The Distribution of Talent across Contests

Ghazala Azmat *
and
Marc Möller †

Abstract

This paper considers the choice between two modes of competition; strong competition for large but few prizes and weak competition for small but many prizes. Intuition suggests that contestants sort according to abilities with strong competition attracting the most able contestants. We show, both theoretically and empirically, that this intuition fails to hold in general. The allocation of talent across contests depends on the overall distribution of abilities. Sorting indeed exists when high abilities are rare. However, as high abilities become more frequent, they tend to shy away from competition. When high abilities are sufficiently frequent, weak competition attracts the stronger field. We provide evidence for our results by using a large panel data set about the race choice of professional marathon runners. Our results have implications for the allocation of firms across markets, workers across firms, students across schools, and scientists across R&D contests.

JEL *classification*: D82, M52, J33
Keywords: Competition, contests, 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. Agents of differing ability compete for valuable but limited resources by exerting effort. In many cases, agents not only decide on how much effort to exert but also on where to compete. For example, college graduates select amongst the promotion tournaments offered by alternative employers. Architects choose design contests, parents select schools for their children, and marathon runners pick races. Given that competition is one of the core topics of economics, it is surprising that the issue of contest choice has attracted relatively few attention.

Contests typically differ in the strength of competition they induce. Some contests offer large but few prizes inducing their participants to exert high levels of effort. Other contests offer small but many prizes and are characterized by lower effort levels. For example, in a labor market contest, some employers offer steep hierarchies where few employees are promoted but wage increases are large. Other employers offer flat hierarchies where the benefits of becoming promoted are smaller but incurred by a larger group of workers. In an educational setting, some schools' class grades may be more inflated than others, making a top grade less valuable but easier to achieve. Finally, in design contests and marathons the strength of competition is determined by the allocation of the monetary prize budget.

Do the most competitive contests attract the most able contestants? In this paper, we show both theoretically and empirically, that this is not always the case. We demonstrate that the allocation of abilities across contests depends in a systematic way on the overall distribution of ability within the population of contestants. Our results offer new insight into the distribution of talent across contests and have important implications for contest regulators, organizers, and participants. They imply, for example, that the level of competitive balance achieved by a school regulator; the quality of workers attracted by a firm's promotion policy; and the opposition expected by a marathon runner, all depend, in a systematic way, on the overall distribution of abilities.

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 greater efforts, i.e. stronger competition. Modeling competition as an all–pay auction, we first show that a contestant's payoff is increasing in the contest's overall prize budget but decreasing in the likelihood of meeting high ability opponents. Making the contest more competitive by increasing the size but decreasing the number of prizes leads to higher payoffs when the probability to meet high ability opponents is small but lower payoffs when this probability is large.

Our main theoretical result shows that the fraction of high ability contestants who choose strong competition is decreasing in the total number of high ability contestants. When this number becomes sufficiently large, weak competition attracts an even larger fraction of high ability contestants than strong competition. The intuition for this result is as follows. Strong contests offer greater prizes than weak contests but induce higher costs of effort. The difference in effort costs is increasing in the number of (potential) high ability rivals. Hence, as high abilities become more frequent, strong contests become less attractive. When high abilities become sufficiently frequent, most of the potential prize gains are spent in form of effort cost. The fact that weak contests mitigate competition then becomes so important that high abilities prefer weak contests, even though in weak contests their rivals are expected to be more able.

We provide empirical support for our results by investigating the race choice of professional marathon runners. With more than 20,000 observations we examine contest choice over three decades. There are three main features that make marathons the ideal setting to test our theory. First, marathon running is strongly dominated by highly talented African runners. This enables us to use a runner's origin as a proxy for ability which, unlike finishing times, is independent of effort and prize considerations. Second, due to the abolishment of the amateur rule by the International Olympic Committee, marathons started to offer prize money in the early 1980s. This has lead to a dramatic increase in the participation of African runners and hence, has altered the overall distribution of abilities. Finally, for historic reasons, a large fraction of the prize money is offered by a small number of Top races (Berlin, Boston, Chicago, London, and New York).

In accordance with our theory, we find that the likelihood of a high ability runner to participate in a Top race is increasing in the race's prize budget but decreasing in the expected opposition, i.e. the number of Africans who participated in last year's edition. To increase the probability of participation by 10%, a contest would need to increase its total prize budget by $6,667 or reduce the proportion of African participants by 13%. These results allow us to determine the "price" that contestants assign to opposition. Our calculations show that Top runners are willing to forgo potential prize winnings of $2,567\$$ for each African opponent they are able to avoid.

In line with our main theoretical result, we find that, as the overall fraction of Africans in the population of runners increases, more Africans shy away from competition. In particular, a 1% increase in the overall fraction of Africans, leads to a 1.03% reduction in the fraction of Africans entering a Top race. According to our estimates, Top races would have to increase their prize budgets by 11.4% in order to maintain their attractiveness to Africans after a 10% increase in the overall fraction of Africans.

The paper is organized as follows. In the remainder of this section we review the related literature. Section 2 introduces the theoretical model and describes how competition depends on a contest's prizes and the abilities of its participants. Section 3 contains our theoretical results about the contestants' choice of contest. Our empirical strategy and results are the subject of Section 4. Section 5 concludes.

Related Literature

Since competition is ubiquitous, the issue of sorting across contests considered in this paper is fairly general and relates to a variety of literatures. In the following we discuss links with research on labor tournaments, contest design, auction theory, and school choice.

The idea that firms may use their compensation schemes to screen between workers of differing ability can be traced back to Stiglitz (1975) and Lazear (1986). Empirical studies have shown that these selection effects can be equally important as incentive effects. For example, Lazear (2000) shows that about half of the increase in output due to a company's change in compensation scheme can be accounted to the attraction of more able workers. In contrast to our setup, these papers focus on the choice between fixed salaries and piece rates so that relative comparisons are absent and a worker's firm choice is independent of other workers' choices.

Lazear and Rosen (1981) were the first to propose relative comparisons in form of pairwise labor tournaments as a way to provide workers with optimal incentives to exert effort. In order to allow for contests with a differing number of prizes, Yun (1997) has extended their setup by considering tournaments with a general number of workers. Yun's interest is in the characterization of the contracts that induce efficient self–selection *and* first best levels of effort. Yun shows that the optimal labor tournaments offer many large prizes to high ability workers and few small prizes to low ability workers. In contrast, we are interested in the extent to which workers self–select in dependence of the distribution of abilities, and the trade–off is between tournaments with large but few and tournaments with small but many prizes.

Our work is also related to the recent literature on contest design.[1] This literature typically considers a single contest and takes the set of participants as exogenously given. Only a few papers endogenize the set of participants by studying the influence of entry fees, but they maintain the focus on a single contest (Taylor (1995), Fullerton and McAfee (1999)). The observation that contest design not only influences the players effort choices but also their choice of contest has been the starting point of our own recent work.

In Azmat and Möller (2009) we have considered contests which compete for the

[1]The design issues considered by this literature include the allocation of prizes (Barut and Kovenock (1998), Krishna and Morgan (1998), Moldovanu and Sela (2001), Rosen (2001), Szymanski and Valletti (2005)), the division of contests into sub–contests (Moldovanu and Sela (2007)), simultaneous versus sequential designs (Clark and Riis (1998)), and optimal seeding in elimination tournaments (Groh et al. (2008)).

participation of a group of homogenous contestants via the selection of prize structures. Our main result has shown that equilibrium prize structures become steeper as the noisiness of contest outcomes increases. We have provided empirical evidence for this result by showing that longer and hence less predictable road running contests offer steeper prize structures. However, since contestants were assumed to be homogenous the distribution of talent across contests was beyond the scope of our theoretical analysis. Moreover, since our dataset was restricted to a single year it was not possible to consider how the runners' race choice responds to changes in the overall distribution of abilities.

Selection effects have also attracted attention in the literature on competing auctions (McAfee (1993), Peters and Severinov (1997), Burguet and Sakovics (1999)). In contrast to our model, auctions are assumed to offers a single and identical prize (one unit of a homogenous good) and bidders self–select in dependence of the auctions' reservation prices. An exception in this respect is the recent paper by Moldovanu et al. (2008) which allows auctions to vary in their supply. However, while our results are driven by the dependence of a contestant's effort on his set of opponents, in their setup bidders simply bid their valuation independently of their choice of auction.

Finally, our interest in the allocation of talent is shared by the literature on educational policies and school choice (see for example Arnott and Rowse (1987)). The distinguishing feature of this line of research is the assumption that students not only value their individual ranking within a school but also benefit from high quality peers. For example, Damiano, Hao, and Suen (2010) derive the equilibrium distribution of talent in dependence of the size of this peer effect. They show that the segregation of talent is increasing in the peer effect. Our analysis demonstrates that segregation is an issue even in the absence of peer effects, and that segregation may go either way.

On the empirical side, a few recent papers have reported evidence for selection effects in tournaments. Leuven et al. (2008) conduct a field experiment where students self–select into tournaments that award a single first prize of differing size to the student with the highest exam score. They find that more able students enter contests with higher prizes. Harbring and Irlenbusch (2003) consider an experiment in which players choose efforts in an all–pay auction. Interpreting those players who chose to provide zero effort as non–participants, they find that the number of participants is increasing in the number of prizes. Less related are two recent studies which consider self–selection between a tournament and alternative compensation schemes. Eriksson et al. (2009) design a laboratory experiment where subjects choose between a pairwise tournament and a piece–rate payment. They find that more risk averse subjects are less likely to choose tournaments. Dohmen and Falk (2011) consider the choice between a pairwise tournament and a fixed payment in a real–effort experiment. They find that those subject who choose a tournament have higher ability, lower degrees of risk aversion, and a more optimistic relative self–assessment. Neither of these studies tests for selection effects across tournaments with a differing number of prizes nor determines the allocation of ability in dependence of its overall distribution.

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 $i \in \{S, W\}$. A contest of type i offers

$M_i \in \{1, 2, \ldots, N\}$ identical prizes of size $b_i > 0$ and a performance–independent payment $w_i \geq 0$ to each of its participants.[2] In a labor tournament setting, w_i could represent the workers base wage while b_i measures the wage increase or bonus for those who become promoted. In the example of school choice w_i could be seen as the benefit of graduating from school i while b_i represents the value of obtaining a top–grade. Some sports contests not only announce prizes but pay athletes for their attendance.

Contests of type S award higher ($b_S > b_W$) but fewer ($M_S < M_W$) prizes than contests of type W. As will become clear below, contests of type S induce stronger competition, i.e. higher levels of effort, than contests of type W. We therefore denominate contests of type S as *strong contests* and contests of type W as *weak contests.* 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. Players differ with respect to their constant marginal cost of effort c. There are two types of players, $j \in \{L, H\}$. 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 known to all players.

In each contest, players compete for prizes by exerting effort. Here we follow an extensive literature on contest design (see for example Clark and Riis (1998) or Moldovanu and Sela (2001)) by assuming that this competition takes the form of an all–pay auc-

[2]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–idential prizes is still missing. A first step into this direction has been made by Cohen and Sela (2008).

tion. In particular, we suppose that in each contest, prizes are awarded to the players with the highest levels of effort while ties are broken randomly. A player of type $j \in \{L, H\}$ who exerts effort $e_j \geq 0$ in a contest of type $i \in \{S, W\}$ receives the payoff $U_j^i = w_i + b_i - c_j e_j$ if he wins one of the M_i prizes and $U_j^i = w_i - c_j e_j$ otherwise.

Following Clark and Riis (1998), we assume that players are able to observe their opponents' abilities at the time of their effort choice. This assumption holds for example in our empirical setting since marathon runners are informed about the identity and past performance of their opponents. It also holds in the labor tournament or educational setting where workers/students learn the abilities of their peers over the course of the contest. Without this assumption efforts would still be higher in strong contests than in weak contests and a player's payoff would still be decreasing in the (expected) abilities of his opponents (see Moldovanu and Sela (2001)). We therefore contemplate that our results would remain valid when players face uncertainty about the ablities of their opponents not only at the time of contest choice but also at the time of competition.

The timing of events is as follows: (1) Players simultaneously choose which type of contest to enter. (2) Each contest accepts as many high ability players as possible and fills any remaining slots with low ability players.[3] Players are assumed to have zero outside options which implies that those players who were turned down by the contest of their choice enter a contest of the other type. (3) Players observe their opponents'

[3] This assumption requires organizers to observe the players' abilities whenever the number of applicants exceeds the number of slots. While in some of our examples the contestants' abilities are observable fairly well (e.g. architects, athletes), in others entry tests are often employed to select the most able amongst the applicants (e.g. workers, students). While our assumption simplifies the algebra, our results are robust to the introduction of random rationing.

abilities and compete for prizes by choosing their effort levels simultaneously.

In the remainder of this section we analyze the players' competition in dependence of a contest's prize structure. We employ the results of Clark and Riis (1998) which will be reviewed below. 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 i, N players will compete for M_i identical prizes. Players value a prize identically at b_i but differ in their marginal costs of effort c_j. A players's payoff from obtaining a prize net of his effort costs is $b_i - c_j e$. 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 they attach to obtaining a unit $v_j = b_i/c_j$.

Clark and Riis (1998) show that the equilibrium of an all–pay auction with heterogenous 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 $n = 1, 2, \ldots, M$ with the highest valuations and zero for all other players.

This result has the following implications for our model. In a contest of type i, players may attach two different values to obtaining one of the M_i prizes. A low ability player assigns the value $v_L = b_i/c_L$ whereas a high ability player has valuation $v_H = b_i/c_H > v_L$. The equilibrium can therefore be described in terms of the number of high ability players H_i in a contest of type i.

For example, if $H_i \geq M_i + 1$ then the result of Clark and Riis (1998) implies that expected payoffs are w_i for all players. In contrast, if $H_i \leq M_i$ then expected payoffs are w_i for all low ability players and $w_i + b_i(1 - \frac{c_H}{c_L})$ for all high ability players. We summarize these findings in the following:

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

3 Distribution of talent

In this section we determine the equilibrium allocation of abilities across the two modes of competition. In general, players make their contest choice contingent on the contests' characteristics and the expected abilities of their opponents. As a consequence, 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.

Consider first the contest choice of low ability players. In our model competition takes the extreme form of an all pay–auction. Lemma 1 has shown that as a consequence a low ability player's expected payoff in a contest of type i is

$$E[U_L^i] = w_i. \tag{1}$$

Independently of the contest's prize structure and the opponents' ability, a low ability player's potential benefit from winning a prize is exactly compensated by his effort cost and his expected payoff is therefore given by the contest's performance independent payment. Hence low ability players simply prefer the contest with the highest w_i and are indifferent when $w_S = w_W$.

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

$$G(M_i, p_i) = \sum_{m=0}^{M_i-1} \binom{N-1}{m} (p_i)^m (1-p_i)^{N-1-m}. \tag{2}$$

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

$$E[U_H^i] = w_i + b_i(1 - \frac{c_H}{c_L})G(M_i, p_i). \tag{3}$$

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

Proposition 1 *A high ability player's expected payoff from entering a contest of type i, is increasing in the performance independent payment w_i, and the number M_i and size b_i of prizes, but decreasing in the probability p_i with which opponents have high ability. An increase in the steepness of contest i's prize structure leads to higher payoffs when $p_i < \bar{p}_i$ and to lower payoffs when $p_i > \bar{p}_i$. A low ability player's expected payoff is increasing in w_i but independent of the remaining variables.*

While the first claim of Proposition 1 is straight forward, the second claim requires some explanation. Suppose that we increase the steepness of contest i's prize structure by raising b_i and lowering M_i. Contest i 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.

What do these preferences imply for the equilibrium allocation of talent? Since contests have a priority for high ability players, in equilibrium low ability players simply fill the slots that have remained unfilled by high ability players. 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 the number of high ability players, hN, is smaller than the number of slots available in each type of contest, $\frac{N}{2}$, then the probability with which a slot of type S is filled with a high ability player is given by $p_S = 2hq$ while a slot of type W is filled with probability $p_W = 2h(1-q)$. If instead $hN \geq \frac{N}{2}$ then for sufficiently small or sufficiently large values of q high ability players exhibit excess demand for one type of contest. For example when $q \geq \frac{1}{2h}$ then all slots in contests of type S become filled with high ability players and all remaining high ability players have to enter contests of type W. In dependence of the values of h and q, the probabilities p_S and p_W with which a slot of type S or W is filled with a high ability player can be determined as

$$p_S = \begin{cases} 2h-1 & \text{if} \quad h \geq \frac{1}{2} \text{ and } q \leq 1-\frac{1}{2h} \\ 1 & \text{if} \quad h \geq \frac{1}{2} \text{ and } q \geq \frac{1}{2h} \\ 2hq & \text{otherwise} \end{cases} \tag{4}$$

$$p_W = \begin{cases} 1 & \text{if} \quad h \geq \frac{1}{2} \text{ and } q \leq 1-\frac{1}{2h} \\ 2h-1 & \text{if} \quad h \geq \frac{1}{2} \text{ and } q \geq \frac{1}{2h} \\ 2h(1-q) & \text{otherwise.} \end{cases} \tag{5}$$

The equilibrium allocation of abilities across contests is determined by the contest choice of high ability players which depends on

$$\Delta \equiv b_S G(M_S, p_S) - b_W G(M_W, p_W) + \frac{w_S - w_W}{1 - \frac{c_H}{c_L}}. \tag{6}$$

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^*.*

Intuition suggests 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 $w_S = w_W$ and 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 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

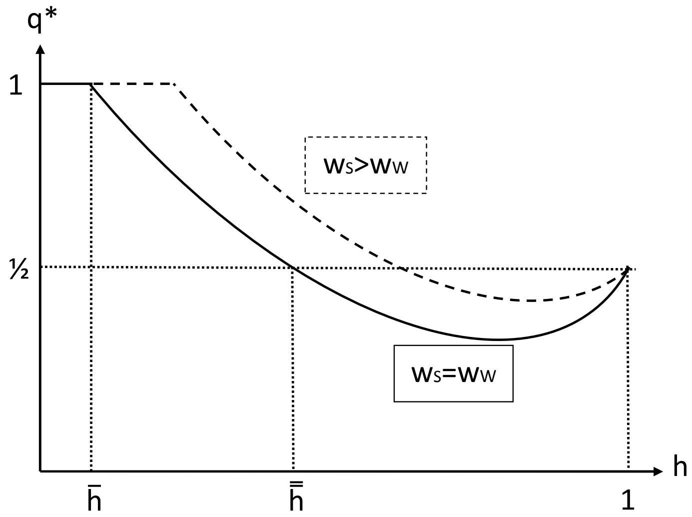

Figure 1: The fraction q^* of high ability players who choose strong competition in dependence of the overall fraction h of high ability players in the population.

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. Note that $\bar{h} < \frac{1}{2}$ implies that complete sorting can never occurr when high ability players are equally frequent as low ability players.

The intuition for this result is as follows. Strong competition offers high potential prizes while weak competition is 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, effort considerations become so important that high ability players prefer weak competition over strong competition 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 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 whereever $q^* < 1$.

In the following section we will test the results contained in Propositions 1 and 2 by considering the race choice of professional marathon runners. Our empirical analysis provides strong support for our theory, especially for the negative dependence of sorting on the overall distribution of abilities presented in Figure 1.

4 Empirical Section

In this section we will test the predictions of the model using a large panel dataset of international 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 budgets and prize structures should be observable and differ across contests. There are three features that make professional marathon running the ideal setting to test our model.

First, in 1986, the International Olympic Committee abolished its amateur rule. According to this rule, only amateurs were allowed to participate in the Olympic games. Amateur runners, by construction, were forbidden to accept prize money. As a consequence, marathons were reluctant to award prize money. Figure 2 shows that since the lifting of the ban, prize money has been increasing steadily.

Second, in part as a result of the lifting of the ban and the emergence of sizable prize budgets, the participation of African runners has been increasing over time.[4] This

[4] Running means big money for Africans. The average income of a Kenyan is approximately \$100 a month. The average prize money for winning a major marathon is \$100,000.

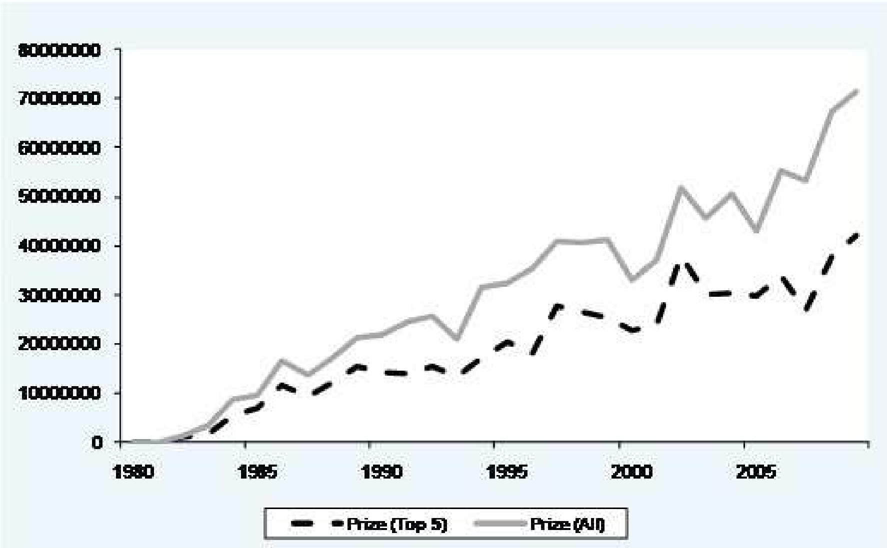

Figure 2: Prize money aggregated over Top5 races and all 34 races, respectively. Prize money is measured in US$ and is for men's marathons only.

is of particular relevance for our analysis since Africans are highly able runners. For example, in 2009, 93 out of the 100 fastest male marathon runners were African, of which 62 were Kenyan and 26 were Ethiopians.[5] In the same year more than 70% of total prize money was awarded to Kenyan or Ethiopian runners. While nowadays long distance running is strongly dominated by African runners, this is a recent phenomenon. In 1986, less than 10% of races were won by African runners, compared with over 80% in 2009. As can be seen in Figure 3, the entry of African runners has lead to a change in the distribution of abilities. Taking the ratio of the race time of the average runner finishing in the Top 20 over the average winning time as a measure of competitiveness, Figure 3 shows that marathon running has become more competitive over the years.

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

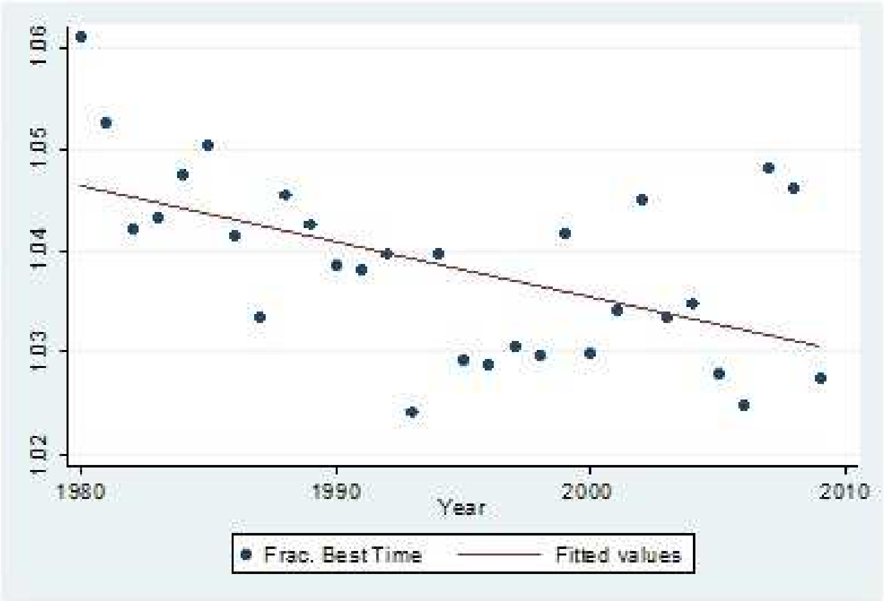

Figure 3: The competitiveness of (male) marathon running. Competitiveness is defined as the ratio of the average finishing time of a Top20 runner over the average winning time of the year.

While in the early 1980s, winners had a comparative advantage of around 6%, this advantage has decreased to only 2% in the late 2000s. In the light of our theoretical model this change can be interpreted as an increase in the fraction of high ability contestants.

Finally, the third feature that makes marathon running the ideal setting to test our theory is the fact that a small number of races award a large share of the overall prize money. For historical reasons, the (Top5) marathons in Berlin, Boston, Chicago, London, and New York are the most prestigious races and have gained the highest international recognition. As can be seen from Figure 2 the Top5 marathons award more than 50% of the total prize money offered by the 34 races contained in our dataset.

A marathon runner therefore faces the tradeoff that is at the heart of our theoretical model: Participate in a Top5 race where prizes are high but competition is strong or choose another race where prizes are low but competition is weak.

There are further reasons to use marathon data rather than for example data on labor tournaments. While in marathons prizes and performance are easily observed, a firm's pay structure and a worker's individual perfomance are hardly available. Moreover, professional runners run two marathons per year. This makes contest choice sufficiently infrequent to have an important effect on expected earnings and sufficiently frequent to guarantee learning and the establishment of equilibrium behaviour. A runner's outside option, i.e. the prize he could have won in another race assuming identical performance, can be readily determined. In our dataset $X\%$ of the prize winners could not have earned a higher prize in any other marathon. In comparison, a worker's outside option is difficult if not impossible to determine. A final advantage of marathons is that, unlike firms, they are fairly homogenous in their setup.

As a brief preview of our results consider a runner's choice between participating in a Top5 race and entering one of the next five most important races (Hamburg, Honolulu, Frankfurt, Paris, and Rome). Figure 4 shows the proportion of Africans that choose to participate in a Top5 race as a function of the overall fraction of Africans in the Top10 races in a given year. If we identify Africans as high ability runners then Figure 4 is in close accordance with the predictions of the theoretical model depicted in Figure 1. As the overall number of African runners increases, Africans shy away from competition by entering less competitive marathons. In the preceeding sections we will provide more rigorous evidence for this relationship.

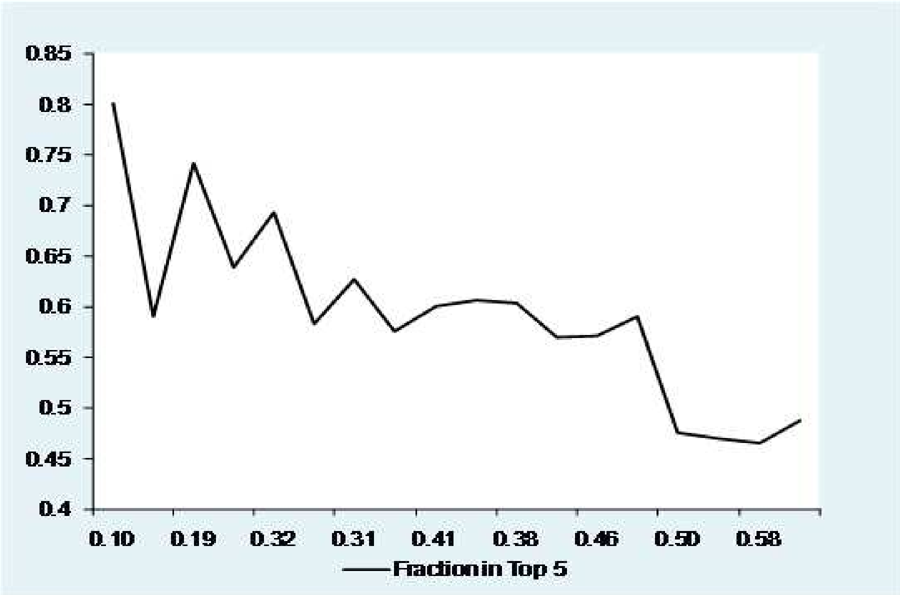

Figure 4: The proportion of African runners entering a Top5 race as a function of the overall fraction of African runners in the Top10 races. For men's marathons only.

4.1 Data Description

We use data from the Association of Road Running Statisticians containing detailed race and runner information for the largest international marathons from 1980 to 2009. We restrict attention to the 34 marathons for which we have complete data. With the exception of a few races these marathons contain the ones offering the highest prize budgets.[6] We have information for both, male and female races. For each race, we observe the date, location, total prize money, as well as the prize distribution. At the runner level, we can identify the top twenty finishers for each race and in many cases, we can identify the top forty or fifty finishers. Overall, we have a sample of over 20,000

[6]Some well paying races are relatively new (Dubai) while for others information about prizes is not available (Rotterdam).

runners over the 30 years. 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 the Top5 races and for all other races. From this table, we can see that there are stark differences between these race categories. In particular, Top5 races award around five times as much prize money as the other races ($228,206 compared with $43,947). Different from our theoretical model, the prizes awarded by a marathon are not identical but decreasing in rank. However, as in the model, a marathon's prize budget may be more or less concentrated. Similar to the strong contests of our model, the prize structure of a Top5 race is more concentrated towards the first prize. In particular, in the Top5 races first prizes are on average 93% larger than second prizes compared to 88% in all other races. Since in Top5 races prizes are higher and more concentrated we will identify Top5 races as strong contests and all other races as weak contests. Further motivation for this identification can be derived from the fact that in Top5 races winning times are on average X minutes faster than in all other races. Part of this difference can be explained by the fact that, in accordance with the model, the prizes offered by a Top5 race induce higher effort levels. The remaining part is due to selection effects. From Table 1 we can see that the fraction of high ability runners is considerably larger in Top5 races. This holds no matter whether we identify high abilitiy runners by African origin or as those runners whose finishing time is within 1%, 5%, or 10%, respectively, of the fastest finishing time of the year.[7] For example, 19% of race participants in the

[7]Since some race courses are faster than others we adjust all finishing times using a conversion factor constructed by the Association of Road Running Statisticians. This is done throughout the

Top5 races are African compared to 15% in the other races.[8] Similarly, 30% of runners in the Top5 races have a finishing time within 5% of the best finishing time, compared with only 7% in all other races.

Table 2 compares the descriptive statistics for the runners who are amongst the fastest 100 finishers for a given year in their gender category, with all other runners. As one would expect, there are large differences in the prize money and the finishing times. The average finishing time of a Top100 runner is 2 hours and 20 minutes (includes male and female runners), compared with 2 hours and 29 minutes for all other runners. As a consequence, prize money is higher. The average prize won by a Top100 runner is $18,933 compared with only $3,691 for all other runners. We can see that runners only enter between one and two races per year, suggesting that races must be selected carefully for runners to maximize their earnings. Finally, the percentage of Africans is considerably higher amongst Top100 runners (25% compared with 15%). This lends further support to our claim that Africans constitute the most able runners.

Our analysis proceeds in two steps. In the first step we test Proposition 1 by considering the choice between a Top5 and a Non–Top5 race at the individual runners' level. In the second step we test Proposition 2 by looking at the aggregate distribution of runners between these two race categories.

4.2 Individual race choice

To test Proposition 1, we investigate how a runner's expected payoff from a marathon, and hence his probability of entering, depends on the marathon's characteristics. Let-

entire analysis.

[8]We focus on Kenyan and Ethiopian runners since they are the most outstanding. Our results remain unchanged when we extend the set of African countries.

ting P_{ijt} denote the probability with which runner i enters race j in time period t, we estimate the following equation:

$$P_{ijt} = \alpha_0 + \alpha_B B_{jt} + \alpha_C C_{jt} + \alpha_A A_{jt-1} + X_i \beta + \varepsilon_{ijt}. \tag{7}$$

The variables of interest are the marathon's total prize budget, B_{jt}, the concentration of the prize structure, C_{jt}, and the proportion of African runners in the previous period, A_{jt-1}. Due to the dominance of African runners, A_{jt-1} can serve as an indicator of the level of opposition to be expected. We also include a vector of runner characteristics, X_i, which includes, age, nationality, gender, and a dummy variable indicating whether the race takes place on the runner's home turf. In addition, we control for time trends and race fixed effects.

Since we are interested in the runners' choice between entering a Top5 race or a Non–Top5 race, we let P_{ijt} take the value 1 whenever j is a Top5 race and the value 0 otherwise. According to Proposition 1, P_{ijt} should be increasing in B_{jt} and decreasing in A_{jt-1}. Hence α_B should be positive and α_A should be negative. The predictions of Proposition 1 with respect to the concentration of the contest's prize structure are more complicated. C_{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. We will address this issue later in the section and start by focusing on the other two components.

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 African they could have also been included in A_{jt-1}. In order to avoid the resulting endogeneity problem we restrict the analysis to the Top100 Non-African runners in a given year.[9] In particular, we estimate whether an increase in the fraction of African runners in race j in the previous year, reduces the likelihood with which a Top100 Non–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 without controls for runner characteristics, respectively. Column 3 includes trends (allowing for differential trends for females) and race fixed–effects. Overall, we find that an increase in expected opposition, leads to a decrease in the entry of high ability runners. This persists in all specifications. Total prize money, as we might expect, has a strong and positive effect on entry but with the inclusion of race fixed effects, the relation is no longer significant. This is most likely due to the high persistence of prize money differences across races.

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 African runners and to the Top100 Non-African 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.[10] This translates into an upward biased estimate of α_A. To avoid this problem, we need to instrument for the entry of African runners, A_{jt-1}. In other words, we need some exogenous variation in the entry of 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.[11] Both variables are correlated

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

[10]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.

[11]We use Kenian data since most of the high–ability African runners come from there. We also try alternative measures such as rainfall and commodity prices in Ethiopia and an average over all

with the number of Africans who compete in a given year but uncorrelated with race characteristics. Moreover, the race choice of Non-Africans 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. Brückner and Ciccone (2010) show that negative rainfall shocks, which are interpreted as economic recessions, are followed by significant improvement in democratic institutions. In addition, it has been documented by Deaton and Miller (1995) and Deaton (1999) that commodity price downturns cause rapidly worsening economic conditions in many Sub-Saharan African economies.

We construct an international commodity price index for Kenya following Deaton (1999) and Brückner and Ciccone (2010). Using the International Monetary Fund (2011) monthly international commodity price data for commodities exported by Kenya for the period 1986 to 2009 and Kenya's export share of these commodities taken from Deaton for 1990, we calculate the commodity price index. The rainfall data cover the period 1986 to 2009 and is taken from NASA Global Precipitation Climatology Project (GPCP). The rainfall data is that of Brückner and Ciccone (2010) but covers a longer time period.

We may also be concerned that race organizers adjust the total prize budget, B_{jt}, to attract high ability runners. We may expect that if entry falls, race organizers will increase prize money, such that the coefficient on B_{jt} is biased downwards. We deal with this by instrumenting the value of a race's prize budget using the exchange rate

Sub-Saharan countries, respectively, and the results are similar.

of the country where the race takes place relative to a currency basket. In order to construct a currency basket, we use the annual Special Drawing Rights (SDR) basket provided by the International Monetary Fund. The currency value of the SDR is determined using a basket of major currencies (U.S. dollar, Euro, Japanese yen, and pound sterling).[12]

In column 1 of Table 4, the first stage estimates show that rainfall and commodity prices are strongly related to the participation of Africans in international marathons. In particular, positive rainfall shocks and commodity price upturns (related to improved economic conditions) increase the number of Africans entering marathon competitions. This is intuitive since most runners, in particular the younger ones, rely on the support of sponsors, part of which are local businesses or regional government agencies. As expected, in column 2, we see that exchange rates are strongly related to total prize money.

In Table 5, we present the results for the IV estimates. We find that an increase in the proportion of Africans participating in a Top5 race in the previous year, is associated with lower entry of Top100 Non-African runners in the current year. 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 Non–African runner who expects 1% more African opposition is 0.77% less likely to choose a Top5 race. We also find that an increase in the total prize money is associated with a positive and significant effect on the entry of Top100 Non–African runners. In contrast to the OLS regression this effect persists in all specifications. Using both

[12]The weights assigned to each currency in the SDR basket are adjusted to take account of changes in the share of each currency in world exports of goods and services and international reserves.

instruments, our estimation predicts that an additional $1,000\$$ in total prize money raises the likelihood that a Top100 Non–African runner participates in a Top5 race by 1.5%. Given that African opposition refers to the fraction of African runners amongst the first 20 finishers, the participation of one additional African opponent is equivalent to a 5% increase in African opposition. This implies that Top100 Non–African runners are willing to forgo potential prize winnings of $2,567\$$ for each African opponent they are able to avoid.

4.3 Distribution of talent across marathons

Proposition 2 is concerned with the equilibrium distribution of talents across contests. To test Proposition 2, we analyze whether an increase in the overall number of African runners leads to a more balanced distribution of talent across races. More specifically, we test the following equation:

$$A_t^{T5} = \alpha_0 + \alpha_1 A_t + \alpha_2 B_t^{T5} + \varepsilon_t. \tag{8}$$

The dependent variable, A_t^{T5}, is the proportion of Africans in a given season who participate in Top5 races. The main variable of interest is the overall proportion of African runners in a given season, A_t. The variable B_t^{T5} denotes the proportion of the total prize money that is awarded in the Top5 races. We control for gender and time trends, as well as differential trends for male and female races.

Table 6 shows the estimates for equation (8). Since in our theoretical model the number of strong contests is identical to the number of weak contests, we first restrict our analysis to the Top10 races. In particular, columns 1 to 4, consider the runners' allocation between a Top5 race and the next five most important races. In columns 5 to 8, we consider the runners' allocation across all 34 races. The results are similar for both samples.

Overall, we find that as the overall proportion of Africans increases by 1%, the proportion of Africans entering a Top5 race falls by 1.03%. On the other hand, a 1% increase in the proportion of prize money awarded by the Top5 races, leads to an increase in the proportion of Africans entering a Top5 race (1.17% for Top10 races and 0.49% for all 34 races). It is reassuring that these effects persist when we control for time trends, gender and differential trends across gender.

We check the robustness of these results by using an alternative proxy for talent. Rather than using African 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 African origin is therefore that this definition of high ability is independent of prize money considerations. Table 7 shows that our main results still hold when we repeat the analysis for the alternative measure of ability based on rankings. We identify high ability runners as those runners who have a finishing time within 1% of the fastest finishing time during the season.[13] We also look at those finishing within 5% and 10% 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 African runners is changing over time.

[13]The identification of high ability runners is done separately for men and women and finishing times are adjusted for differences in race courses.

5 Conclusion

How do contestants choose in which contest to compete? And how do they value potential prize offerings relative to expected opposition? Do they prefer contests 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 the overall distribution of talent 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. We have provided empirical support for our results using data about the race choice of professional marathon runners. Making use of an organizational change and the resulting inflow of high ability contestants, we have shown that high ability runners have become less focussed on the top races.

Appendix 1 - Proofs

Proof of Proposition 1

It is immediate that E_H^i is increasing in w_i and decreasing in p_i. To prove the last claim of Proposition 1, increase the steepness of contest i's prize structure by letting

$\tilde{M}_i < M_i$ and $\tilde{b}_i > b_i$ and consider

$$\begin{aligned}\frac{E_H^i - \tilde{E_H^i}}{1 - \frac{c_H}{c_L}} &= b_i G(M_i, p_i) - \tilde{b}_i G(\tilde{M}_i, p_i) \\ &= b_i \sum_{m=0}^{M_i - 1} \binom{N-1}{m} p_i^m (1-p_i)^{N-1-m} - \tilde{b}_i \sum_{m=0}^{\tilde{M}_i - 1} \binom{N-1}{m} p_i^m (1-p_i)^{N-1-m} \\ &= b_i \text{Prob}(\tilde{M}_i \leq H_i \leq M_i - 1) - (\tilde{b}_i - b_i)\text{Prob}(H_i \leq \tilde{M}_i - 1).\end{aligned} \tag{9}$$

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

Proof of Proposition 2

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

$$E_p[H|H \leq M-1] = \sum_{m=0}^{M-1} \binom{N-1}{m} p^m (1-p)^{N-1-m} m \tag{10}$$

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

$$E_p[H] = p(N-1) \tag{11}$$

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

$$\Delta = b_S G(M_S, 2hq) - b_W G(M_W, 2h(1-q)). \tag{12}$$

We have

$$\begin{aligned}
\frac{dG(M,p)}{dp} &= \sum_{m=0}^{M-1} \tbinom{N-1}{m}[mp^{m-1}(1-p)^{N-1-m} - (N-1-m)p^m(1-p)^{N-2-m}] && (13)\\
&= \sum_{m=0}^{M-1} \tbinom{N-1}{m} p^{m-1}(1-p)^{N-2-m}[m-(N-1)p] && (14)\\
&= \frac{1}{p(1-p)}\left\{E_p[H|H \le M-1] - E_p[H]G(M,p)\right\} < 0. && (15)
\end{aligned}$$

It follows that

$$\frac{d\Delta}{dq} = 2h\left[b_S \frac{dG(M_S, p_S)}{dp} + b_W \frac{dG(M_W, p_W)}{dp}\right] < 0. \tag{16}$$

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

$$\Delta(q=0) = b_S - b_W G(M_W - 1, 2h) > 0. \tag{17}$$

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

$$\Delta(q = 1) = b_S G(M_S - 1, 2h) - b_W. \tag{18}$$

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

$$\begin{aligned} h\frac{d\Delta}{dh} &= \left[b_S p_S \frac{dG(M_S, p_S)}{dp} - b_W p_W \frac{dG(M_W, p_W)}{dp} \right] && (19) \\ &= \frac{b_S}{1 - p_S} \{ E_{p_S}[H | H \leq M_S - 1] - E_{p_S}[H] G(M_S, p_S) \} && (20) \\ &- \frac{b_W}{1 - p_W} \{ E_{p_W}[H | H \leq M_W - 1] - E_{p_W}[H] G(M_W, p_W) \}. \end{aligned}$$

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

$$\begin{aligned} \frac{h}{b_W G(M_W, p_W)} \frac{d\Delta}{dh} &= \frac{1}{1 - p_S} \{ \frac{E_{p_S}[H | H \leq M_S - 1]}{G(M_S, p_S)} - E_{p_S}[H] \} && (21) \\ &- \frac{1}{1 - p_W} \{ \frac{E_{p_W}[H | H \leq M_W - 1]}{G(M_W, p_W)} - E_{p_W}[H] \}. \end{aligned}$$

The first term is negative and strictly decreasing in p_S. Hence for $p_S \geq p_W$ we can find an upper bound by setting $p_S = p_W$ to get

$$\frac{h(1 - p_W)}{b_W G(M_W, p_W)} \frac{d\Delta}{dh} \leq \frac{E_{p_W}[H | H \leq M_S - 1]}{G(M_S, p_W)} - \frac{E_{p_W}[H | H \leq M_W - 1]}{G(M_W, p_W)}. \tag{22}$$

The right hand side is negative for all p_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 $\Delta(q) = b_S G(M_S - 1, p_S) > 0$. 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 $\Delta(q) = -b_W G(M_W - 1, p_W) < 0$. 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 (12), 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

Table 1: Descriptive Statistics (Races Level)

	Top5 Races					All other Races				
Variable	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Total Prize	216	228,206	116,439	15,500	603,500	735	43,947	42,534	1,000	287,000
Prize Steepness (H3)	216	0.40	0.05	0.33	0.70	629	0.41	0.06	0.33	0.75
Prize Steepness (1:2)	216	0.52	0.08	0.33	0.83	629	0.53	0.08	0.24	0.86
Fraction of Africans	228	0.19	0.18	0	0.65	1295	0.14	0.22	0	1
Fraction of HA (1%)	228	0.03	0.06	0	0.40	1295	0.00	0.02	0	0.25
Fraction of HA (5%)	228	0.29	0.26	0	1	1295	0.07	0.16	0	1
Fraction of HA (10%)	228	0.67	0.29	0	1	1295	0.35	0.36	0	1

Notes: Means and standard deviations for the Top5 races and all other 34 races, respectively. The sample period is 1986 to 2009. “Total Prize” is the total prize awarded in a race (in US$). “Prize Steepness (H3)” is the Hirfendel index of the top three prizes. “Prize Steepness (1:2)” is the ratio between the first and second prize. “Fraction of HA” refers to the fraction of runners finishing within 1%, 5% and 10% of the best time of the year, respectively.

Table 2: Descriptive Statistics (Runner level)

	Rank within Top 100					All other Runners				
Variable	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Female	4551	0.50	0.50	0	1	12377	0.43	0.49	0	1
Age	4551	29.77	4.56	16	50	12377	31.08	5.63	16	56
African	4551	0.24	0.43	0	1	12377	0.15	0.36	0	1
No. Races	4551	1.43	0.60	1	5	12377	1.27	0.55	1	6
Prize	3342	18,933	26,032	200	255,000	4876	3,691	4,921	40	100,000
Finish Time	4551	2:20	0:10	2:02	2:41	12377	2:29	0:14	2:04	3:36

Notes: Means and standard deviations for runners who ranked in the top 100 in a given year (separate for men and women) and all other runners, respectively. The sample period is 1986 to 2009 and uses all 34 races. “No. of Races” is the number of races run in a given year. “Prize” is the amount a runner is awarded (on average) in a race (in US$). “Finish Time” refers to time in (hours:minutes). The finishing times have been converted using ARRS conversion factors to ensure that times are comparable.

Table 3: Enter Top Five Race (OLS)

VARIABLES	OLS Enter Top5	OLS Enter Top5	OLS Enter Top5
Prop. Africans (t-1)	-0.1783***	-0.1147***	-0.0655*
	[0.035]	[0.040]	[0.035]
Total Prize ('00000)	0.1807***	0.1813***	0.0047
	[0.005]	[0.005]	[0.006]
H3 Prize	-0.6394***	-0.7068***	0.1206
	[0.120]	[0.122]	[0.089]
Female		0.0531***	-0.0016
		[0.014]	[0.020]
Age		-0.0002*	-0.0001**
		[0.000]	[0.000]
Nationality		-0.0010***	-0.0003*
		[0.000]	[0.000]
At Home		0.0597	-0.0011
		[0.079]	[0.047]
Trend			-0.0001
			[0.001]
Trend*Female			0.0002
			[0.0012]
Constant	0.7142***	0.7579***	0.3119**
	[0.049]	[0.051]	[0.143]
Race Fixed Effects	No	No	Yes
Observations	4153	4137	4137
R-Squared	0.224	0.229	0.738

Notes: * denotes significance at the 10% level, ** denotes significance at the 5% and *** denotes significance at the 1% level. “Prop. Africans (t-1)”, is the proportion of Africans (out of all runners) in a given race in the previous year. The proportion of Africans is calculated separately for men and women. “Total Prize (‘00000)” is measured in US$. “H3 Prize” measures the prize structure steepness (i.e., the Herfindahl-Hirschman index, calculated for the top three prizes).

Table 4: First Stage IV

VARIABLES	Prop. Africans (t-1)	Total Prize ('00000)
Commodity Price Index (t-1)	0.0010***	
	[0.000]	
Log Rainfall (t-1)	0.0728***	
	[0.027]	
Exchange Rates		0.0003***
		[0.000]
Constant	-1.1673***	-2.8839***
	[0.176]	[0.384]
Controls	Yes	Yes
Trend	Yes	Yes
Trens*Female	Yes	Yes
Race Fixed Effects	Yes	Yes
Observations	4,137	4,137
R-Squared	0.684	0.723
F-test of excl. instr.	44.41	28.49

Notes: * denotes significance at the 10% level, ** denotes significance at the 5% and *** denotes significance at the 1% level. "Commodity Price Index (t-1)" is the commodity prices in Kenya the previous year, using international commodity price data from IMF. "Log Rainfall" is the log of the amount of rainfall in Kenya in the previous year. "Exchange Rate" is the exchange rate of the country of the race relative to the exchange rates of the others. We construct a basket of currencies using the annual Special Drawing Rights (SDR) basket provided by the IMF.

Table 5: Enter Top Five Race (IV)

VARIABLES	IV Prop. Africans (t-1) Enter Top 5	IV Total Prize ('00000) Enter Top 5	IV Both Enter Top 5
Prop. Africans (t-1)	-0.7860***	-0.1032**	-0.7678***
	[0.255]	[0.043]	[0.238]
Total Prize ('00000)	0.0084	0.2006**	0.1543**
	[0.006]	[0.078]	[0.065]
H3 Prize	0.1505	-0.3642*	-0.213
	[0.093]	[0.218]	[0.183]
Female	0.0435*	-0.0003	0.0415*
	[0.026]	[0.023]	[0.025]
Age	-0.0001*	-0.0002**	-0.0002**
	[0.000]	[0.000]	[0.000]
Nationality	-0.0003*	-0.0008***	-0.0006***
	[0.000]	[0.000]	[0.000]
At Home	0.0054	0.0071	0.0111
	[0.049]	[0.053]	[0.047]
Trend	0.0157***	-0.0139**	0.0044
	[0.006]	[0.006]	[0.007]
Trend*Female	-0.0092***	0.0030*	-0.0064*
	[0.003]	[0.002]	[0.003]
Constant	-0.0629	0.8798***	0.7894***
	[0.178]	[0.057]	[0.063]
Race Fixed Effects	Yes	Yes	Yes
Observations	4137	4137	4137
R-Squared	0.712	0.664	0.739

Notes: * denotes significance at the 10% level, ** denotes significance at the 5% and *** denotes significance at the 1% level. “Prop. Africans (t-1)”, is the proportion of Africans (out of all runners) in a given race in the previous year. The proportion of Africans is calculated separately for men and women. “Total Prize (‘00000)” is measured in US$. “H3 Prize” measures the prize structure steepness (i.e., the Herfindahl-Hirschman index, calculated for the top three prizes). The instruments for the “Prop. Africans (t-1)” and “Total Prize (‘00000)” are described in the notes of Table 4.

Table 6: Proportion of Africans in Top 5 Races

	Top 10 Races			All 34 Races		
VARIABLES	Prop. Afr. (T5)	Prop. Afr. (T5)	Prop. Afr. (T5)	Prop. Afr. (T5)	Prop. Afr. (T5)	Prop. Afr. (T5)
Prop. Africans	-0.7742***	-0.3551**	-1.0272**	-0.6214***	-0.5321***	-1.2260**
	[0.187]	[0.171]	[0.501]	[0.131]	[0.125]	[0.471]
Prop. of Prize (Top 5)		1.1128***	1.1749***		0.4822***	0.4887***
		[0.190]	[0.195]		[0.139]	[0.138]
Female	-0.0894*	-0.0734*	-0.2516	-0.009	-0.0297	0.0303
	[0.050]	[0.042]	[0.153]	[0.035]	[0.033]	[0.113]
Trend			0.0125			0.0250*
			[0.017]			[0.015]
Trend*Female			0.0014			-0.0099
			[0.008]			[0.006]
Constant	0.8727***	-0.2134	-0.1688	0.4619***	0.1331	-0.0424
	[0.097]	[0.202]	[0.280]	[0.065]	[0.113]	[0.175]
Observations	79	79	79	79	79	79
R-squared	0.19	0.448	0.471	0.275	0.375	0.399

Notes: * denotes significance at the 10% level, ** denotes significance at the 5% and *** denotes significance at the 1% level.Years from 1986 until 2009. The dependent variable, “Prop. Afr. (T5)”, is the proportion of Africans (out of all Africans) in the Top 5 races. “ Prop. Africans”, is the proportion of Africans (out of all runners). The proportion of Africans is calculated separately for each race season. The proportion of Africans is calculated separately for men and women. “Prop. of Prize (Top 5)” is the proportion of total prize money in the Top% races. Prizes are measured is US$.

Table 7: Proportion of High Ability Runners in Top 5 Races

	Top 10 Races			Tope 34 Races		
VARIABLES	Prop. HA (T5) 1%	Prop. HA (T5) 5%	Prop. HA (T5) 10%	Prop. HA (T5) 1%	Prop. HA (T5) 5%	Prop. HA (T5) 10%
Proportion High Ability	-1.9806***	-0.2709*	-0.1143	-4.4436**	-0.7164***	-0.2960***
	[0.704]	[0.158]	[0.144]	[2.142]	[0.212]	[0.109]
Proportion of Prize (Top 5)	0.3174*	1.0407***	1.1402***	1.2970***	0.7094***	0.4466***
	[0.175]	[0.125]	[0.118]	[0.284]	[0.139]	[0.082]
Female	-0.1633	-0.0056	-0.0409	0.1434	-0.0994	-0.1557*
	[0.128]	[0.105]	[0.121]	[0.221]	[0.111]	[0.082]
Trend	-0.0195**	-0.0136**	-0.0035	0.0177	-0.0170**	-0.0163***
	[0.008]	[0.007]	[0.007]	[0.014]	[0.007]	[0.005]
Trend*Female	0.0104*	0.0025	-0.0008	-0.009	0.006	0.0062**
	[0.006]	[0.005]	[0.004]	[0.010]	[0.005]	[0.003]
Constant	1.0461***	0.0791	-0.1408	-0.368	0.4347*	0.5654**
	[0.257]	[0.241]	[0.333]	[0.376]	[0.238]	[0.219]
Observations	79	79	79	79	79	79
R-squared	0.31	0.716	0.692	0.354	0.603	0.618

Notes: * denotes significance at the 10% level, ** denotes significance at the 5% and *** denotes significance at the 1% level.Years from 1986 until 2009. The dependent variable, "Prop. HA. (T5)", is the proportion of high ability runners in the Top 5 races. "Prop. HA", is the proportion of high ability runners (out of all runners). The proportion of high ability runners is calculated separately for each race season. The proportion of high ability runners is calculated separately for men and women. High ability runners, refers to the proportion of runners finishing within 1%, 5% and 10% of the best time of the season, respectively. "Prop. of Prize (Top 5)" is the proportion of total prize money in the Top% races. Prizes are measured is US$.

References

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

[2] Arnott, R., Rowse, J. (1987) Peer Group Effects and Educational Attainment. *Journal of Public Economics* **32**, 287-305.

[3] Barut, Y., Kovenock, D. (1998) The Symmetric Multiple Prize All–Pay Auction with Complete Information. *European Journal of Political Economy* **14**(4), 627-644.

[4] 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.

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

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

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

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

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

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

[11] Deaton, A., Miller, R., (1995) International Commodity Prices, Macroeconomic Performance and Politics in Sub–Saharan Africa. *Princeton Studies in International Finance*, No. **79**.

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

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

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

[15] Groh, C., Moldovanu, B., Sela, A., Sunde, U. (2008) Optimal Seedings in Elimination Tournaments. *Economic Theory*, forthcoming.

[16] Harbring, C., Irlenbusch, B. (2003) An Experimental Study on Tournament Design. *Labour Economics* **10**, 443-464.

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

[18] Lazear, E., P., Rosen, S. (1981) Rank–Order Tournaments as Optimal Labor Contracts. *Journal of Political Economy* **89**(5), 841-864.

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

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

[21] Leuven, E., Oosterbeek, H., Sonnemans, J., van der Klaauw, B.(2008) Incentives versus Sorting in Tournaments: Evidence form a Field Experiment. *Journal of Labor Economics*, forthcoming.

[22] McAfee, R., P. (1993) Mechanism Design by Competing Sellers. *Econometrica* **61**, 1281-1312.

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

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

[25] Moldovanu, B., Sela, A., Shi, X. (2008) Competing Auctions with Endogenous Quantities. *Journal of Economic Theory*, **141**, 1-27.

[26] Peters, M., Severinov, S. (1997) Competition among Sellers Who Offer Auctions Instead of Prices. *Journal of Economic Theory* **75**, 114-179.

[27] Rosen, S. (2001) Prizes and Incentives in Elimination Tournaments. *American Economic Review* **76**(4), 701-715.

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

[29] Taylor, C., R. (1995) Digging for Goldon Carrots: An Analysis of Research Tournaments. *American Economic Review* **85**(4), 872-890.

[30] 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.