The Distribution of Talent across Contests

Ghazala Azmat *
and
Marc Möller †

Abstract

Do the contests with the largest prizes attract the most able contestants? Do contestants avoid competition? In this paper we show that the distribution of abilities plays a crucial role in determining contest choice. Positive sorting exist only when the proportion of high ability contestants is sufficiently small. As this proportion increases, contestants shy away from competition and sorting decreases. Eventually, contests with smaller prizes attract stronger participants, i.e. there exists negative sorting. 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 between prizes and sorting.

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

*Department of Economics and Business, University Pompeu Fabra, Ramon Trias Fargas 25-27, 08005 Barcelona, Spain. Email: ghazala.azmat@upf.edu, Tel.: +34-93542-1757, Fax: +34-93542-1766.

†Department of Economics, University of Bern, Schanzeneckstrasse 1, 3001 Bern, Switzerland. Email: marc.moeller@vwi.unibe.ch, Tel.: +41-(0)31631-8078.

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 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.

It has been shown both theoretically (Clark and Riis (1998), Moldovanu and Sela (2001)), as well as empirically (Ehrenberg and Bognanno (1990)), that effort choices are sensitive to the size and the allocation of a contest's prize budget. Competition is especially intense in contests that offer large but few prizes. For example, working hours of up to 80 h/week are common practice at McKinsey & Company, whose up–or–out policy promises large wage increases but entails that "25% of the firm is new every year."[1] In the US pharmaceutical industry, where the benefits from patent protection are substantial but restricted to the few drugs that obtain approval, annual R&D costs have been reported to exceed 25 billion US$ (DiMasi et al. (2003)).

While the relationship between prizes and effort seems to be well understood, little is known about their influence on contest selection. Other incentive schemes, such as piece rates, have been praised for their capacity to screen workers according to their ability (Stiglitz (1975) and Lazear (1986)). Moreover, the productivity gains associated with the self–selection of the most able workers have been shown to be as important as those related to incentive effects (Lazear (2000)). Several studies have therefore emphasized the importance of contest design for the attraction of high quality participants. For example, Burguet and Sakovics (1999) argue that auctions may use their reserve prices to screen buyers according to their valuations. Similarly, Fullerton and McAfee (1999) propose entry

[1]Interview with former managing director, Rajat Gupta, Academy of Management Executive, 2001, Vol. 15, No. 2.

fees as a means to improve the quality of participants in a research tournament. Since prizes are an inseparable feature of any competitive environment, an important question is whether they serve a similar purpose. In this paper we consider, both theoretically as well as empirically, whether prizes induce contestants to sort according to their abilities.

One might expect that the contest with the largest prize attracts the most able contestants. However, there are several reasons why this intuition may fail to hold. First, contests with smaller prizes might be preferred since they induce lower effort costs. This phenomenon has been observed in the pharmaceutical industry where an increasing number of companies, including the biggest, have started to develop generic drugs for which benefits are smaller but approval is less costly to obtain.[2] Second, contests with a higher number of prizes might be preferred since they offer a greater chance of success. This is in line with a flattening of corporate hierarchies documented by Rajan and Wulf (2006) which can be interpreted as the companies' attempt to win the "war for talent" by offering workers a higher chance of retainment (Michaels et al. (2001)). Finally, unlike in a standard screening framework (Mirrlees (1971), Spence (1973)), contest choice constitutes a *strategic* decision since contestants have an interest to avoid strong opponents.

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 amongst 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 marginal costs of effort than low ability contestants. Strong contests offer greater but fewer prizes than weak contests and therefore induce higher efforts, i.e., stronger competition. 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

[2]For example, through its generic unit Greenstone, Pfizer develops generic versions of existing drugs, including its own. See "More Generics Slow Rise in Drug Prices" New York Times, August 8, 2007.

ability contestants become sufficiently frequent, weak competition attracts an even larger share of high ability contestants than strong competition. Hence the common perception that larger prizes attract stronger competitors fails to hold in general.

We take advantage of an unusually clean opportunity to empirically investigate the extent of sorting across contests. With around 20,000 observations, we examine contest choice of professional marathon runners over three decades. The set-up allows us to abstract from a number of identification problems present in other types of data. For example, while in a labor-market setting it is often impossible to disentangle between firm and worker types and to abstract from complementarities in team–work, in marathons individual performance is readily available. In addition, there are three features that make marathons the ideal setting to test our model. First, marathon running is strongly dominated by a small group of highly talented East-African runners, mainly from Kenya and 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. Second, five Major marathons (Berlin, Boston, Chicago, London, and New York), offer more than 50% of the total available prize money but also the lowest chances of success. 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. Finally, due to the abolishment of the amateur rule by the International Olympic Committee, marathons started to offer prize money in the mid–1980s. This led to a substantial increase in the participation of Kenyan and Ethiopian runners and hence, altered the overall distribution of abilities.

In accordance with our theory, we find that the likelihood of an elite runner to participate in one of the Major marathons is increasing in the race's prize budget and decreasing in the expected number of high ability opponents, measured by the number of Kenyan and Ethiopian participants in last year's race. Following Brückner and Ciccone (2010), we use exogenous variation in local economic conditions to predict participation of Kenyan and Ethiopian runners. Our analysis allow us to determine the "price" that contestants assign to opposition. Our estimates show that elite runners are willing to forgo potential prize winnings of $25,500$$ for each high ability opponent they are able to avoid.

In line with our main theoretical result, we find that, as the number of high ability participants increases, they become more likely to avoid competition. In particular, as the share of Kenyan and Ethiopian runners increases by 10 percent, the fraction of high ability runners who choose to participate in Major races falls by 10.3 percent. According to our estimates, Major races would have to increase their prize budgets by 8.8 percent to maintain their attractiveness to high ability runners.

These results constitute the first evidence for tournament selection effects. Previous studies have focused on the choice between tournaments and alternative incentive schemes using experimental setups. For example, Dohmen and Falk (2011) report the results of a real–effort experiment in which subjects choose between a pairwise tournament and a fixed payment. They find that apart from having higher ability, subjects who choose a tournament have lower degrees of risk aversion and a more optimistic self assessment. Eriksson et al. (2009) reports similar results for the choice between a pairwise tournament and piece–rates. Our results complement these findings by considering the frequently encountered choice between tournaments of differing prize structure.

2 The model

We consider a continuum of contests with mass one. Each contest allows for $N \geq 3$ participants. There are two types of contests $j \in \{S, W\}$. For reasons explained below they are denominated as strong contests (S) and weak contests (W). A contest of type j offers $M_j \in \{1, 2, \ldots, N-1\}$ identical prizes of size $b_j > 0$ and a performance–independent (i.e. fixed wage) payment $w_j \geq 0$ to each of its participants.[3] In a labor tournament

[3]The assumption that a contest's prizes are all identical makes the model tractable. A general description of competition for the case of $N \geq 3$ heterogeneous players and $M > 1$ non–identical prizes is

setting, w_j could represent the workers base wage, while b_j measures the wage increase or bonus for those who become promoted.

Contests of type S award higher ($b_S > b_W$) but fewer ($M_S < M_W$) prizes than contests of type W.[4] Apart from differences in their payment structures, contests are assumed to be identical. For simplicity we assume that both types of contests exist in equal fractions. Our results remain qualitatively unchanged when this assumption is relaxed.

There is a continuum of risk–neutral players with mass N.[5] Players differ with respect to their constant marginal cost of effort c. There are two types of players, $i \in \{L, H\}$. A high ability player's marginal cost of effort is $c_H > 0$, while low ability players have marginal cost $c_L > c_H$. A fraction $h \in (0, 1)$ of players has high ability and the distribution of abilities is common knowledge amongst players.

In each contest, players compete for prizes by exerting effort. We follow an extensive literature on contest design (see for example Clark and Riis (1998) or Moldovanu and Sela (2001, 2006)) by assuming that contests are perfectly discriminating. This means that in each contest, prizes are awarded to the players with the highest levels of effort while ties are broken randomly.[6] A player of type $i \in \{L, H\}$ who exerts effort $e \geq 0$ in a contest of type $j \in \{S, W\}$ will receive the payoff $U_i^j = w_j + b_j - c_i e$ if he wins one of the M_j prizes and $U_i^j = w_j - c_i e$ otherwise.

The model has two stages. In the first stage, players simultaneously choose which type of contest to enter. Once they have entered a contest, players observe the abilities of their opponents. In the second stage, players compete by simultaneously selecting their effort levels.

When the number of players who choose a particular type of contest exceeds the

still missing. A first step into this direction has been made by Cohen and Sela (2008).

[4]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 high prizes and a tournament with few small prizes.

[5]The implications of risk aversion are discussed in Section 4.

[6]Alternatively, winners could be determined stochastically, i.e. in dependence of efforts and random factors. For a discussion of this case see Section 3.

number of available slots players need to be rationed. As a rationing rule we assume that each contest accepts as many high ability players as possible and fills any remaining slots with low ability players.[7] We show below that, in equilibrium, high ability players are never rationed, i.e. their allocation is driven entirely by preferences and not by the rationing rule. Those players who were turned down by the contest of their choice enter a contest of the other type. This is optimal since players have zero outside options and expected payoffs are strictly positive in both types of contest.

3 Individual contest choice

In this section we determine the players' expected payoff from participating in a contest in dependence of the contest's payment/prize structure and its set of participants. Since the total number of players matches the total number of contest slots, in equilibrium each contest will have N participants. Hence in a contest of type j, N players will compete for M_j identical prizes. Players value a prize identically at b_j but differ in their marginal costs of effort c_i. Since players are risk–neutral and effort costs are linear, the model is equivalent to a multi–unit all–pay auction where bidders have identical costs but differ in the value $v_i = b_j/c_i$ they attach to obtaining a unit.[8]

Clark and Riis (1998) show that the equilibrium of an all–pay auction with heterogeneous players and M identical prizes is necessarily in mixed strategies. This equilibrium is unique when all players have different valuations. When some players' valuations are identical, multiple equilibria might exist but equilibria are payoff–equivalent (see Baye, Kovenock, and de Vries (1996)). When players are ordered according to their valuations, i.e. $v_n \geq v_m$ for all $n < m$, then expected payoffs are $v_n - v_{M+1}$ for the players with the

[7] Note that the above assumption requires organizers to observe the players' abilities. While in a sports context the athletes' abilities can be deducted from past performances, in the labor market entry tests or interviews are typically employed to select the most able amongst the applicants.

[8] In an M–unit all–pay auction a bidder who bids x_i and values the object at v_i obtains the utility $v_i - x_i$ if his bid is amongst the M highest bids. Otherwise his utility is $-x_i$. To match the auction with our contest identify bids with efforts and multiply utilities by c_i.

M highest valuations and zero for all other players.

This result has the following implications for our model. In a contest of type j, players may attach two different values to obtaining one of the M_j prizes. A low ability player assigns the value $v_L = b_j/c_L$ whereas a high ability player has valuation $v_H = b_j/c_H > v_L$. A high ability player therefore expects a payoff (net of performance independent payments) equal to $c_H(v_H - v_L) = b_j(1 - \frac{c_H}{c_L})$ if the number of high ability players is smaller or equal to the number of prizes and zero otherwise. For low ability players, expected payoffs can never exceed the contest's performance independent payment. The following lemma summarizes these findings.

Lemma 1 *Suppose that H_j high ability players and $N - H_j$ low ability players participate in a contest of type $j \in \{S, W\}$. A high ability players's expected payoff is $E[U_H^j|H_j] = w_j + b_j(1 - \frac{c_H}{c_L})$ if $H_j \leq M_j$ and $E[U_H^j|H_j] = w_j$ if $H_j > M_j$. A low ability player's expected payoff is $E[U_L^j|H_j] = w_j$ irrespective of H_j.*

The fact that the expected payoffs of low ability players are independent of prizes is due to our assumption that contests are perfectly discriminating. Low ability players choose positive effort levels and win a prize with positive probability but their expected prize winnings are exactly compensated by their effort costs. If contests involved a random element then expected payoffs of low ability players would depend on prizes but this dependence would still be weaker than for high ability players. This difference can be understood as a Spence–Mirrlees single–crossing condition which reflects the fact that high ability players have a stronger preference for prizes (as opposed to effort cost savings) than low ability players. It gives rise to the possibility of sorting. Note however, that this interpretation requires a player's set of opponents to be fixed, i.e. the players' strategic interaction is neglected. Also note that the players' incentive to sort is strongest when contests are perfectly discriminating. Hence our insight that the most competitive prize structures may fail to attract the most able contestants will extend to the case where contest outcomes are random. For a detailed investigation of the relationship between a contest's prize structure and its randomness see Azmat and Möller (2009).

We are now ready to describe a player's individual preferences at the time of contest choice. From Lemma 1 it immediately follows that for a low ability player the expected payoff from entering a contest of type i is independent of the set of opponents and given by $E[U_L^j] = w_j$. Hence low ability players simply prefer the contest with the highest performance independent payment, w_j, and are indifferent when $w_S = w_W$.

Next consider high ability players. Let p_j denote the likelihood with which an opponent in a contest of type j has high ability. The probability with which a high ability player obtains a payoff in excess of w_j in contest j equals the probability with which the player meets at most $M_j - 1$ high ability opponents. It is given by

$$G(M_j, p_j) \equiv \sum_{m=0}^{M_j-1} \binom{N-1}{m} (p_j)^m (1-p_j)^{N-1-m}. \tag{1}$$

A high ability player's expected utility from entering a contest of type j is given by

$$E[U_H^j] = w_j + b_j(1 - \frac{c_H}{c_L}) G(M_j, p_j). \tag{2}$$

It depends on the contest's overall prize budget via w_j, the allocation of prizes via M_j and b_j, and the (expected) strength of his opponents represented by p_j. In the Appendix we prove the following intuitive result:

Proposition 1 *A high ability player's expected payoff from entering a contest of type j, is increasing in the performance independent payment w_j, and the number M_j and size b_j of prizes, but decreasing in the probability p_j with which opponents have high ability. Payoffs are increasing in the steepness of contest j's prize structure when $p_j < \bar{p}_j$ but decreasing when $p_j > \bar{p}_j$.*

While the first claim of Proposition 1 is straight forward, the second claim requires some explanation. Suppose that we increase the steepness of contest j's prize structure by raising b_j and lowering M_j. Contest j then awards higher but fewer prizes. Hence winners earn higher rewards but competition becomes fiercer leading to higher effort costs. When the probability to meet high ability opponents is small, high ability players prefer higher (though fewer) prizes due to their comparative advantage over low ability players. In contrast, when the probability to meet high ability opponents is large, high ability players prefer more (though smaller) prizes due to their mitigating effect on competition and the resulting decrease in effort costs.

4 Distribution of talent

What do the players' individual preferences imply for the equilibrium allocation of talent? Since players make their choice contingent on the expected abilities of their opponents, a player's contest choice depends directly on the choices of all other players. This distinguishes the present model from standard models of sorting where the choices of other players matter only indirectly, i.e. via their influence on the beliefs about the players' types. In the last section we saw that in our setup, the contest choice of low ability players depends exclusively on the contests' performance independent payments. In particular, the choice of low ability players is independent of the behavior of high ability players. This allows us to concentrate on the contest choice of high ability players, making the model tractable.

Suppose that a fraction $q \in [0, 1]$ of the high ability players choose strong contests while the remaining fraction $1 - q$ choose weak contests. If both fractions, q and $1 - q$, are sufficiently small, i.e. if $\max\{q, (1-q)\}hN < \frac{N}{2}$, then all high ability players are able to enter the contest of their choice. The probability with which a slot of type S is filled with a high ability player is then given by $p_S = 2hq$ while a slot of type W is filled with probability $p_W = 2h(1 - q)$. If instead $\max\{q, (1-q)\}hN \geq \frac{N}{2}$ then high ability players would exhibit excess demand for one type of contest. In this case $p_S = 1$ and $p_W = 2h - 1$

or vice versa. The contest choice of high ability players is determined by the difference between their expected payoff from entering a strong contest and their expected payoff from entering a weak contest. From (2) this difference is proportional to

$$\Delta \equiv b_S G(M_S, p_S) - b_W G(M_W, p_W) + \frac{w_S - w_W}{1 - \frac{c_H}{c_L}}. \tag{3}$$

High ability players strictly prefer a contest of type S (W) when $\Delta > 0$ ($\Delta < 0$) and are indifferent when $\Delta = 0$. We are now able to state our main result:

Proposition 2 *If contests offer identical performance independent payments ($w_S = w_W$) then in the unique equilibrium a fraction q^* of high ability players enter strong contests and the following holds: $q^* = 1$ for all $h \in (0, \bar{h}]$ where $\bar{h} < \frac{1}{2}$; $q^* \in (\frac{1}{2}, 1)$ and strictly decreasing in h for all $h \in (\bar{h}, \bar{\bar{h}})$; $q^* \leq \frac{1}{2}$ for all $h \in [\bar{\bar{h}}, 1)$. An increase in w_S, M_S, or b_S and a decrease in w_W, M_W, or b_W all lead to an upward shift in q^*.*

One may expect that contests which offer higher but fewer prizes should be more attractive to high ability players. Proposition 2 shows that this intuition fails to hold in general. The equilibrium allocation of talent depends on the overall distribution of talent within the population of players as can be seen in Figure 1.

When the fraction h of high ability players is small, i.e. $h \leq \bar{h}$, then *all* high ability players choose strong competition, i.e. there is *complete* (positive) sorting of abilities. For intermediate values of h, i.e. $\bar{h} < h < \bar{\bar{h}}$, high ability players are still more likely to choose strong competition than weak competition but sorting is only *partial* and strictly decreasing in h. When h is large, i.e. $h \geq \bar{\bar{h}}$, strong competition attracts *less* high ability players than weak competition, i.e. sorting is reversed (negative). Note that $\bar{h} < \frac{1}{2}$ implies that complete sorting can never occur when high ability players are equally frequent as low ability players. Also note that in the equilibrium described by Proposition 2, high ability players are never rationed. If in one type of contest, all slots would be filled with high ability players, then all high ability players would strictly prefer the other type of contest. Only in the limit, as $h \to 1$, both types of contests become filled by high ability participants and $q^* \to \frac{1}{2}$.

The intuition for Proposition 2 is as follows. Strong contests offer high potential prizes while weak contests mitigate competition and are characterized by low effort costs. From the viewpoint of a high ability player, effort considerations become more important as the likelihood to meet high ability rivals increases and the comparative advantage over low ability players becomes less likely to play a role. When high abilities become sufficiently frequent, the mitigation of competition becomes so valuable that high ability players prefer weak contests over strong contests even though rivals in the former are expected to be more able than rivals in the latter.

An increase in w_S, M_S, or b_S, or a decrease in w_W, M_W, or b_W, raises the payoff that players expect in contests of type S relative to type W. This leads to an upward shift in q^*. As can be seen from Figure 1, the range where sorting is complete becomes larger and q^* increases wherever $q^* < 1$.

Finally, let us discuss the possible influence of risk aversion on the players' contest choice. From the viewpoint of a high ability player, each type of 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 $q > \frac{1}{2}$, 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. As a consequence, q^* can be expected to be lower, i.e. risk aversion leads to a decrease in sorting. Note that this discussion ignores the fact that risk aversion will also influence the way in which players compete. It has been shown for example, that (in a contest with a single prize) risk aversion decreases the effort of low ability contestants but increases the effort of high ability contestants (Fibich et al. (2006)). Since weak contests reduce effort costs by mitigating competition, we therefore contemplate that, as before, risk aversion makes weak contests become more attractive for high ability players (and less attractive for low ability players). A thorough investigation of the effect of risk aversion would require an extension of the work of Clark and Riis (1998) to the case of risk averse players and is beyond the present analysis.

5 Empirical Framework

In this section we will test the predictions of the model using a large panel dataset of international city marathons. Testing the model requires a setting in which individual abilities are observable and the distribution of abilities is subject to changes. In addition, prize structures should be known and should differ across contests. Marathon data offers several advantages over, for example, data on labor tournaments. While prizes and performance are easily observed in marathons, a firm's pay–structure and a worker's individual performance are hardly available. While marathons are fairly homogeneous in their setup, firms differ in dimensions other than their pay–structure.[9] Finally, while professional runners choose two marathons per year, employment relations are established less frequently, making equilibrium behavior less likely to emerge. Beyond these advantages, there are three important reasons for why marathons in particular constitute the ideal setting to test our theory.

First, a surprisingly high fraction of the best marathon runners are of East–African origin. In 2009, 62 of the 100 fastest (male) marathon runners were Kenyan and 26 were Ethiopian.[10] In the same year, more than 70 percent of the available prize money was won by Kenyan or Ethiopian runners. This dominance, unparalleled in other sports, has been explained by genetic, social, nutritional, and geographical factors (Noakes, 1985). It allows us to identify the most able contestants by origin, which, unlike past performance,

[9] Some marathons have faster (flatter) race courses than others. To make marathons comparable, we adjust all finishing times using a conversion factor constructed by the Association of Road Running Statisticians. This is done throughout the entire analysis.

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

is independent of prize and effort considerations.

Second, East–African runners were hardly present in international marathons until the mid 1980s. Running had become hugely popular in East–Africa in 1960, when the Ethiopian marathon runner Abebe Bikila became the first African to win an Olympic gold medal. However, due to the amateur rule of the International Olympic Committee, runners were not allowed to compete for money and city marathons refrained from offering money prizes until the abolishment of the rule in 1986. As a result, East–African runners participated almost exclusively in Olympic games and World–Championships. In the 1980s, as a response to the abolition to the amateur rule, city marathons started to award money prizes and prize budgets have increased ever since (see Figure 2). Today, with winning prizes well above 100,000$, running means big money to athletes from East–Africa. For example, in 2009 the per capita income in Kenya was approximately 1,000$ per annum. Not surprisingly, the number of East-African runners that compete internationally has increased steadily since the mid 1980s. In light of our theoretical model, this change can be interpreted as an increase in the fraction of high ability contestants. It has made the sport more competitive by decreasing the gap between winners and losers. This can be seen in Figure 3 which depicts the ratio of the fastest race time 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 5 percent, this advantage has decreased to less than 3 percent in the late 2000s.

The third important feature of marathon running is the fact that five races have obtained a special status comparable to the Grand Slam tournaments in tennis. The marathons in Berlin, Boston, Chicago, London, and New York have the longest tradition, the highest prize budgets, and the largest number of participants. Collectively, the group annually attracts more than 5 million on-course spectators, 250 million television viewers, and 150,000 participants. Its economic impact has been claimed to lie above

400$ million.[11] As can be seen from Figure 2, the five marathons award more than 50 percent of the total prize money available in the 35 races in our dataset. In 2006 they launched the World Marathon Majors, a new point ranking offering a $1 million prize to the best performing runner of the series. In the following they will therefore be referred to as "Major" marathons while all other races will be denoted as "Minor" marathons. Major marathons are not only characterized by high prizes but also by a high number of runners competing for each prize. A marathon runner therefore faces the trade-off that is at the heart of our theoretical model: Participate in a Major marathon where prizes are high but competition is strong or choose a Minor marathon where prizes are low but competition is weak.

As a brief preview of our results, Figure 4 depicts the distribution of East–African runners across the two race categories. In order to compare with the predictions of the theoretical model depicted in Figure 1, we focus on the ten most important marathons (five Major, five Minor). As the overall proportion of East–African participants increases, the share which chooses a Major rather than a Minor marathon decreases. Beyond a certain fraction of high ability runners, sorting is reversed, i.e. we observe negative sorting.

5.1 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.[12] We restrict attention to the 35 most relevant marathons.[13] 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

[11] For more details see http://worldmarathonmajors.com/US/about/.

[12] We are grateful to Ken Young from the Association of Road Racing Statisticians for kindly providing us with the data.

[13] 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.

times. For each race, we observe the date, location, as well as 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 less than twenty prizes, for each race our data therefore contains 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 separately show the descriptive statistics 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 $26,371$\$). However, they also have around three times as many competitors ($22,332$ compared with $6,838$).[14] In our theoretical model contests differ in the number of prizes, M_j, but have the same number of competitors, N. What matters for the contest choice of a high ability contestant is the likelihood to be amongst the M_j most able competitors. Since most marathons award less than 10 prizes, but the number of competitors is much larger in Major marathons, the likelihood to be amongst the 10 most able runners is lower in a Major marathon. Major marathons offer higher prizes but the chances of winning are lower. We can therefore identify Major races as strong contests and Minor races as weak contests. Further motivation for this identification can be obtained by considering the allocation of prize budgets. In contrast to our theoretical model, the prizes awarded by a marathon are not identical but decreasing in rank. This decrease turns out to be steeper in Major races. 57% of the Major races have a prize allocation that is steeper than the average compared to only 35% for Minor races.[15]

The two types of races also differ in the quality of the runners they attract. From

[14]These numbers include amateur runners but the comparison of the size of the elite fields is similar.

[15]Steepness is measured by the Herfindahl-Hirschman index, calculated for the top three prizes.

Table 1 we can see that, on average, over the years the fraction of high ability runners has been considerably larger in the Major races. This holds no matter whether 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 5 percent of the year's best, compared with only 8 percent in the Minor races. As a consequence, winning times in Major races are on average 8 minutes faster which is equivalent to a 2.6km lead. Part of this difference can be explained by the fact that, in accordance with the model, the prizes offered by a Major race induce higher effort levels. The remaining part is due to selection effects, which will be the focus of our analysis.

Table 1 also compares the descriptive statistics for East–African and Non–East–African runners. For male runners, the comparison shows that runners from East–Africa are faster (2:14 compared to 2:17) and win higher prizes ($8,307$\$ compared to $3,360$\$) than runners from other origins. 28 percent of East African runners have (adjusted) finishing times within 5% of the years's fastest time, compared to only 21 percent for Non–East–African runners. For female runners the differences are even larger. These numbers lend support to our identification of East–African runners as high ability contestants.

5.2 Individual contest choice

5.2.1 OLS Analysis

To test Proposition 1, 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:

$$P_{ijt} = \alpha_0 + \alpha_A A_{jt-1} + \alpha_B B_{jt} + \alpha_S S_{jt} + X_i \beta + \varepsilon_{ijt}. \tag{4}$$

Since we are interested in the runners' choice between entering a Major and a Minor race, we let P_{ijt} take the value 1 whenever j is a Major race and the value 0 otherwise. The variable A_{jt-1} denotes the proportion of East–African runners amongst the race's top twenty finishers in the previous year. Due to the dominance of East–African runners, A_{jt-1} serves as a measure of the level of opposition to be expected. The variable B_{jt} denotes the marathon's total prize budget. S_{jt} is a dummy variable whose value is 1 when the race's prize structure is steeper than the average. We also include a vector of control variables, X_i, containing the runner's age, nationality, gender, and ranking in the previous year, and 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 trends and race fixed effects.[16]

According to Proposition 1, P_{ijt} should be increasing in B_{jt} and decreasing in A_{jt-1}. The predictions of Proposition 1 with respect to the steepness of the contest's prize structure are more complicated. S_{jt} should have a positive effect on P_{ijt} when A_{jt-1} is relatively small but a negative effect when A_{jt-1} is relatively large. For this reason, we will also look at the interaction of S_{jt} with A_{jt-1}.

Since Proposition 1 is concerned with the preferences of high ability contestants, we restrict attention to the race choice of the top ranked runners. However, since many of these runners are East–African they could have also been contained in A_{jt-1}. In order to avoid the resulting endogeneity problem, we restrict the analysis to the Top100 male and Top100 female runners in a given year with origins different from Kenya or Ethiopia.[17] In particular, we estimate whether an increase in the fraction of East–African runners in race j in the previous year, reduces the likelihood with which a Top100 Non–East–African runner enters race j in the current year.

In Table 3, we present the results. Column 1 and 2 contain the results with and

[16] The results remain unchanged if year dummies are included in place of a linear time trend.
[17] Our results are robust with respect to changes in the cut–off point for our definition of "high-ability".

without controls, respectively. Column 3 includes trends and race fixed–effects. Overall, we find that an increase in expected opposition, leads to a significant decrease in the entry of high ability contestants. This persists in all specifications. Total prize money has a strong and positive effect on entry. We postpone the discussion of the size of these effects until the instrumental variable analysis below.

5.2.2 IV Analysis

An important concern is that 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. This translates 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}. 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.[18] 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 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, it has been documented

[18] This is preferable to using the countries' GDP as an instrument since GDP is subject to world trends.

by Deaton (1999) that commodity price downturns cause rapidly worsening economic conditions in Sub-Saharan African economies. We therefore 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, in particular the younger ones, 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 cover the period 1986 to 2009 and is taken from the NASA Global Precipitation Climatology Project.

We may also be concerned that race organizers adjust the total prize budget, B_{jt}, to keep their race attractive for high ability contestants. If entry falls, race organizers may increase prize money. As a consequence the coefficient on B_{jt} will 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.[19]

In Column 1 of Table 4, 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

[19] This basket contains U.S. Dollar, Euro, Japanese Yen, and Pound 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.

runners competing internationally. In Column 2, we see that, as predicted, exchange rates are strongly related to total prize money. Both (sets) of instruments are strong, with high F-statistics.

In Table 5, we present the results for the IV estimates. As in the OLS regression, we find that entry is negatively affected by expected opposition. However, the effect is stronger than in the OLS regressions, suggesting that α_A is, indeed, upward biased when using OLS. Using both instruments our estimation predicts that a Top100 runner who expects 10 percent more opposition is 5.1 percent less likely to choose a Major race. We also find that an increase in total prize money is associated with a positive and significant effect on the entry of Top100 runners. Using both instruments, our estimation predicts that an additional $100,000$$ in total prize money raises the likelihood that a Top100 runner participates in a Major race by 10.4 percent. Given that opposition refers to the proportion of East–African runners amongst the first twenty finishers in the previous year, a 10 percent increase in opposition is equivalent to the participation of two additional East–African runners. This implies that Top100 runners are willing to forgo potential prize winnings of $25,500$$ for each high ability opponent they are able to avoid. Finally, with regard to the effect of prize steepness on entry, we see that the coefficient on the interaction between steepness and expected opposition is negative, as predicted by the model. However, this result is not statistically significant.

When controlling for runners' characteristics, we find that runners who were more highly ranked in the previous year, are more likely to enter a Major race in the current year. In particular, within the Top100 runners, the highest ranked runner is 9 percent more likely to enter a Major race then the lowest ranked runner. Hence there exists evidence for a tendency of contestants to sort according to abilities. How this tendency is influenced by the overall distribution of abilities is the subject of our next estimation.

5.3 Distribution of talent

While Proposition 1 was concerned with the individual preferences of contestants, Proposition 2's focus is on the equilibrium distribution of contestants 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.

In equilibrium, each player chooses a best response to the contest choice of all other players by entering the contest that maximizes his expected payoff. In our data a runner's outside option, i.e., the prize he could have won in another race assuming identical performance, can be readily determined. A surprisingly high fraction of runners turns out to choose a best response. We find that 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 where their prize would have been higher. This suggests that in our framework, contestants choose contests carefully and in order to maximize their expected prize winnings. Contest choice is repeated over time and learning seems to have lead to the establishment of equilibrium behavior.

To test Proposition 2, 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:

$$S_t^M = \alpha_0 + \alpha_1 HA_t + \alpha_2 B_t^M + t + \varepsilon_t. \tag{5}$$

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 2, 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 2, sorting should be increasing in B_t^M. We control for time trends as well as whether the year was an Olympic year. Since marathons can be divided into spring–races 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 6 shows the estimates for equation (5). 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 10 races increases by 1 percent, the share of East–Africans who choose a Major marathon decreases by 1.03 percent. The effect is even stronger, 1.23 percent, when all 35 races are considered. These results constitute evidence for the decrease in sorting depicted in Figure 1. As expected, we also find evidence for a positive relation between sorting and prize budget differences. In particular, a 1 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 1.17 percent for the top 10 races and by 0.49 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 10 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, 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 a ranking of 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 runners who have (adjusted) finishing times within 1 percent of the fastest finishing time during the season.[20] We also look at those finishing within 5 percent and 10 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 the composition of the group of East–African runners is changing over time.

Table 7 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 the stronger the more able the runners under consideration. In particular, a 10 percent increase in the proportion of high ability runners, reduces sorting by 46, 7, or 3 percent when high ability refers to runners within 1, 5, or 10 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.

6 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

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

with high prizes and strong opposition over contests with low prizes and weak opposition? And how do these preferences depend on their abilities? In this paper we have provided both theoretical as well as empirical insight into these questions.

We have shown 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. Our analysis has allowed us to determine the "prize" that contestants are willing to pay for a decrease in opposition and that organizers have to award to guarantee their contest's attractiveness.

In future research we plan to further expand our understanding of contest selection. One issue of interest is the influence of peer effects on contest choice. These effects are common in models of school choice where students are assumed to care not only about their ranking within their class but also about the average quality of their peers (Damiano et al. (2010)). While competition is not explicitly modeled by the peer effect literature, in a contest setting, the relative value of rankings would be directly determined by the contest's prize structure.

Appendix 1 - Proofs

Proof of Proposition 1

It is immediate that E_H^j is increasing in w_j, b_j, and M_j, but decreasing in p_j. To prove the last claim of Proposition 1, increase the steepness of contest j's prize structure by letting $\tilde{M_j} < M_j$ and $\tilde{b_j} > b_j$ and consider

$$\begin{aligned}
\frac{E_H^j - \tilde{E_H^j}}{1 - \frac{c_H}{c_L}} &= b_j G(M_j, p_j) - \tilde{b_j} G(\tilde{M_j}, p_j) && (6)\\
&= b_j \sum_{m=0}^{M_j - 1} \binom{N-1}{m} p_j^m (1-p_j)^{N-1-m} - \tilde{b_j} \sum_{m=0}^{\tilde{M_j}-1} \binom{N-1}{m} p_j^m (1-p_j)^{N-1-m}\\
&= b_j \text{Prob}(\tilde{M_j} \leq H_j \leq M_j - 1) - (\tilde{b_j} - b_j) \text{Prob}(H_j \leq \tilde{M_j} - 1).
\end{aligned}$$

The first term represents the advantage of the flatter prize structure. When the number of opponents H_j turns out to be between $\tilde{M_j}$ and $M_j - 1$ then the flatter prize structure guarantees a positive payoff, b_j, 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 $\tilde{M_j} - 1$ then payoffs are positive for both prize structures but the steeper prize structure offers an extra payoff $\tilde{b_j} - b_j > 0$. Note that the likelihood ratio $\text{Prob}(H_j \leq \tilde{M_j} - 1)/\text{Prob}(\tilde{M_j} \leq H_j \leq M_j - 1)$ is strictly decreasing in p. It converges to 0 for $p_j \to 1$ and to ∞ for $p_j \to 0$. Hence there exists a $\bar{p}_j \in (0, 1)$ such that $E_H^j - \tilde{E_H^j} \geq 0$ if and only if $p_j > \bar{p}_j$. The steeper prize structure $(\tilde{M_j}, \tilde{b_j})$ guarantees a higher payoff if and only if the likelihood p_j with which opponents have high ability is smaller than $\bar{p}_j$. ■

Proof of Proposition 2

In a contest where an opponent has high ability with probability p, let

$$E_p[H|H \leq M-1] = \sum_{m=0}^{M-1} \binom{N-1}{m} p^m (1-p)^{N-1-m} m \tag{7}$$

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

$$E_p[H] = p(N-1) \tag{8}$$

denote the (unconditional) expected number of high ability opponents.

Consider first the case where $h < \frac{1}{2}$. In this case the number of high ability players falls short of the number of slots in each type of contest. Hence a strictly positive fraction of slots in each type of contest are filled with low ability contestants so that $p_S = 2hq \in (0,1)$ and $p_W = 2h(1-q) \in (0,1)$. The equilibrium is determined by

$$\Delta = b_S G(M_S, 2hq) - b_W G(M_W, 2h(1-q)). \tag{9}$$

We have

$$\begin{aligned}
\frac{dG(M,p)}{dp} &= \sum_{m=0}^{M-1} \tbinom{N-1}{m} [mp^{m-1}(1-p)^{N-1-m} - (N-1-m)p^m(1-p)^{N-2-m}] && (10)\\
&= \sum_{m=0}^{M-1} \tbinom{N-1}{m} p^{m-1}(1-p)^{N-2-m}[m-(N-1)p] && (11)\\
&= \frac{G(M,p)}{p(1-p)} \{E_p[H|H \leq M-1] - E_p[H]\} < 0. && (12)
\end{aligned}$$

It follows that

$$\frac{d\Delta}{dq} = 2h \left[b_S \frac{dG(M_S, p_S)}{dp} + b_W \frac{dG(M_W, p_W)}{dp} \right] < 0. \tag{13}$$

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

$$\Delta(q=0) = b_S - b_W G(M_W, 2h) > 0. \tag{14}$$

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

$$\Delta(q=1) = b_S G(M_S, 2h) - b_W. \tag{15}$$

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

$$\begin{aligned} h\frac{d\Delta}{dh} &= \left[b_S p_S \frac{dG(M_S, p_S)}{dp} - b_W p_W \frac{dG(M_W, p_W)}{dp}\right] && (16) \\ &= \frac{b_S G(M_S, p_S)}{1 - p_S}\{E_{p_S}[H|H \leq M_S - 1] - E_{p_S}[H]\} && (17) \\ &- \frac{b_W G(M_W, p_W)}{1 - p_W}\{E_{p_W}[H|H \leq M_W - 1] - E_{p_W}[H]\}. \end{aligned}$$

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

$$\begin{aligned} \frac{h}{b_W G(M_W, p_W)} \frac{d\Delta}{dh} &= \frac{1}{1 - p_S}\{E_{p_S}[H|H \leq M_S - 1] - E_{p_S}[H]\} && (18) \\ &- \frac{1}{1 - p_W}\{E_{p_W}[H|H \leq M_W - 1] - E_{p_W}[H]\}. \end{aligned}$$

It is one of the properties of the binomial distribution that the difference between the unconditional and the tail conditional mean 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 $q^* \geq \frac{1}{2}$ it holds that $p_S \geq p_W$ we can therefore find an upper bound for the first term by setting $p_S = p_W$ to get

$$\frac{h(1 - p_W)}{b_W G(M_W, p_W)} \frac{d\Delta}{dh} \leq E_{p_W}[H|H \leq M_S - 1] - E_{p_W}[H|H \leq M_W - 1] < 0. \quad (19)$$

The last inequality followed from $M_S < M_W$. Hence we have shown that at any equilibrium such that $q^* \geq \frac{1}{2}$ and hence $p_S^* \geq p_W^*$ it holds that $\frac{d\Delta}{dh}|_{q=q^*} < 0$. Together with $\frac{d\Delta}{dq} < 0$ this implies that q^* is strictly decreasing in $h \in (\bar{h}, \frac{1}{2})$ as long as $q^* \geq \frac{1}{2}$.

It remains to consider the case where $h \geq \frac{1}{2}$. For $q \leq 1 - \frac{1}{2h}$ we have $p_W = 1$ and $p_S \in (0, 1)$ so that $\Delta(q) = b_S G(M_S, p_S) > 0$. Hence in equilibrium it has to hold that $q^* > 1 - \frac{1}{2h}$. Similarly for $q \geq \frac{1}{2h}$ we find $p_S = 1$ and $p_W \in (0, 1)$ so that $\Delta(q) = -b_W G(M_W, p_W) < 0$. Hence in equilibrium it has to hold that $q^* < \frac{1}{2h}$. For $1 - \frac{1}{2h} < q < \frac{1}{2h}$, $\Delta(q)$ is given by (9), and the equilibrium q^* is the unique solution to $\Delta(q^*) = 0$ in $(1 - \frac{1}{2h}, \frac{1}{2h})$. Hence all the arguments used in the case where $h < \frac{1}{2}$ remain valid. In particular q^* is strictly decreasing in $h \in (\frac{1}{2}, 1)$ as long as $q^* \geq \frac{1}{2}$.

Hence we can conclude that there exists a $\bar{\bar{h}} \in (\bar{h}, 1]$ such that $q^*(h)$ is strictly decreasing in $(\bar{h}, \bar{\bar{h}})$ and $q^* \leq \frac{1}{2}$ for all $h > \bar{\bar{h}}$. ■

Appendix 2 - Tables and Figures

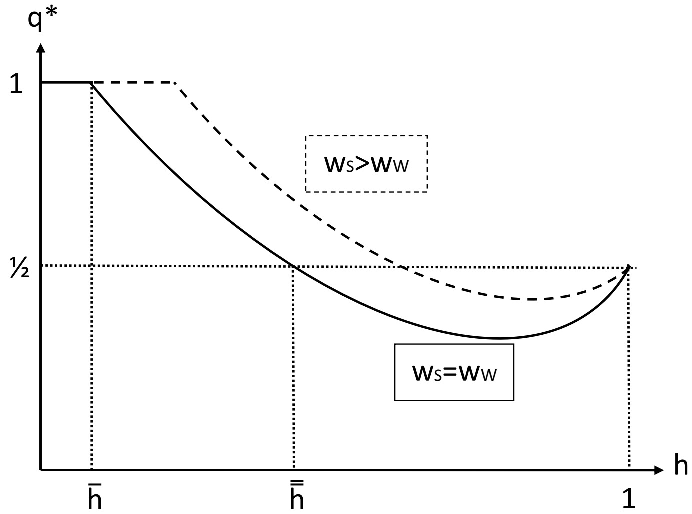

Figure 1: **Equilibrium.** The fraction q^* of high ability players who choose strong competition in dependence of the overall fraction h of high ability players in the population.

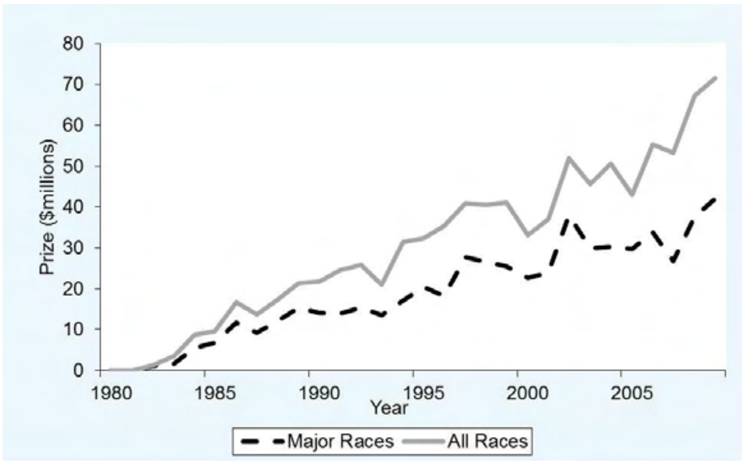

Figure 2: **Total Prize Money in Marathons.** Prize money is measured in US$(millions) and is for men's marathons only. It is aggregated over the 5 Major marathons and over all 35 marathons in the dataset, respectively.

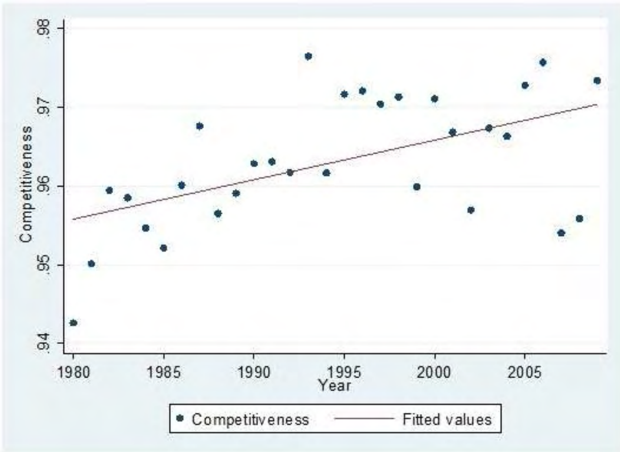

Figure 3: **Competitiveness of Marathon Running.** Competitiveness is defined as the ratio of the best (male) winning time of a year over the average finishing times of top 20 (male) finishers in all races. Finishing times are adjusted for race–course differences.

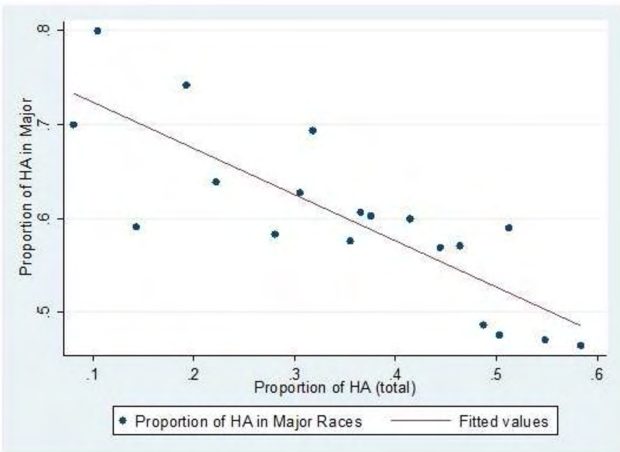

Figure 4: **Contest Choice.** "Proportion of HA (total)" is the proportion of high ability (East–African) runners in the 10 races under consideration. "Proportion of HA in Major" is the share of high ability runners who chose a Major race (Berlin, Boston, Chicago, London, New York) rather than a Minor race (Hamburg, Honolulu, Frankfurt, Paris, Rome). For men's marathons only.

	Major Races			Minor Races		
Variable	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
Total Prize ($)	238	221,689	126,466	1381	26,371	40,460
Steep Prize	238	0.57	0.5	1381	0.35	0.48
No. of Participants	236	22,332	10,143	859	6,838	6,462
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	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. The sample period is 1986 to 2009. "Total Prize" is the sum of prizes awarded in a race. "Steep Prize" takes value 1 if the Herfindahl-Hirschman 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. This data was collected separately from various sources, including ARRS, Wikipedia, 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 Runners			All other Runners		
Variable	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
Age	2892	28.78	4.54	7515	30.63	4.84
No. Races	2892	1.42	0.6	7515	1.27	0.53
Prize ($)	2892	8,307	19,698	7515	3,360	10,406
Finishing Time	2892	02:14	00:05	7515	02:17	00:05
Fraction HA (1%)	2892	0.02	0.15	7515	0.01	0.12
Fraction HA (5%)	2892	0.28	0.45	7515	0.21	0.41
Fraction HA (10%)	2892	0.73	0.45	7515	0.67	0.47
	Female Runners					
	East African Runners			All other Runners		
Variable	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
Age	646	27.69	4.44	7729	31.76	6.03
No. Races	646	1.45	0.59	7729	1.31	0.59
Prize ($)	646	13,539	27,229	7729	4,031	12'175
Finishing Time	646	02:33	00:08	7729	02:41	00:09
Fraction HA (1%)	646	0.01	0.1	7729	0.01	0.07
Fraction HA (5%)	646	0.17	0.37	7729	0.06	0.25
Fraction HA (10%)	646	0.5	0.5	7729	0.26	0.44

Table 2: **Descriptive Statistics (Runners)**. Means and standard deviations (by gender category) for East–African and Non–East–African 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 a runner obtains (on average) per race. “Finishing Times” have been adjusted using ARRS conversion factors to ensure that race courses are comparable. “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.

VARIABLES	OLS Enter Major	OLS Enter Major	OLS Enter Major	OLS Enter Major
Opposition_{t-1}	-0.5957***	-0.5743***	-0.1137***	-0.1172***
	[0.030]	[0.034]	[0.033]	[0.040]
Total Prize ('00000$)	0.2432***	0.2313***	0.0140**	0.0139**
	[0.005]	[0.005]	[0.006]	[0.006]
Steep Prize	-0.003	-0.0017	0.0144	0.0132
	[0.012]	[0.012]	[0.009]	[0.012]
Opposition_{t-1}·Steep Prize				0.0067
				[0.042]
Female		0.0041	0.0172	0.0172
		[0.012]	[0.019]	[0.019]
Age		-0.0003***	-0.0001**	-0.0001**
		[0.000]	[0.000]	[0.000]
Nationality		-0.0005*	-0.0006***	-0.0006***
		[0.000]	[0.000]	[0.000]
At Home		0.1041	0.0087	0.0088
		[0.077]	[0.053]	[0.053]
Rank_{t-1}		-0.0020***	-0.0010***	-0.0010***
		[0.000]	[0.000]	[0.000]
Trend			-0.0027**	-0.0027**
			[0.001]	[0.001]
Trend*Female			-0.0009	-0.0009
			[0.001]	[0.001]
Olympic Year			0.0032	0.0032
			[0.009]	[0.009]
Constant	0.3122***	0.4497***	0.1637	0.1633
	[0.009]	[0.018]	[0.135]	[0.135]
Race Fixed Effects	No	No	Yes	Yes
Observations	5492	5469	5469	5469
R-Squared	0.349	0.363	0.717	0.717

Table 3: **Probability to Enter Major Race (OLS)**. *,**,*** denotes significance at the 10%, 5%, 1% level, respectively. Estimations use linear probability model. The sample is restricted to the runners who were amongst the top 100 Non–East–African runners in the previous year. "Opposition_{t-1}", is the fraction of East-African runners amongst the top20 finishers of the race in the previous year. "Rank_{t-1}" is the ranking of the runner in the previous year (between 1 and 100). "Olympic Year" takes value 1 in years 1988, 1992, 1996, 2000, 2004, and 2008 and 0 in all other years.

VARIABLES	$Opposition_{t-1}$	Total Prize ('00000$)
Commodity Price Index $Kenya_{t-1}$	0.0018***	
	[0.000]	
Log Rainfall $Kenya_{t-1}$	0.0917*	
	[0.048]	
Commodity Price Index $Ethiopia_{t-1}$	-0.0005***	
	[0.000]	
Log Rainfall $Ethiopia_{t-1}$	0.1022***	
	[0.034]	
Exchange Rate		0.2865***
		[0.022]
Constant	-0.7351***	-1.0323***
	[0.165]	[0.261]
Controls	Yes	Yes
Trend	Yes	Yes
Trend·Female	Yes	Yes
Race Fixed Effects	Yes	Yes
Observations	5469	5369
R-Squared	0.649	0.743
F-Test of Excl. Instr.	26.25	88.14

Table 4: **First Stage IV**. *,**,*** denotes significance at the 10%, 5%, 1% level, respectively. "Commodity Price Index Kenya (Ethiopia)" uses international commodity price data from IMF. All variables indexed by $t-1$ relate to the previous year. "Exchange Rate" is the exchange rate of the country of the race relative to the Special Drawing Rights currency basket provided by the IMF.

VARIABLES	IV (Opposition$_{t-1}$) Enter Top 5	IV (Total Prize) Enter Major	IV (Both) Enter Major	IV (Both) Enter Major
Opposition$_{t-1}$	-0.5894**	-0.1200***	-0.5145**	-0.4922**
	[0.244]	[0.035]	[0.240]	[0.249]
Total Prize ('00000$)	0.0157***	0.1309***	0.1042**	0.1049**
	[0.006]	[0.047]	[0.045]	[0.045]
Steep Prize	0.0149	-0.0207	-0.0119	-0.0089
	[0.010]	[0.017]	[0.016]	[0.019]
Opposition$_{t-1}$·Steep Prize				-0.019
				[0.056]
Female	0.0343	0.0292	0.0416**	0.0412*
	[0.021]	[0.020]	[0.021]	[0.021]
Age	-0.0001**	-0.0001**	-0.0001**	-0.0001**
	[0.000]	[0.000]	[0.000]	[0.000]
Nationality	-0.0005***	-0.0007***	-0.0007***	-0.0007***
	[0.000]	[0.000]	[0.000]	[0.000]
At Home	0.0141	0.0067	0.0116	0.0113
	[0.054]	[0.055]	[0.053]	[0.053]
Rank$_{t-1}$	-0.0011***	-0.0008***	-0.0009***	-0.0009***
	[0.000]	[0.000]	[0.000]	[0.000]
Trend	0.007	-0.0096***	0.0003	-0.0001
	[0.005]	[0.003]	[0.006]	[0.006]
Trend·Female	-0.0064**	0.0002	-0.0047	-0.0045
	[0.003]	[0.001]	[0.003]	[0.003]
Olympic Year	0.0033	0.015	0.0121	0.0121
	[0.009]	[0.010]	[0.010]	[0.010]
Constant	0.1071	0.9027***	0.8978***	0.8975***
	[0.141]	[0.079]	[0.080]	[0.080]
Race Fixed Effects	Yes	Yes	Yes	Yes
Observations	5469	5368	5368	5368
R-Squared	0.706	0.695	0.717	0.717

Table 5: **Probability to Enter Major Race (IV)**. *,**,*** denotes significance at the 10%, 5%, 1% level, respectively. For definition of variables see Table 3.

	Top 10 Races				All 35 Races			
VARIABLES	Sorting	Sorting	Sorting	Sorting	Sorting	Sorting	Sorting	Sorting
Fraction 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]
Prize Major		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*	-0.009	-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 6: **Sorting of High Ability Runners (Origin)**. *,**,*** denotes significance at the 10%, 5%, 1% level, respectively. Years from 1986 until 2009. The dependent variable, "Sorting", is the proportion of East African runners who entered a Major rather than a Minor race. "Fraction HA (Origin)", is the overall fraction of East African runners in the races under consideration. Both variables are calculated separately for each race season (spring, autumn) and gender category. "Prize Major" is the proportion of the overall prize money awarded in the Major races. For definition of variables see Table 3.

	Top 10 Races			All 35 Races		
VARIABLES	Sorting	Sorting	Sorting	Sorting	Sorting	Sorting
Fraction HA (1%)	-1.9589***			-4.6357**		
	[0.707]			[2.148]		
Fraction HA (5%)		-0.2751*			-0.7163***	
		[0.159]			[0.214]	
Fraction HA (10%)			-0.1194			-0.3075***
			[0.146]			[0.110]
Prize Major	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	0.006	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***	0.1	-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 7: **Sorting of High Ability Runners (Finishing Times)**. *,**,*** denotes significance at the 10%, 5%, 1% level, respectively. Years from 1986 until 2009. The dependent variable, "Sorting", is the proportion of high ability runners who entered a Major rather than a Minor race. High ability runners have (adjusted) finishing times within 1%, 5%, or 10% of the year's fastest time in their gender category. "Fraction HA (1%, 5%, 10%)", is the overall fraction of high ability runners in the races under consideration. Both variables are calculated separately for each race season (spring, autumn) and gender category. "Prize Major" is the proportion of the overall prize money awarded in the Major races. For definition of variables see Table 3.

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. (1993) Rigging the Lobbying Process: An Application of the All–Pay Auction. *American Economic Review* **83**(1), 289-294.

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

[4] Burguet, R., Sákovics, J. (1999) Imperfect Competition in Auction Designs. *International Economic Review* **40**(1), 231-247.

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

[9] DiMasi, J., Hansen, R., Grabowski, H. (2003) The Prize of Innovation: New Estimates of Drug Development Costs. *Journal of Health Economics* **22**, 151-185.

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

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

[14] Fullerton, R., L., McAfee, R., P. (1999) Auctioning Entry into Tournaments. *Journal of Political Economy* **107**(3), 573-605.

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

[16] Lazear, E., P. (1986) Salaries and Piece Rates. *Journal of Business* **59**(3), 405-431.

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

[18] Michaels, E., Handfield–Jones, H., Axelrod, B. (2001) *The War for Talent*, Harvard Business School Press, Boston.

[19] Miguel, E., S. Satyanath, and E. Sergenti 2004: Economic Shocks and Civil Conflict: An Instrumental Variables Approach, *Journal of Political Economy*, **112**, 725–753.

[20] Mirrlees, J., A. (1971) An Exploration in the Theory of Optimum Income Taxation. *Review of Economic Studies* **38**, 175-208.

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

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

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

[24] Rajan, R., Wulf, J. (2006) The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies. *Review of Economics and Statistics* **88**(4), 759-773.

[25] Spence, A., M. (1973) Job Market Signaling. *Quarterly Journal of Economics* **87**, 355-374.

[26] Stiglitz, J., E. (1975) Incentives, Risk, and Information: Notes Towards a Theory of Hierarchy. *Bell Journal of Economics* **6**(2), 552-579.

[27] 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.