THE PREDICTIVE POWER OF NOISY ELIMINATION TOURNAMENTS

Dmitry Ryvkin

CERGE-EI

Charles University
Center for Economic Research and Graduate Education
Academy of Sciences of the Czech Republic
Economics Institute

WORKING PAPER SERIES (ISSN 1211-3298)
Electronic Version

252

Working Paper Series 252
(ISSN 1211-3298)

The Predictive Power of Noisy Elimination Tournaments

Dmitry Ryvkin

CERGE-EI
Prague, March 2005

ISBN 80-7343-045-2 (Univerzita Karlova v Praze, CERGE)
ISBN 80-7344-034-2 (Národohospodářský ústav AV ČR, Praha)

THE PREDICTIVE POWER OF NOISY ELIMINATION TOURNAMENTS

Dmitry Ryvkin[1]

CERGE-EI
Charles University and Academy of Sciences of the Czech Republic

March 11, 2005

Abstract
An elimination tournament matches players pairwise and promotes the winners to a subsequent round where the procedure is repeated. In the presence of idiosyncratic noise the tournament turns into a probabilistic mechanism that reveals the ranking of players imperfectly. I assess theoretically the power of such a mechanism to determine the *ex ante* best player as the winner, as a function of the number of players, their distribution of type, and the noise level. I consider also various seeding strategies and show that for large and small noise (as compared to the variance of ability distribution among players), seeding and other control parameters of tournament design tend to play no role, whereas for intermediate noise level the predictive power depends strongly on the control parameters and therefore can be systematically manipulated by the principal.

Keywords: elimination tournaments, noise, seeding, ability distributions, design economics
JEL Classification: C73, C90, D21

Abstrakt
Ve vyřazovacím turnaji se hráči utkávají ve dvojicích a vítěz postupuje do dalšího kola, kde se procedura znovu opakuje. Z přítomnosti idiosynkratického šumu se z turnaje stává pravděpodobnostní mechanismus, který nedokáže vždy zcela odhalit správné pořadí jednotlivých hráčů. V závislosti na počtu hráčů, rozdělení jejich typů a intenzitě šumu hodnotím sílu jednotlivých mechanismů ustanovit *ex ante* nejlepšího hráče vítězem. Uvažuji různé nasazovací strategie, které se ukazují podstatnými při ovlivňování predikční síly mechanismu. Rovněž ukazuji že pro příliš vysoké i příliš nízké úrovně šumu (ve srovnání s variancí rozdělení hráčských dovedností) nasazování i další kontrolní parametry designu turnaje hrají pouze zanetbatelnou roli, zatímco pro střední urovně šumu předikční síla velmi závisí na kontrolních parametrech a tedy může být systematicky ovlivňována organizátorem.

[1]E-mail: dmitry.ryvkin@cerge-ei.cz. Address: P.O. Box 882, Polítických vězňů 7, 11121 Prague 1, Czech Republic. The author thanks Andreas Ortmann and Dirk Engelmann for their helpful comments.

1 Introduction

It is often unavoidable to rank agents. Prominent examples are governmental agencies that seek to allocate money to the best research group, or managers who want to promote the best worker, or employers who want to hire the best candidate. In such cases, principals extensively use *tournaments*.

A tournament is a mechanism that belongs to a special class of principal-agent games, or games of asymmetric information, which necessarily involve *comparisons* rather than cardinal measurements of players' performance. In a tournament, incentives are created by prizes that are usually fixed in advance (for an exception, see Baye and Hoppe 2003), with the allocation of prizes being based on the *ex post ranking* of players. The tournament, therefore, *reveals* the ranking of its participants.

There are various reasons why tournaments are often preferred to other forms of incentive provision (such as piece-rate compensation): filtering out common productivity shocks, low monitoring costs, the ability to measure relative performance when absolute performance is difficult to evaluate. For a review on tournaments in labor markets, see Prendergast (1999) and Lazear (1999). For a discussion of rent-seeking and innovation tournaments, see Taylor (1995), Fullerton and McAfee (1999), Baye and Hoppe (2003). Tournaments are also the essence of sporting events (for a recent review see Szymanski 2003).

The problem of optimal tournament design, which belongs to the newly emerging field of design economics (see e.g. Roth 2002), arises naturally in the various contexts mentioned above. This problem usually includes the design of the following elements: (i) the tournament format, i.e. the rules by which players are matched and the ranking is determined; and (ii) the prizes, i.e. how many prizes should be awarded, and what their values are. The solution, of course, depends on the organizer's objectives and the characteristics of the participating players.

In economic environments, the organizer's objectives usually are the maximization of total output or maximization of total effort (Gradstein and Konrad 1999). In sporting tournaments, especially in team sports, some other objectives are often present, such as win maximization, or competitive balance maximization (see Szymanski 2003). Players usually differ by their ability (or productivity), and hence their disutility of effort.

In most of the existing literature the issue of the selection of the optimal tournament format is ignored. Authors mostly consider *contests*, i.e. tournaments where all players perform only once. In contest tournament models, it is usually of interest what the optimal prize structure is under various forms of heterogeneity of players. The analogue of a contest in sports is, for example, a marathon race. Also, perfectly discriminating contests are, in fact, equivalent to what is known as *all-pay auctions*.

Few authors analyze more complicated tournament formats. Moldovanu and Sela (2002) consider *parallel contests*, in which players are divided into $t > 1$ equal groups, and compete for a prize V/t within each group (V is the total value of the prize). It is shown that for sufficient heterogeneity and/or sufficient convexity of players' effort cost functions such an architecture is preferred to the contest where the single prize V is awarded. This result shows that at least in some circumstances more complicated tournament formats are superior, and therefore it makes sense to study them more carefully.

Clearly, the contest format cannot model tournaments in hierarchical organizations, where winners are promoted to higher levels and compete in many stages. These formats, known as *up-or-out* rules, are common e.g. for lawyers (see O'Flaherty and Siow (1995), or in the academy. Rosen (1986) used the *elimination* format to study such a tournament.

The elimination (also known as knock-out or Olympic) tournament format is organized as follows. There are R rounds. In the first round $N = 2^R$ players are matched in pairs, and the winner of each pair advances to the next round. In the second round the remaining 2^{R-1} players are matched in pairs again, and the winners again advance to the next round. The procedure is repeated R times until only one player is left, who is the winner of the whole tournament. In sports, the rounds have names: the R-th round (where only one match is played) is the *final*; in the $(R-1)$-th round, two matches called *semifinals* are played; in the $(R-2)$-th round four matches called *quarterfinals* are played, etc.

In the model by Rosen (1986) all players of the elimination tournament get prizes (the winner gets W_1; the loser of the final gets $W_2 < W_1$; the losers of the semifinals get $W_3 < W_2$, etc.). It is shown that in order to maintain a high effort level throughout all stages of the tournament the principal has to increase the wage differential $W_i - W_{i+1}$ towards the top. The increasingly high wage differentials at the top of hierarchical organizations are also reported empirically by Bognanno (2001); Ehrenberg and Bognanno (1990) report high prizes at the top of multistage sporting tournaments.

Gradstein and Konrad (1999) construct a fairly general format of a multi-stage tournament, in which arbitrary groupings of winners occur at every stage. The probability to be the winner of a group is modelled by the logit-like function of players' effort levels (see Tullock 1980; Rosen 1986 uses a generalized form of it for two players in every match). It is shown that in order to maximize the total effort level in the tournament (i) within each stage, group sizes must be equal, and (ii) the group size must be the same across rounds. The optimal number of rounds is a corner solution, i.e. it is either maximal or minimal possible (the elimination format or the contest format, respectively), depending on the discriminatory power of the logit contest success function.

These results show that the elimination tournament is an important benchmark format, which might be preferred to other formats in some situations. In addition to creating incentives, tournaments provide principals with *information* on the ranking of agents. This information is mainly ignored in the literature, notwithstanding its usefulness. I argue that there are situations where accurate ranking information might be the organizer's only concern.

When a tournament is used to create incentives for higher effort, and the organizer's objective is the total output (or effort) maximization, it is actually unimportant *who* specifically is rewarded by the first prize. There will be no loss for the organizer if the revealed best agent is not indeed the best but performed better than others at that particular time due to randomness. In such cases, a contest is probably the most efficient solution. However, in some circumstances, especially when a tournament is used for promotion purposes, it is the organizer's objective to reward *truly the best* player, because the long-term productivity of the firm is at stake. A natural objective for the organizer then becomes the *predictive power* optimization.

The notion of predictive power can be easily understood if one thinks about a tournament as an *estimator* of the true ranking of players. Indeed, because of the heterogeneity of abilities there exists an underlying (and unknown) true ordering of players by their ability level, and it is this true ordering that the organizer wants to know, e.g. in order to promote the best player. Due to many environmental factors that are impossible to measure and take into account, the performance of players at each particular moment fluctuates around their true effort levels (or ability levels, given the correspondence between the two). The ranking revealed by the tournament is, therefore, in general different from the true ordering; it *estimates* the true ordering, with some *bias* and *variance*. If the tournament is fair (in the sense O'Keeffe, Viscusi, and Zeckhauser 1984 use this term, that is, the principal treats all agents equally), the estimator is clearly unbiased. The variance is hard to define, because it is not clear what is the *distance* between two different rankings, but for the purposes of this paper it is not necessary. I will concentrate on the simplest case assuming that the principal only cares about estimating *the best* player correctly, and her objective is therefore to maximize the probability for the best player to win the tournament.

The probability for a particular player to advance in the elimination tournament was calculated by Knuth (1987). Assumptions made in that paper are, however, very restrictive. First, it is assumed that in any match where ranks differ by more than one, the stronger player wins unambiguously. Second, the players are re-shuffled randomly before every round, which means that the impact of seeding is ignored.

The importance of seeding was noted by Rosen (1986), who, however, considered only the averaged random seeding case. As I show, it is possible to analyze the problem of seeding in a consistent manner. In fact, for three prominent distributions of players' abilities, and $N = 4, 8$ players, I find numerically the *optimal* seeding, which maximizes the probability of the best player winning.

Horen and Reizman (1985) showed that in general the predictive power maximizing seeding only exists for $N = 4$ players. For $N = 8$ and larger (i.e. for tournaments of 3 or more rounds) the optimal seeding depends on the configuration of players' heterogeneity.

The present paper contributes to the literature on tournaments by (i) considering the elimination tournament of heterogeneous players with arbitrary pairwise winning probabilities; (ii) considering the new organizer's objective, the predictive power, which I argue to be important in applications; (iii) using a new approach to modelling the heterogeneity of players, the ability distribution and noise generating the winning probabilities; (iv) analyzing the influence of seeding on the tournament result as measured by the predictive power.

The paper is organized as follows. In Section 2 the winning probabilities are discussed, and a model is presented to parameterize them in terms of ability distribution and noise. In Section 3 the elimination tournament model is set up, and the key questions to be addressed are more precisely formulated. In Section 4 the recurrence relation for the predictive power is obtained, which is the first major result of the paper. In Section 5, the problem of optimal seeding is discussed. In Section 6, the results for the predictive power for three prominent distributions of players' abilities are presented and discussed, which constitutes the second major result. In Section 7, the results for the same distributions are presented for seeded players. Section 8 concludes.

2 Distribution of abilities, noise, and winning probabilities

In an elimination tournament, players are matched pairwise. Depending on seeding and the way the tournament unfolds, there is a possibility that any two players $i, j \in \{1, \ldots, N\}$ are matched. The fundamental quantities that characterize the probabilities of the advancement of players in the tournament are, therefore, the *winning probabilities* w_{ij} defined as

$$w_{ij} = \Pr\{\text{player } i \text{ beats player } j\}. \tag{1}$$

No ties are possible,[2] so $w_{ij} = 1 - w_{ji}$, and only $N(N-1)/2$ of the winning probabilities are independent.

The winning probabilities can be given exogenously, for example, through past statistics, or they can be obtained through a Bayesian updating procedure. In either case, they are statistically well-defined quantities. Parameterization of a tournament in terms of winning probabilities is, however, too complex. Instead, it would be desirable to specify a small number of intuitively interpretable parameters and derive the winning probabilities from some underlying statistical *model*.

One way to model the winning probabilities is to specify a functional form (as done by e.g. Rosen 1986 and many other authors) for w_{ij} as a function of effort levels and/or abilities of players i and j. This approach is appealing for its simplicity, but there is no obvious reason why a particular functional form (usually logit) should be chosen in the present context. Another way to model w_{ij} is to assume, following e.g. Lazear and Rosen (1981), that every player's performance y_i equals her effort level μ_i plus a random noise term: $y_i = \mu_i + \epsilon_i$. The optimal effort level for every player in such models is determined through utility maximization, and it is assumed that players have heterogeneous costs of

[2] The elimination tournament format requires that be broken, as opposed to other formats, e.g. the round-robin format.

effort (or, equivalently, heterogeneous prize valuations). It is reasonable to assume that heterogeneity in costs of effort in fact corresponds to the heterogeneity in abilities, and under natural assumptions on the shapes of utility and cost functions the optimal effort levels are determined uniquely by abilities. For example, in sporting events, issues of asymmetric information are less important, and hence can be assumed away. Specifically, it is likely that abilities and efforts are highly correlated (see, for example, the empirical study by Ehrenberg and Bognanno 1990). Moldovanu and Sela (2002) set the performance equal to effort, $y_i = \mu_i$, but assume the cost function of effort in the form $C(\mu_i) = c_i\gamma(\mu_i)$, with some distribution of c_i in the population, which, following the arguments above, translates into a distribution of abilities. O'Keeffe *et al.* (1984) model the performance as $y_i = A_i\mu_i$, where A_i is player's ability, and directly assume a distribution of abilities in the population. Given the correspondence between effort and ability, one of the two characteristics is redundant. It is more appealing to use abilities because (i) they are the underlying reason for different effort levels; (ii) it does not require any assumptions on the shape of utility and cost functions; (iii) some natural and empirically testable assumptions can be made about the distribution of abilities in the population. Reed (2001) proposed a model that explains why the Pareto distribution might be observed in many seemingly unrelated situations, with the ability distribution being one of them. Other "natural" candidates and well-established benchmarks are the normal distribution, and the uniform distribution.

Thus, I assume that there is a probability density function (pdf) $f(x)$ of abilities x in the population of players. Let $F(x)$ denote the corresponding cumulative density function (cdf). In order to set a scale of ability, suppose that $f(\cdot)$ is normalized so that $\mathrm{Var}[x] = 1$. N players with abilities $x_1, \ldots, x_N$ are drawn independently from the population. In order to infer how a tournament estimates the ranking of players, it is convenient to order players by their ability so that $x_1 > x_2 > \ldots > x_N$. It is further assumed that there exists an idiosyncratic normally distributed noise ϵ_i with mean zero and variance σ^2 that additively distorts the performance of players, so that in every match player i's output is $y_i = x_i + \epsilon_i$. Then, given the abilities of players i and j, it is easy to calculate the probability for player i to beat player j:

$$w_{ij}(x_i, x_j) = \Phi\left(\frac{x_i - x_j}{\sigma\sqrt{2}}\right). \tag{2}$$

Here $\Phi(\cdot)$ is the cdf of the standard normal distribution. Note that Eq. (2) is not an arbitrary functional form for winning probabilities, but a consequence of an intuitively appealing statistical model. The reason why exactly the normal density is most likely to be the distribution of noise is, essentially, the central limit theorem. The probabilities (2) are conditional on the given levels of ability x_i and x_j. In order to obtain the unconditional winning probabilities, one needs to average over all possible realizations of $x_1, \ldots, x_N$, taking into account the fact that the players are ordered by their ability:[3]

$$w_{ij} = N! \int_{-\infty}^{\infty} dx_1 f(x_1) \int_{-\infty}^{x_1} dx_2 f(x_2) \ldots \int_{-\infty}^{x_{N-1}} dx_N f(x_N) w_{ij}(x_i, x_j). \tag{3}$$

It is shown in Ryvkin and Ortmann (2004) that Eq. (3) can be simplified to yield

$$w_{ij} = K_{ij} \int_0^1 dz_1 \int_0^{z_1} dz_2 (1 - z_1)^{i-1} (z_1 - z_2)^{j-i-1} z_2^{N-j} \Phi\left(\frac{F^{-1}(z_1) - F^{-1}(z_2)}{\sigma\sqrt{2}}\right). \tag{4}$$

Here $K_{ij} = N!/[(i-1)!(j-i-1)!(N-j)!]$, and $F^{-1}(\cdot)$ is the inverse cdf of abilities in the population,

[3]As mentioned in Ryvkin and Ortmann (2004), there may be alternative ways to think about the parameterization of the winning probabilities. For example, one can totally avoid the notion of the average winning probability w_{ij}, and explore various tournament outcomes in terms of the conditional probabilities $w_{ij}(x_i, x_j)$, and only in the end average over all possible realizations of $x_1, \ldots, x_N$. Still, the average winning probabilities w_{ij} are well-defined real-world quantities, while modelling them is a separate issue. One can think about the model parameters (e.g. in this case the distribution of players' abilities, the noise level, and the number of players) as the ones *generating* a particular configuration of average winning probabilities.

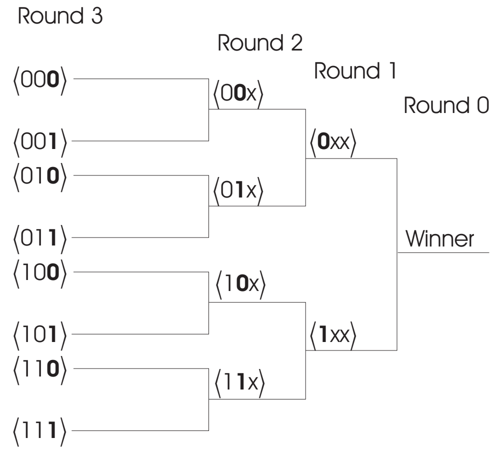

Figure 1: The binary tree representation for an elimination tournament of $N = 8 = 2^3$ players and $R = 3$ rounds. The players are enumerated by binary numbers. Symbol "x" denotes unknown bits of advancing players. In each match the key differing bits are shown bold.

3 General setup

An *elimination tournament* of $N = 2^R$ players can be conveniently represented in the form of a binary tree. Consider a binary tree with $N = 2^R$ terminal nodes. Enumerate the terminal nodes by R-bit binary numbers $b^{(k)} = \langle b_1^{(k)}...b_R^{(k)}\rangle$, where $b_i^{(k)} \in \mathcal{B} = \{0, 1\}$, and $k = 1, \dots, N$. Let $\mathcal{B}^R$ denote the set of the terminal nodes.

Let $\mathcal{P} = \{1, \dots, N\}$ be a set of N players. A *seeding function* $f_s : \mathcal{P} \to \mathcal{B}^R$ assigns a terminal node to every player. From now on, I refer to players by their position in the seeding (i.e., by the terminal nodes $b^{(k)}$ they are assigned to). In the end, the true "identities" of players can be restored as $i = f_s^{-1}(b^{(k)})$.

The elimination tournament is a sequential procedure that consists of R rounds $r = R, \dots, 1$.[4] In the R-th round matches $(b^{(1)}, b^{(2)})$, $(b^{(3)}, b^{(4)})$,..., $(b^{(N-1)}, b^{(N)})$ are played. Note that within every match the players' binary numbers differ only in the R-th bit. Indeed, every matched pair $(b^{(2k-1)}, b^{(2k)})$, $k = 1, \dots, 2^{R-1}$, in the binary representation looks like $(\langle b_1^{(2k-1)}...b_{R-1}^{(2k-1)}0\rangle, \langle b_1^{(2k-1)}...b_{R-1}^{(2k-1)}1\rangle)$.

[4] For convenience, I enumerate rounds backwards in time. As mentioned earlier, it is also common in sports to call the rounds close to the root of the tree by the number of matches ("brackets") in them, e.g. the final ($r = 1$, $2^{1-1} = 1$ match), semifinals ($r = 2$, $2^{2-1} = 2$ matches), quarterfinals ($r = 3$, $2^{3-1} = 4$ matches), etc., so that the exponents of 2 refer to the backward enumeration of rounds.

The winners of the R-th round's matches are promoted to the next round, $R-1$, where they are matched according to the flow along the branches of the tree towards the root. Formally, the matching occurs as follows: two players in round $R-1$ are matched if and only if their first $R-2$ bits coincide, and their $(R-1)$-th bits are different (see Fig. 1).

The procedure is repeated recursively. In round $R-i$ the winners from round $R-i+1$ are matched. Within each match first $R-i-1$ bits of players' binary numbers coincide, and the $(R-i)$-th bits are different. The tail bits are not important for matching, since $R-i$ bits suffice to enumerate the remaining players in the $(R-i)$-th round. Those bits, however, have to be dragged behind in order to preserve the initial identities of all surviving players.

In round 1, there is just one (final) match. The players of this match necessarily have their 1-st bits different. The winner of the final match is the winner of the whole tournament.

Fig. 1 shows the tree representation for an elimination tournament of 8 players.

Following Ryvkin and Ortmann (2004) and Section 2, suppose every match (i,j) (between players i and j in the original enumeration $\mathcal{P}$) is a Bernoulli trial that defines a random variable p_{ij} to be 1 with probability w_{ij} (player i beats player j) and 0 with probability $w_{ji} = 1 - w_{ij}$. Clearly, the winning probabilities w_{ij} can be translated into the winning probabilities among the seeding-enumerated players: $w_s^{(b^{(i)}b^{(j)})} = w_{f_s^{-1}(b^{(i)})f_s^{-1}(b^{(j)})}$, with the inverse transformation being $w_{ij} = w_s^{(f_s(i)f_s(j))}$.

In this setting, it is possible to formulate all kinds of problems involving probabilities of certain events. Here I will consider two such problems.

Problem 1. Given the winning probabilities w_{ij} and the seeding function $f_s(\cdot)$, find the probability A_r^i for player i to reach round r of the tournament. Particularly, find the probability $\rho_i \equiv A_0^i$ of winning the tournament.

Problem 2. Given the winning probabilities w_{ij}, find the seeding function $f_s^i(\cdot)$, such that the probability A_0^i for player i to win the tournament is maximized. Particularly, explore the existence and uniqueness of such a maximum.

4 The predictive power

Below I show how to calculate the probability for a player to advance up to a given round of the tournament, including the probability to win the whole tournament (i.e. to advance to round 0).

Let $A_r(b^{(i)})$ denote the probability for the player seeded as $b^{(i)}$ to advance to the r-th round. In the r-th round player $b^{(i)}$ can only be matched with a player whose binary number has the following properties: (i) the first $r-1$ bits are the same as in $b^{(i)}$; (ii) the r-th bit is different from that of $b^{(i)}$. Thus, the players who can potentially be matched with player $b^{(i)}$ in the r-th round have a binary number of the form $\langle b_1^{(i)} \dots b_{r-1}^{(i)} \overline{b_r^{(i)}} c_{r+1} \dots c_R \rangle$. Here the overline denotes a flipped bit: $\overline{0} = 1$ and $\overline{1} = 0$; $c_k \in \mathcal{B}$ are arbitrary bits. Also, the player to be matched with $b^{(i)}$ needs herself to survive until the r-th round. Thus, the following Theorem can be formulated.

Theorem. The probability $A_r(b^{(i)})$ for the player seeded $b^{(i)}$ to reach the r-th round solves the recurrence relation

$$A_{r-1}(b^{(i)}) = \sum_{c_{r+1},\dots,c_R \in \mathcal{B}} A_r(\langle b_1^{(i)} \dots b_{r-1}^{(i)} \overline{b_r^{(i)}} c_{r+1} \dots c_R \rangle) w_s^{(b^{(i)} \langle b_1^{(i)} \dots b_{r-1}^{(i)} \overline{b_r^{(i)}} c_{r+1} \dots c_R \rangle)} \qquad (5)$$

with initial condition $A_R(b^{(i)}) = 1$.

This recurrence relation is very intuitive. The probability of advancing to the next round is the sum of probabilities of beating all possible candidates for the match multiplied by their corresponding probabilities of reaching this point in the game. The $R \times N$ numbers $A_r(b^{(i)})$ completely define the probabilistic properties of the tournament.

The initial conditions simply imply that all N players with probability 1 start the tournament in the R-th round.

5 The optimal seeding problem

Suppose the tournament organizer has information about the *ex ante* ranking of players, and her objective is to seed them so that the predictive power (i.e. the probability ρ_1 for player 1 to win the tournament) is maximized.

The restrictions imposed so far on the winning probabilities regard their being conditioned on the *ex ante* ranking of players. Thus, it is required that for all $i, j, k \in \mathcal{P}$

$$\text{(a) } w_{ij} \geq w_{ik} \text{ for } j > k; \qquad \text{(b) } w_{ij} \geq 1/2 \text{ for } i < j. \tag{6}$$

Condition (6a) implies that any player i has a higher probability of winning against a weaker player. Condition (6b), in fact, is a definition of the term "stronger": it means that a stronger player in a match wins with a probability not less than 1/2.

Note that N players can be seeded in $N!$ various ways, of which, however, only $n_s = N!/2^{N-1}$ are non-trivially different. In what follows, by "seeding" I mean the whole set of equivalent seedings.

Horen and Reizman (1985) show that for an elimination tournament of $N = 4$ players conditions (6a,b), and additional no-tie condition $w_{ij} + w_{ji} = 1$, imply that the probability of the best player winning (the predictive power) is maximized by the seeding $\{(1,4),(2,3)\}$, [5] i.e. there exists a universal optimal seeding for 4 players.

[5]This can be proven by a direct check. For 4 players, there are three non-trivially different seedings: $\{(1,2),(3,4)\}$ [seeding (1)], $\{(1,3),(2,4)\}$ [seeding (2)], and $\{(1,4),(2,3)\}$ [seeding (3)]. The probability to win the tournament for player 1 equals for the three seedings $\rho_1^{(1)} = w_{12}(w_{13}w_{34} + w_{14}w_{43})$, $\rho_1^{(2)} = w_{13}(w_{12}w_{24} + w_{14}w_{42})$, and $\rho_1^{(3)} = w_{14}(w_{12}w_{23} + w_{13}w_{32})$. Then, it follows from conditions (6) and the identity $w_{ij} + w_{ji} = 1$ that

$$\rho_1^{(3)} - \rho_1^{(1)} = w_{14}w_{12}\left[w_{32}\left(\frac{w_{13}}{w_{12}} - 1\right) + w_{34}\left(1 - \frac{w_{13}}{w_{14}}\right)\right] \geq 0,$$

$$\rho_1^{(3)} - \rho_1^{(2)} = w_{14}w_{13}\left[w_{24}\left(1 - \frac{w_{12}}{w_{14}}\right) - w_{23}\left(1 - \frac{w_{12}}{w_{13}}\right)\right] \geq 0.$$

These two inequalities constitute the proof. Also, Horen and Reizman (1985) show that (i) the optimal seeding for 4 players is *fair*, i.e. the better a player is the higher is her probability of winning the tournament (in our notation, $\rho_1 \geq \rho_2 \geq \rho_3 \geq \rho_4$); (ii) the optimal seeding maximizes the probability that players 1 and 2 meet in the final under the necessary and sufficient condition that $w_{14}/w_{13} \geq w_{24}/w_{23}$;

A natural generalization of the optimal seeding for 4 players to the case of N players is

$$\{(1, N), (N-1, N-2), (N-3, N-4), \ldots, (3, 2)\}. \tag{7}$$

Seeding (7) looks appealing because it matches player 1 with weakest possible players at earlier stages, leaving better players for later stages, and thus, decreasing their survival probabilities. On the other hand, the better players meet their most powerful rivals at earlier stages.

Unfortunately, the situation becomes more complicated for $N > 4$ players. Already for $N = 8$, Horen and Reizman (1985) find that there is no universal optimal seeding, and seeding (7) is optimal only in some situations. Specifically, they show that there are 8 seedings,[6] seeding (7) being one of them, that can be optimal, depending on the structure of the matrix of winning probabilities w_{ij}. This somewhat counterintuitive result arises because of the multi-round structure of the tournament. For example, it may be optimal (in terms of ρ_1) to modify seeding (7) by bringing some of the better players to the upper half of the tree in exchange for increasing the probability of meeting weaker players at later stages.[7]

Another reason why there is no single optimal seeding for $R \geq 3$ rounds is that conditions (6) are too general. Indeed, it is not a coincidence that seeding (7) looks very appealing. All examples given by Horen and Reizman (1985) which fail (7) as the optimal seeding for 8 players, involve a violation of something that can be loosely called "smoothness" of the winning probabilities w_{ij}. For example, seeding $\{(1,8),(7,5),(6,4),(3,2)\}$, which differs from seeding (7) by the swap $5 \leftrightarrow 6$, is optimal for $w_{12} = w_{13} = 0.5$,

(iii) if the teams have "values" $v_1 \geq v_2 \geq v_3 \geq v_4$, then the optimal seeding maximizes the expected value of the winner $\sum_{i=1}^{4} \rho_i v_i$.

[6]The 8 seedings are $\{(1,8),(7,6),(5,4),(3,2)\}$, $\{(1,8),(7,5),(6,4),(3,2)\}$, $\{(1,8),(6,5),(7,4),(3,2)\}$, $\{(1,7),(6,5),(8,4),(3,2)\}$, $\{(1,7),(6,5),(8,3),(4,2)\}$, $\{(1,8),(6,5),(7,3),(4,2)\}$, $\{(1,8),(7,5),(5,3),(4,2)\}$, and $\{(1,8),(7,6),(5,3),(4,2)\}$.

[7]Horen and Reizman (1985) show that if player 1 has to be matched with the winner of a match between two other players, weakening the better of the two increases the probability of player 1 winning, but weakening the worse player of the two may actually decrease player 1's chances. Also, weakening *both* players may decrease player 1's chances.

$w_{14} = w_{15} = w_{16} = 0.9$, $w_{17} = w_{18} = 1$; $w_{23} = w_{24} = 0.5$, $w_{25} = w_{26} = 0.9$, $w_{27} = w_{28} = 1$, etc. In this example one can see that for every player, all other players can be divided into two distinct groups: the players of practically same ability, and the players who are almost surely defeated. Also, there is lack of transitivity. Players 1 and 2 are of the same ability, but player 1 beats player 4 with probability 0.9, while players 2 and 4 are again equal.

All other examples of seeding (7) being non-optimal, as provided by Horen and Reizman (1985), they have the same "degenerate" features. This suggests that if certain regularity conditions are imposed on w_{ij}, seeding (7) may become the only optimal seeding, at least for the case of 8 players. The question of what these regularity conditions are remains open. It is relatively straightforward, albeit cumbersome, to write down *some* sufficient conditions for any particular R, but of a much more considerable interest would be fairly simple, necessary, and sufficient conditions for arbitrary R. If such conditions exist, and what they are, is a challenging mathematical problem.

Section 7 presents the results of a numerical exploration of how seeding affects the predictive power in elimination tournaments. For the winning probabilities generated, as described in Section 2, by three prominent distributions of players' abilities and noise, I show that seeding (7) is optimal for $N = 8$ players for all values of noise intensity σ^2. I further show that for $N = 16, 32, 64, 128$ and 256 players, the predictive power calculated using seeding (7) is consistently larger than the average predictive power calculated for random seeding. In general, it is shown that for intermediate noise levels (not too large, and not too small compared to the variance of the ability distribution) seeding significantly affects the predictive power.

6 Results: no seeding

Here I start the analysis of the results with the no seeding case. Since the predictive power depends dramatically on seeding, the notion of "no seeding" implies "random seeding",

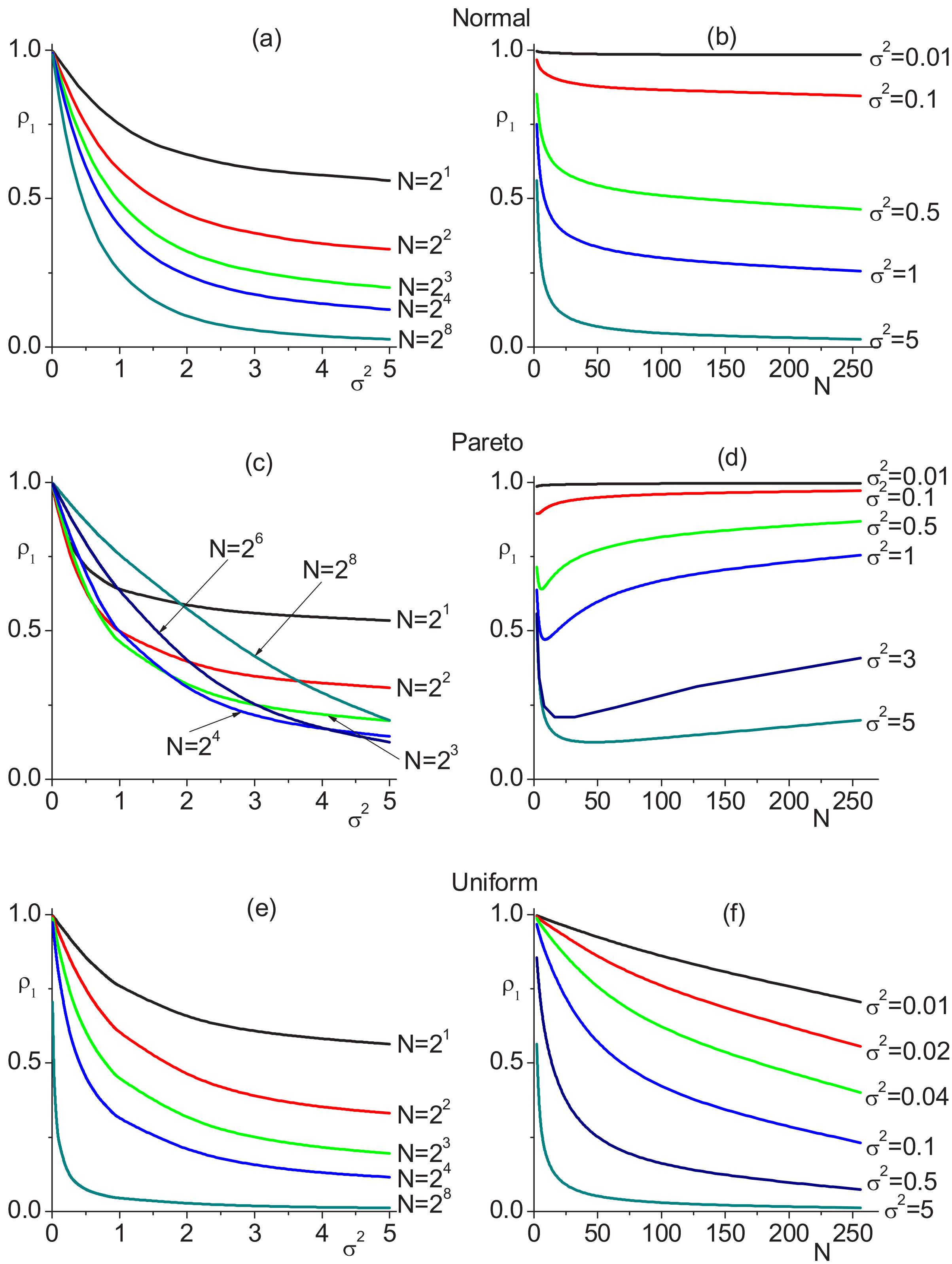

Figure 2: The predictive power ρ_1 as a function of the number of players N and noise intensity σ^2 for three ability distributions. There is no seeding. The figures are obtained using Eq. (5) averaged over a large number of random seedings.

or, strictly speaking, averaging over all possible seedings with equal weights. As already mentioned, out of $N!$ various seedings for N players only a small fraction are non-trivially different. The number of non-trivially different seedings, $n_s = N!/2^{N-1}$, is still a very large number for large N. It is possible to go through *all* the different seedings for $N = 4$ ($n_s = 3$) and $N = 8$ ($n_s = 315$), but not for $N = 16$ ($n_s = 638,512,875$) and larger. Therefore in calculations the averaging over seedings has to be done approximately, i.e. by randomly generating a large number K of different seedings and calculating the average over them. For sufficiently large K the relative statistical error of this approximation will on average decrease as $K^{-1/2}$.

In this Section I discuss the random seeding results for the predictive power ρ_1 as a function of the number of players N and the noise intensity σ^2 for three prominent distributions of players' abilities $f(x)$: normal, Pareto, and uniform. The dependence of ρ_1 on the parameters is shown in Fig. 2. The winning probabilities w_{ij} are calculated for each constellation of parameter values using Eq. (4). The predictive power is then obtained using recurrence relation (5). All three distributions are chosen so that their variance equals 1, to establish a unified scale of ability variations. It is therefore natural to distinguish three cases: (i) small noise, $\sigma^2 \ll 1$, (ii) intermediate noise, $\sigma^2 \sim 1$, and (iii) large noise, $\sigma^2 \gg 1$.

In the small noise case there is little uncertainty as to who is the winner in every match or in the tournament as a whole, i.e. ρ_1 is close to 1. This is reflected in the results (see Fig. 2). Of course, as N increases, the predictive power more and more differs from 1.

For the other extreme case, large noise, the winning probabilities w_{ij} all converge to 1/2, and the difference in ability among players does not matter any more. As a result, the predictive power goes to its "indifference" limit $\rho_1(\sigma^2 \to \infty) = 1/N$, also in agreement with the calculations.

Interestingly, for the Pareto distribution of abilities, ρ_1 exhibits nonmonotonicity as a function of the number of players N: when the number of players grows beyond a certain threshold, the strongest player (player 1) compensates her loss (on average) in the winning probability incurred by additional players by a gain from her ability "increasing" (on average) sufficiently due to the long tail of the Pareto density.

7 Results: seeding

In this Section I present and discuss the results for seeding scheme (7) introduced in Section 5. First, in order to illustrate numerically that for the "smooth" winning probabilities generated from the three distributions of players' abilities, seeding (7) is indeed the optimal seeding, I present the results for predictive power ρ_1 for *all* possible seedings. Of course, this can only be done for small N, such as 4 and 8. For $N = 4$, seeding (7) is known to be optimal for arbitrary w_{ij} satisfying conditions (6), therefore I concentrate here on the case of $N = 8$.

Figure 3 shows the predictive power ρ_1 for $N = 8$ players. The 315 non-trivially different seedings are ordered along the X-axes by predictive power. The (red) dashed lines show the predictive power calculated for seeding (7), ρ_1^{c}, which are also given in the panels. It is seen that seeding (7) indeed yields the highest predictive power for all σ^2.

It is seen from the figures that for small noise the variation in the predictive power is small. This result reflects the general fact that the uncertainty is low, and therefore seeding (or any other control parameter) is not really significant. The variation in the predictive power also becomes small in the large noise limit, because the players are effectively equalized by noise, and it matters less and less how they are seeded. The variation, however, is very significant in the intermediate noise regime, $\sigma^2 \sim 1$. In this regime seeding becomes an important control instrument in an organizer's hands.

The influence of seeding, as opposed to the no seeding case, is summarized in Figure 4, which presents the *difference* between seeding (7) and random seeding predictive powers, $\rho_1^{\mathrm{c}} - \rho_1^{\mathrm{rnd}}$, as a function of N and σ^2. The difference exhibits a universal behavior as a

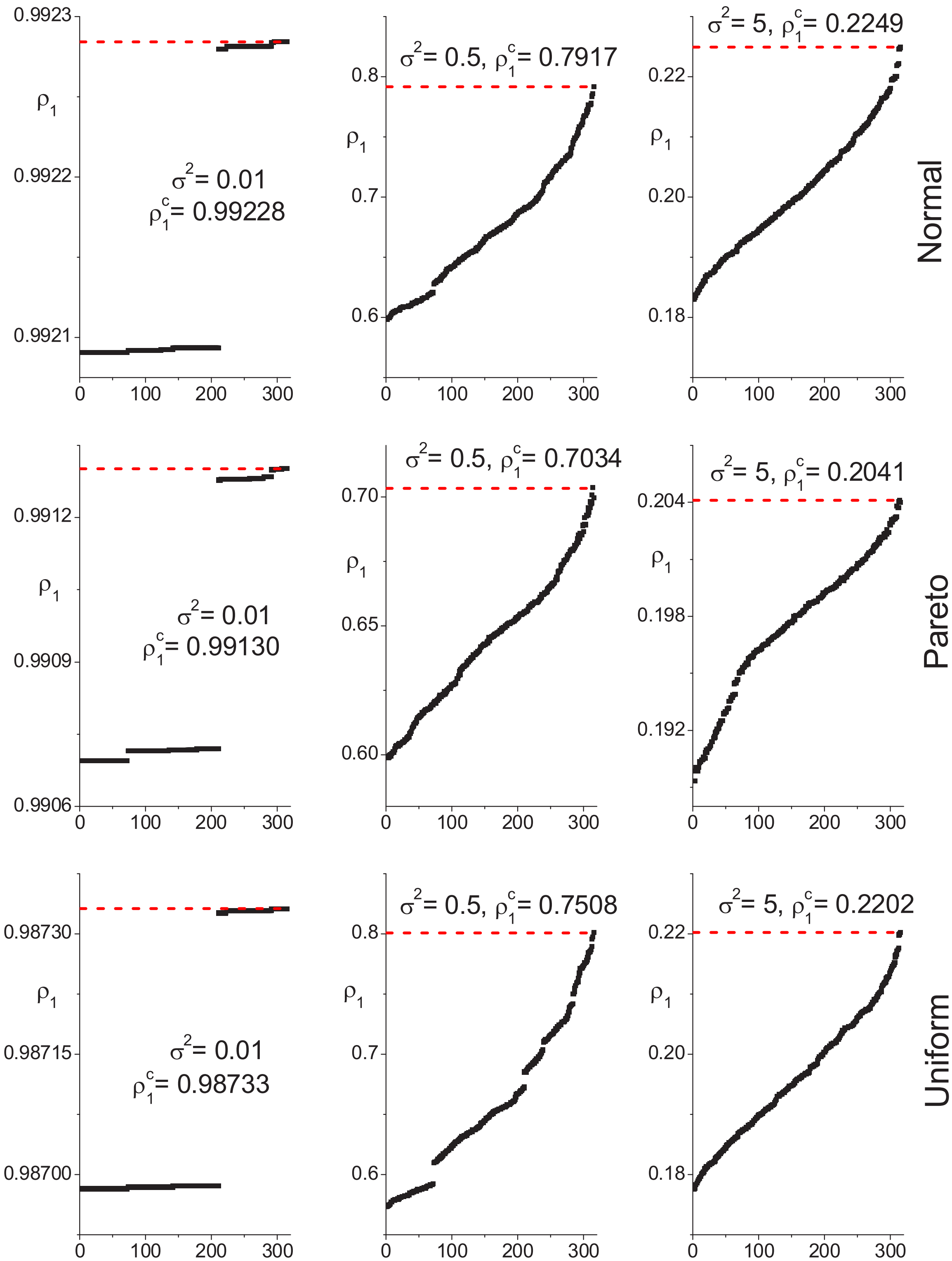

Figure 3: The predictive power ρ_1 for $N = 8$ players, three distributions of players' abilities, and various values of noise intensity σ^2. All 315 different seedings are ordered by their predictive power along the X-axes. The theoretical values of ρ_1^c for seeding $\{(1,8),(7,6),(5,4),(3,2)\}$ are shown by dashed (red) lines and given in the graphs.

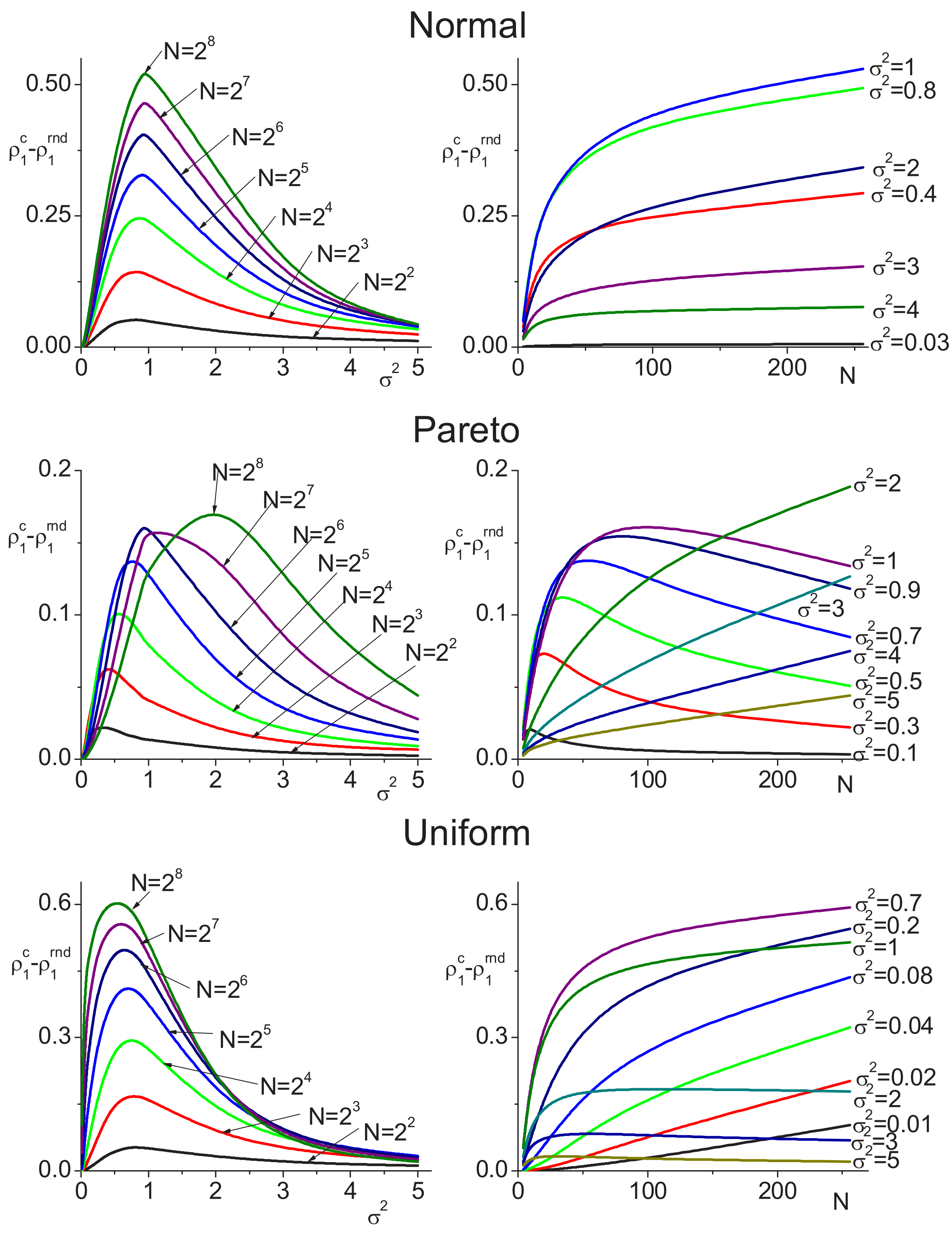

Figure 4: The difference in predictive power between seeding (7) and the average seeding, $\rho_1^{\rm c} - \rho_1^{\rm rnd}$, as a function of the number of players N and noise intensity σ^2 for three distributions of players' abilities. Note that $\rho_1^{\rm c} - \rho_1^{\rm rnd}$ exhibits a universal behavior as a function of σ^2: it starts from nearly zero at small noise intensities, rises to a maximum, and then drops. The reasons are explained in the text.

function of the noise intensity. It starts from zero in the small noise limit, and in the large noise limit it falls off gradually, depending on the number of players. In the intermediate noise regime the difference $\rho_1^{\mathrm{c}} - \rho_1^{\mathrm{rnd}}$ has a single well-defined maximum, which becomes more pronounced as N increases.

Note that seeding scheme (7) does not look appealing in terms of competitive balance since it favors the *ex ante* strongest player too much, and hence confers a disproportional burden on other players. Alternative seeding schemes may be used to induce more balanced results.

8 Conclusions

The present paper is a step towards understanding optimal tournament design. A new quantity, the predictive power, is defined, which is a measure of how well a particular tournament reveals the true ability ordering of players. It is argued that in many contexts maximizing the predictive power is an important real-world objective, whose discussion is missing in the literature. Here I consider a non-trivial tournament format, the elimination tournament; it is the second step (the first one being Ryvkin and Ortmann 2004 on round-robin tournaments) in a series of papers whose ultimate goal is to systematically address the optimal tournament format construction and selection problem. Independently, I argue that the distribution of ability in the population is a plausible model of heterogeneity of players (at least for the purposes at hand). The approach proposed here allows one to avoid choosing specific shapes of utility and cost functions. The uncertainty of outcomes is naturally modelled as noise that distorts players' performance, which, without noise, is assumed to be equivalent to ability.

The predictive power of an elimination tournament is calculated through a recurrence relation. It is shown that for the no seeding case it exhibits nonmonotonicity as a function of the number of players for the Pareto distribution of players' abilities. This has immediate implications, for example, for an organizer who seeks to minimize the predictive power, in terms of the optimal number of players.

The role of seeding in an elimination tournament is also explored. It is shown that for intermediate noise intensities, seeding plays a very significant role and is, therefore, a powerful instrument that allows the organizer to manipulate the outcome of a tournament. In contrast, in the small noise and large noise regimes, seeding plays practically no role. More generally, all design issues are only significant in the intermediate noise regime.

References

Baye, M.R., Hoppe, H.C., 2003. The strategic equivalence of rent-seeking, innovation, and patent-race games. Games and Economic Behavior, 44, 217-226.

Bognanno, M.L., 2001. Corporate tournaments. Journal of Labor Economics, 19, 290-315.

Ehrenberg, R.G., Bognanno, M.L., 1990. The incentive effects of tournaments revisited: evidence from the European PGA Tour. Industrial and Labor Relations Review, 43, 74S-88S; Do tournaments have incentive effects? The Journal of Political Economy, 98, 1307-1324.

Fullerton, R.L., McAfee, R.P., 1999. Auctioning entry into tournaments. The Journal of Political Economy, 107, 573-605.

Gradstein, M., Konrad, K., 1999. Orchestrating rent-seeking contests. The Economic Journal, 109, 536-545.

Horen, J., Reizman, R., 1985. Comparing draws for single elimination tournaments. Operations Research, 33, 249-262.

Knuth, D.E., 1987. Random knockout tournament: Problem 86-2. SIAM Review, 29, 127-129.

Lazear, E.P., 1999. Personnel economics: past lessons and future directions. Presidential address to the society of labor economists, San Francisco, May 1, 1998. Journal of Labor Economics, 17, iv+199-236.

Lazear, E.P., Rosen, S., 1981. Rank-order tournaments as optimal labor contracts. The Journal of Political Economy, 89, 841-864.

Moldovanu, B., Sela, A., 2002. Contest architecture. Discussion paper, University of Mannheim.

O'Flaherty, B., Siow, A., 1995. Up-or-out rules in the market for lawyers. Journal of Labor Economics, 13, 709-735.

O'Keeffe, M., Viscusi, W.K., Zeckhauser, R.J., 1984. Economic contests: Comparative reward schemes. Journal of Labor Economics, 2, 27-56.

Prendergast, C., 1999. The provision of incentives in firms. Journal of Economic Literature, 37, 7-63.

Reed, W.J., 2001. The Pareto, Zipf, and other power laws. Economics Letters, 74, 15-19.

Rosen, S., 1986. Prizes and incentives in elimination tournaments. The American Economic Review, 76, 701-715.

Roth, A., 2002. The economist as engineer: game theory, experimentation, and computation as tools for design economics. Econometrica, 70, 1341-1378.

Ryvkin, D., Ortmann, A., 2004. The predictive power of noisy round-robin tournaments. Working paper, CERGE-EI.

Szymanski, S., 2003. The economic design of sporting contests: a review. Journal of Economic Literature, 41, 1137-1187.

Taylor, C.R., 1995. Digging for golden carrots: an analysis of research tournaments. The American Economic Review, 85, 872-890.

Tullock, G., 1980. Efficient rent seeking. In *Buchanan, J., Toltison, R., Tullock, G., eds. Towards a Theory of Rent Seeking Society*, College Station: Texas A&M University Press, 97-112.

Individual researchers, as well as the on-line and printed version of the CERGE-EI Working Papers Series (including their dissemination) were supported from the following institutional grants:

- Economic Aspects of EU and EMU Entry [Ekonomické aspekty vstupu do Evropské unie a Evropské měnové unie], No. AVOZ70850503, (2005-2010);
- Economic Impact of European Integration on the Czech Republic [Ekonomické dopady evropské integrace na ČR], No. MSM0021620846, (2005-2011);

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.

Published by
Charles University in Prague, Center for Economic Research and Graduate Education (CERGE)
and
Economics Institute (EI), Academy of Sciences of the Czech Republic
CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.
Printed by CERGE-EI, Prague
Subscription: CERGE-EI homepage: http://www.cerge-ei.cz

Editors: Directors of CERGE and EI
Managing editors: Deputy Directors for Research of CERGE and EI

ISSN 1211-3298
ISBN 80-7343-045-2 (Univerzita Karlova v Praze, CERGE)
ISBN 80-7344-034-2 (Národohospodářský ústav AV ČR, Praha)

CERGE-EI
P.O.BOX 882
Politických vězňů 7
111 21 Praha 1
Czech Republic
http://www.cerge-ei.cz