.” *Journal of Labor Economics*, **29**(3), pp. 637-658.

[21] Miguel, E., S. Satyanath, and E. Sergenti (2004) “Economic Shocks and Civil Conflict: An Instrumental Variables Approach.” *Journal of Political Economy*, **112**, 725-753.

[22] Moldovanu, B., Sela, A. (2001) “The Optimal Allocation of Prizes in Contests.” *American Economic Review* **91**(3), 542-558.

[23] Moldovanu, B., Sela, A. (2006) Contest Architecture. *Journal of Economic Theory* **126**(1), 70-97.

[24] Noakes, T. (1985) *The Lore of Running*, Oxford University Press, Oxford.

[25] Siegel, R. (2009) “All-Pay Contests.” *Econometrica* **77**(1), 71-92.

[26] Yun, J. (1997) “On the Efficiency of the Rank-Order Contract under Moral Hazard and Adverse Selection.” *Journal of Labor Economics* **15**(3), 466-494.