Promotions and Incentives: The Case of Multistage Elimination Tournaments

Steffen Altmann, *IZA Bonn and University of Bonn*

Armin Falk, *University of Bonn, IZA, CEPR, and CESifo*

Matthias Wibral, *University of Bonn and IZA*

Promotions play an important role for the provision of incentives in firms. We analyze incentives in multistage elimination tournaments with controlled laboratory experiments. In our two main treatments, we compare a two-stage tournament to a one-stage tournament. Subjects in the two-stage treatment provide excess effort in the first stage, both with respect to Nash predictions and compared to the strategically equivalent one-stage tournament. Additional control treatments confirm that excess effort in early stages is a robust finding and suggest that above-equilibrium effort might be driven by limited degrees of forward-looking behavior and subjects deriving nonmonetary value from competing.

I. Introduction

Promotions play an important role for the provision of incentives in firms and other organizations. Hierarchical position constitutes an im-

We would like to thank the editor, Michael Waldman, whose thoughtful comments helped to improve the paper significantly. We also thank Johannes Abeler, Michael Bognanno, Florian Ederer, Ernst Fehr, Uri Gneezy, Sotiris Georganas, Lars Koch, Matthias Kräkel, Felix Marklein, Robert Oxoby, Bruce Weinberg, and Georg Weizsäcker for helpful discussions. Valuable comments were received from seminar and conference participants in Bonn, Barcelona, Budapest, Cologne, Malaga, Munich, New Orleans, Nottingham, Nuremberg, Raleigh, San Diego, Tübingen, and Warwick. Financial support from the Deutsche Forschungsgemein-

[*Journal of Labor Economics,* 2012, vol. 30, no. 1]

0734-306X/2012/3001-0004$10.00

portant factor of workers' pay even after controlling for differences in human capital and other worker characteristics (Baker, Gibbs, and Holmström 1994; Bognanno 2001). The prospect of being promoted to a better-paid job generates incentives to work hard even if current income is not tied to performance. In virtually all firms those who get promoted compete again for subsequent promotions. In many companies, there are up to a dozen hierarchical levels between the CEO and entry-level management (Belzil and Bognanno 2008; Lazear and Gibbs 2008), and workers advance through several different positions when careers are characterized by long-term employment relationships (Farber 1999). In a seminal contribution, Rosen (1986) has modeled the competition for promotion in such hierarchies as a multistage elimination tournament where in each stage fewer agents are selected for the next step of the career ladder. Incentives generated in such tournaments depend on two important components of the organizational structure: the immediate wage increase for an agent who gets promoted, and the option value of competing in further stages of the tournament and having the chance to earn even higher wages.

Although the importance of multistage elimination tournaments is undisputed, stringent empirical tests of their incentive effects are scarce. In this paper, we provide a step toward closing this gap with the help of controlled laboratory experiments. We study two main questions. First, we are interested in whether people take future promotion possibilities into account when exerting work effort on a given stage of a tournament. Second, we study whether they do so in a way that is consistent with tournament theory. In particular, we analyze whether multistage elimination tournaments and single-stage tournaments that are strategically equivalent are also equivalent in terms of contestants' behavior.

We address these questions with two main treatments. The first is a two-stage tournament (TS) in which four participants compete for being promoted to the second stage. Promotion depends on subjects' output, which is a function of costly effort and an individual noise term. The two subjects with the lowest output levels in the first stage are eliminated from further competition and receive a low wage. The two subjects with the highest output levels in the first stage are promoted, that is, they are allowed to take part in the second stage where they compete against each other again. The subject with the highest second-stage output receives a high wage, whereas the other finalist is paid an intermediate wage. Pa-

schaft through SFB/TR 15, the European Research Council (Starting Grant), and the Institute for the Study of Labor (IZA) is gratefully acknowledged. Part of this research was conducted while two of the authors were visiting the Institute for Empirical Research in Economics at the University of Zurich. The authors thank the members of the institute and especially Christian Ewerhart for their hospitality, encouraging discussions, and financial support. Contact the corresponding author, Steffen Altmann, at altmann@iza.org.

rameters in this treatment are chosen to make the tournament incentive maintaining in the sense that equilibrium effort is identical in both stages (Rosen 1986).

We compare this treatment to a one-stage tournament (OS) in which four subjects compete once for two top positions. In OS, wages for promoted subjects are chosen such that the tournament is strategically equivalent to the first stage of TS. This means that promoted subjects in OS earn the sum of the intermediate wage in TS and the monetary equivalent of the second-stage option value in that tournament, implying that equilibrium effort is the same in both treatments. Comparing OS and TS thus allows testing whether strategic equivalence translates into behavioral equivalence.

Our findings can be summarized as follows: first, average effort in our one-stage treatment is remarkably close to the predictions of tournament theory. This parallels findings of previous experiments on symmetric one-stage tournaments (e.g., Bull, Schotter, and Weigelt 1987; Orrison, Schotter, and Weigelt 2004). Second, behavior in the TS treatment indicates that subjects take the option value of future promotion possibilities into account when deciding on their work effort in multistage tournaments. However, we also observe important departures from theoretical predictions in the TS treatment. Behavior in the first stage of TS differs strongly both from the one-stage treatment and from theoretical predictions. In particular, subjects exert significantly higher efforts in the first stage of the two-stage tournament. Finally, the TS tournament does provide appropriate performance incentives in the second stage: although efforts in the second stage are lower than in the first stage, they do not fall below theoretical predictions.

Most promotion tournaments that we observe in firms and other hierarchical organizations have multiple stages. Our results indicate that the mechanisms of incentive provision in multistage tournaments largely operate as suggested by theory. People do not only respond to differences in prizes, or wages, but are also motivated by the option value generated by future promotion possibilities. However, going from a one-stage tournament to multistage setups seems to make a fundamental difference, as people tend to exert excess effort in early stages of the tournament. This conclusion is corroborated by evidence from a third treatment in which participants compete in a tournament with three stages (THS). Paralleling our results of the two-stage treatment, we observe that participants exert excess effort both in the first and the prefinal stage of the tournament, but that behavior does not differ from equilibrium predictions in the final stage of the tournament. Taken together, our results show that one cannot simply generalize from simple one-stage tournaments to more complex organizational structures.

Our findings also contribute to the growing literature that studies whether economic agents exhibit forward-looking behavior in dynamic setups, more generally (Ameriks, Caplin, and Leahy 2003; Houser, Keane, and McCabe 2004; Chevalier and Goolsbee 2009). For instance, Chevalier and Goolsbee (2009) analyze whether individuals account for future resale opportunities when purchasing durable goods. They show that customers of college textbooks take into account changes in the option value of the books (i.e., changes in the expected resale prices due to new book editions), indicating forward-looking behavior. In our context, excess effort in early stages of the multistage tournaments could be driven by contestants making mistakes and systematically overestimating the option value of later stages. To examine this possibility more directly, we conducted another control treatment (OSL). This treatment is equivalent to our two-stage tournament with the exception that equilibrium behavior in the second stage of the tournament is exogenously induced. We find that excess effort in OSL is less pronounced than in the first stage of our two-stage tournament, with efforts in OSL falling in between those of OS and TS. This suggests that the capacity to correctly perceive option values contributes to excess effort in early stages, but that other factors such as subjects attaching nonmonetary value to achieving certain hierarchical positions or competing in a tournament might be important as well.

Several rationales have been discussed in the literature for why firms employ promotion-based incentives (Gibbons and Waldman 1999). Promotions might be a means to improve the match quality between employees' abilities and job requirements, they can help filter out common productivity shocks, and they are less susceptible to influence activities than other incentive schemes (Lazear and Rosen 1981; Baker, Jensen, and Murphy 1988; Fairburn and Malcomson 2001). In addition, promotions provide a signal on workers' abilities and can therefore play an important role in muting informational asymmetries between the current and other prospective employers (Waldman 1984). This in turn generates additional incentives, for example, for acquiring human capital (Zábojník and Bernhardt 2001).

Our paper complements previous studies that have used field data from executive compensation, sports, or agricultural production to evaluate predictions of these theories (e.g., Ehrenberg and Bognanno 1990; Knoeber and Thurman 1994; Eriksson 1999; Bognanno 2001). The main advantage of an experimental approach is that it allows controlled and exogenous variation of incentives. We can therefore provide causal evidence on the pure incentive effects of different tournament schemes, controlling for the influence of unobservable variables that might affect effort provision and actual promotion decisions in the field, like agents' soft skills or supervisor favoritism (Prendergast and Topel 1996). To address our research questions with observational data, one would need to observe individual measures of compensation and work effort across several hierarchical layers, as well as other factors that potentially affect individuals' perceptions and reactions to option values in dynamic promotion tournaments. A disadvantage of our laboratory approach is that it relies on a relatively homogeneous sample of contestants. For instance, participants in our tournaments might be relatively well trained in solving complex dynamic problems, and they are also likely to differ little in personal characteristics such as competitive inclination, risk attitudes, or cognitive abilities. Moreover, we have to abstract from some features of promotion tournaments in actual firms for the sake of simplicity. For example, there is a natural limit to the size and the time horizon of tournaments in the laboratory.[1]

The remainder of the paper is organized as follows: the next section presents a simple model of multistage elimination tournaments on which our experiment is based. Section III discusses our experimental design and derives hypotheses. Section IV presents and discusses the results of our two main treatments and the control treatments. Section V concludes.

II. A Simple Model of Multistage Elimination Tournaments

We consider a simple elimination tournament in which four identical agents compete for promotion.[2] The promotion decision depends on relative output produced by the agents. Competition consists of two stages: in the first stage, all four agents compete against each other. The two agents with the lowest output levels in that stage receive a wage w_{low} and are eliminated from further competition. The two agents with the highest output levels in the first stage are promoted. That is, they are allowed to take part in the second stage (or "final") where they compete against each other again. The agent who produces more output in the second stage receives a wage w_{high}, whereas the other finalist gets an intermediate wage w_{med}. Note that the decision of who receives w_{high} or w_{med} does not depend on the first-stage output of the finalists.

This two-stage elimination tournament can be modeled as follows. In the first stage of the tournament four agents, $i = 1, 2, 3, 4$, compete against each other. Agents who participate in stage $k \in \{1, 2\}$ individually produce output $y_{i,k}$ according to the production function

$$y_{i,k} = e_{i,k} + \varepsilon_{i,k},$$

where $e_{i,k}$ denotes the effort level that agent i exerts in stage k, and $\varepsilon_{i,k}$ is a random shock faced by agent i in stage k. Shocks are assumed to be drawn independently for each agent in each stage. For simplicity,

[1] See also Falk and Heckman (2009) for a more detailed discussion of advantages and limitations of laboratory experiments.

[2] Most of the assumptions below follow the classic (one-stage) tournament model introduced by Lazear and Rosen (1981).

we assume that $\varepsilon_{i,k}$ is uniformly distributed on the interval $[-q, q]$.[3] Agent i's output in stage k does not depend on previous effort or output and the production technology is identical for all agents in all stages. Agents bear the cost of effort exertion. The cost function is given as follows:

$$C(e_{i,1}, e_{i,2}) = \frac{e_{i,1}^2}{c} + \frac{e_{i,2}^2}{c}.$$

Note that this specification implies separability of costs across stages, that is, in line with Rosen (1986) there is no carryover of costs between stages. Furthermore, we assume that agents are identical and risk neutral with utility functions which are additively separable in wages and effort costs:

$$U_i(w, e_{i,1}, e_{i,2}) = w - C(e_{i,1}, e_{i,2}).$$

For the derivation of equilibrium predictions we restrict our attention to the set of symmetric subgame perfect Nash equilibria. The two-stage tournament can be solved by backward induction. Because (i) the decision who wins the second stage solely depends on the output of the finalists in this stage, (ii) there is no cost carryover between stages, and (iii) the random terms are independently distributed both across stages and agents, the final of our two-stage tournament is equivalent to a one-stage tournament in which two participants compete for a promotion. Given that agent i has reached the second stage where two agents compete for one winner prize $w_{\mathbf{high}}$ and one loser prize $w_{\mathbf{med}}$, he chooses stage 2 effort $e_{i,2}$ in order to maximize an expected utility function of the following form:

$$EU_i(w_{\text{high}}, w_{\text{med}}, e_{i,2}, e_{j,2}) =$$

$$\pi(y_{i,2} > y_{j,2})w_{\text{high}} + [1 - \pi(y_{i,2} > y_{j,2})]w_{\text{med}} - C(e_{i,1}, e_{i,2}).$$

The term $\pi(y_{i,2} > y_{j,2})$ denotes the probability that i's output in stage 2 is greater than the output of agent j. With our assumptions regarding the production function and random terms, this expression can be rewritten as follows:

$$EU_i(w_{\text{high}}, w_{\text{med}}, e_{i,2}, e_{j,2}) =$$

$$F_{\varepsilon_{j,2}-\varepsilon_{i,2}}[e_{i,2} - e_{j,2}](w_{\text{high}} - w_{\text{med}}) + w_{\text{med}} - C(e_{i,1}, e_{i,2}),$$

[3] Virtually all tournament experiments use the uniform distribution, primarily because its concept is easy to understand for experimental subjects. The predictions of the model, however, can be generalized to other distributions of shocks (see Lazear and Rosen 1981).

where $F_{\varepsilon_{j,2}-\varepsilon_{i,2}}[\cdot]$ denotes the cumulative distribution function of the difference between random terms $\varepsilon_{j,2}$, $\varepsilon_{i,2}$. Maximizing $EU_i(\cdot)$ over $e_{i,2}$ yields the following first-order condition:

$$f_{\varepsilon_{j,2}-\varepsilon_{i,2}}(e_{i,2} - e_{j,2})(w_{\text{high}} - w_{\text{med}}) = \frac{\partial C(\cdot)}{\partial e_{i,2}}.$$

Assuming symmetry yields $f_{\varepsilon_{j,k}-\varepsilon_{i,k}}(0) = 1/(2q)$ for $\varepsilon_{i,k}, \varepsilon_{j,k} \sim U[-q, q]$. The symmetric subgame perfect Nash equilibrium of the two-stage tournament thus entails the following second-stage effort level $e_{i,2}^{*,\text{TS}}$ for the two agents who participate in the final:

$$e_{i,2}^{*,\text{TS}} = \frac{(w_{\text{high}} - w_{\text{med}})c}{4q}.$$

Given that both finalists play this equilibrium, the expected utility gain in the final, that is, the continuation value for an agent in the first stage is given as follows:

$$EV_{i,2} = w_{\text{med}} + \frac{1}{2}[w_{\text{high}} - w_{\text{med}}] - \frac{(e_{i,2}^{*,\text{TS}})^2}{c}.$$

An agent who reaches the final earns a wage of w_{med} for sure. By exerting stage 2 effort $e_{i,2}^{*,\text{TS}}$, he has the chance to receive the higher wage w_{high} instead. In the symmetric equilibrium, this occurs with probability 1/2. Moreover, he has to pay the cost of effort exertion in the second stage.

The tournament's first stage can now (in expected values) also be modeled as a one-stage tournament between four agents with two winner prizes $EV_{i,2}$ and two loser prizes w_{low} (see Rosen 1986). The derivation of equilibrium effort for such a tournament follows the same steps as above. Alternatively we can apply a result from Orrison et al. (2004), who show that equilibria of fully symmetric one-stage tournaments are not affected by "organizational replication" for our specification of the production function, cost function, and random terms. This implies that an equilibrium in a tournament with two identical participants and one winner prize is also an equilibrium in a tournament with four identical participants and two winner prizes.

As a shortcut we can therefore use the solution for $e_{i,2}^{*,\text{TS}}$ and simply replace w_{high} and w_{med} with $EV_{i,2}$ and w_{low} to obtain the equilibrium effort level for the first stage:

$$e_{i,1}^{*,\text{TS}} = \frac{(EV_{i,2} - w_{\text{low}})c}{4q} =$$

$$\frac{(w_{\text{med}} - w_{\text{low}} + (1/2)[w_{\text{high}} - w_{\text{med}}] - [(e_{i,2}^{*,\text{TS}})^2/c])c}{4q}.$$

This expression illustrates the two components of incentive provision in multistage tournaments. By winning the first stage and qualifying for the final, an agent receives an immediate wage gain ($w_{\mathbf{med}} - w_{\mathbf{low}}$) but additionally has the option to compete in the final and win the top prize $w_{\mathbf{high}}$. The value of this option is

$$\frac{1}{2}[w_{\text{high}} - w_{\text{med}}] - \frac{(e_{i,2}^{*,\text{TS}})^2}{c}.$$

Several aspects of the model deserve emphasis. First, our design closely follows the original model of elimination tournaments by Rosen (1986) with one notable exception: instead of having two semifinals with two participants each, who compete for one slot in the final, we analyze a setup with four participants competing for two slots in the final. While both variants are theoretically equivalent for symmetric agents, we employ the latter because it allows us to design a one-stage tournament that is procedurally as close as possible to our two-stage tournament. In particular—as will become clear in the next section—both tournaments have the same number of participants and subjects compete for the same number of promotions.

Note also that we concentrate on fully symmetric tournaments, that is, all contestants have identical productivity, effort cost functions, distribution of shocks, and so on. We therefore do not study whether the most able agents are selected for promotion. We deliberately abstract from this important purpose of tournament-based incentive schemes in order to isolate the pure incentive effect of elimination tournaments. Selection on variables other than ability may still occur in our experiment, for instance, if contestants with different personality characteristics perceive incentives in a given tournament differently and therefore provide different levels of effort. We analyze this possibility in more detail in Section IV.F.

III. Design of the Experiment

Our experiment consists of two main treatments and two additional control treatments that allow us to study behavior in multistage tournaments from different angles. For all treatments our benchmark is the prediction of the symmetric subgame perfect Nash equilibrium. In this section, we describe theoretical hypotheses and experimental procedures for our main treatments in detail. Detailed information on our control treatments will be provided in Sections IV.D and IV.E, respectively.

A. Treatments and Hypotheses

Our first treatment TS is a two-stage elimination tournament with four participants competing for $w_{\mathbf{high}}$, $w_{\mathbf{med}}$, and $w_{\mathbf{low}}$ as discussed in the previous section. We compare this treatment to a one-stage tournament (OS) in which four subjects compete for two top positions. The two subjects with the highest output levels receive a wage $w_{\text{med}}^{\text{OS}}$ in OS while the two losers of the competition receive a wage w^{OS}_{low}. The OS treatment fulfills several purposes. First, it serves as a validity check for our results given that a number of studies on one-stage tournaments already exist. In particular, our parametrization of this treatment is very close to a treatment from Orrison et al. (2004).

More importantly, however, the OS treatment allows us to investigate whether one-stage tournaments are behaviorally different from multistage setups. To investigate this question we design OS such that it is strategically equivalent to the first stage of the two-stage tournament TS. As discussed in the previous section, the first stage of a two-stage tournament can be interpreted as a one-stage tournament in which agents compete for the expected value of participating in the second stage. Strategic equivalence between OS and the first stage of TS is thus achieved by keeping the low wage constant ($w^{OS}_{low} = w_{low}$) and choosing

$$w^{OS}_{med} = EV_{i,2} = w_{med} + \frac{1}{2}[w_{high} - w_{med}] - \frac{(e^{*,TS}_{i,2})^2}{c}.$$

In other words, the wage for the promoted agents in the one-stage tournament (w^{OS}_{med}) is equivalent to the wage w_{med} from TS plus the option value of participating in the final of TS. This choice implies that equilibrium effort levels in the OS treatment and in the first stage of the TS treatment are the same. We can therefore test the *behavioral equivalence hypothesis*:

$$e^{OS}_{i,1} = e^{TS}_{i,1}.$$

In addition, we analyze behavior across stages in our two-stage tournament. In the TS treatment wages are chosen such that equilibrium efforts in the first and second stage are equal. Remember that the two elements of incentives in multistage tournaments are the wage spread and the option value of competing for further promotions. In the final stage, the option value is zero because there are no further promotions beyond that stage. To make the tournament in the TS treatment incentive maintaining in the sense of Rosen (1986), this decrease in the option value is offset by an appropriate increase in the wage spread in the second stage:

$$w_{high} - w_{med} = w_{med} - w_{low} + \frac{1}{2}[w_{high} - w_{med}] - \frac{(e^{*,TS}_{i,2})^2}{c}.$$

Comparing behavior across stages thus allows to test the *incentive maintenance hypothesis*:

$$e^{TS}_{i,1} = e^{TS}_{i,2}.$$

Experimental parameters and the resulting equilibrium predictions for our two main treatments are summarized in columns 1 and 2 of table 1.[4]

[4] The remaining treatments reported in table 1 will be discussed in more detail in Secs. IV.D and IV.E.

Table 1
Parameters and Equilibrium Predictions for All Treatments

Treatment	OS	TS	THS	OSL
c	2,250	2,250	2,250	2,250
q	60	60	60	60
w_{low}	5.73	5.73	5.73	5.73
w_{med}	13.62	12.11	12.11	17.57/9.68
w_{high}	...	20	18.49	...
w_{top}	...	...	26.38	...
$e^*_{i,1}$	74	74	74	74
$e^*_{i,2}$	...	74	74	...
$e^*_{i,3}$	...	...	74	...

When deciding on their efforts, subjects could choose any integer $e_{i,k} \in \{0, 1, \ldots, 125\}$. The effort costs in each stage are given by

$$C(e_{i,k}) = \frac{e^2_{i,k}}{2{,}250}$$

and $\varepsilon_{i,k} \sim U[-60, 60]$. The parameters chosen imply equilibrium efforts of 74 in both stages of TS and in OS.[5]

B. Experimental Procedures

The experiment was conducted at the BonnEconLab of the University of Bonn. A total of 320 subjects, 124 in our two main treatments, participated in the experiment.[6] We employed a one-shot between-subjects design, that is, subjects participated in only one of the treatments. All participants were students from various fields of study. The experiment was programmed and conducted with the software z-Tree (Fischbacher 2007); subjects were recruited using the online recruitment system ORSEE (Greiner 2003).

Upon arrival in the laboratory subjects were seated in separate cubicles. Before the tournament started, they received detailed written instructions on the respective treatment they took part in. These were neutrally framed and did not contain potentially value-laden terms like "tournament," "final," "winner," and so on. Subjects were also provided with a cost table that listed the costs for all feasible effort levels. After reading the instructions subjects completed several control questions. The experiment started only after all participants had answered all control questions correctly. During the tournament, subjects simultaneously entered their effort de-

[5] Note that, while in equilibrium all players make positive profits, the range of feasible efforts and the specification of the cost function imply that, in principle, subjects could make losses. However, this only occurred once in our main treatments.

[6] In one session of the OS treatment, we had to conduct one tournament less since the number of participants was lower than expected.

cisions and were then asked to state their expectations about the other participants' efforts on the next screen. This question was not announced beforehand. After the first stage, participants in TS were only informed about the realization of their own random draw and about whether they had been promoted to the second stage. The finalists then again made an effort choice and entered their expectation about the opponent's effort. At the end of the tournament subjects were informed about their earnings and were asked to fill in a questionnaire. The questionnaire included sociodemographic variables, free-form questions about decision making in the experiment, a survey on competitive inclination (Smither and Houston 1992), a survey measuring domain-specific risk attitudes (Dohmen, Falk, Huffman, et al. 2011), a test for cognitive reflection (Frederick 2005), and a German version of the Barratt Impulsiveness Scale (Patton, Stanford, and Barratt 1995; Preuss et al. 2008).

In addition to subjects' decisions in the tournament, we elicited a measure of their risk attitudes using an incentive-compatible lottery procedure adapted from Dohmen, Falk, Huffman, et al. (2011). The procedure involved 15 choices between a binary lottery and a safe option. The lottery was kept constant between choices—subjects always had a 50/50 chance of winning €4 or €0—while the safe option increased from €0.25 to €3.75 in steps of 25 cents. The first decision where a subject chose the safe option thus gives us an upper bound of her certainty equivalent for the lottery at hand.[7] For each subject, one randomly determined choice was paid out at the end of the experiment. Finally, for a subset of participants we elicited an incentivized measure of overconfidence and an incentivized measure of forward-looking behavior. Our measure of overconfidence is based on how subjects judge their performance in a trivia quiz relative to their actual performance. As our measure of forward-looking behavior we used the "Hit 15 Points" game (see Sec. IV.F) introduced by Burks et al. (2008).[8]

The structure of the experimental session ensured that subjects' decisions in the tournament are independent observations. The whole experimental session lasted on average 100 minutes. In our two main treatments, subjects earned an average of €18.75 (€1 ≈ US\$1.25 at the time of experiment), including a showup fee of €4 and a fixed payment of €3 for completing the postexperimental questionnaire. In the tournament, earnings ranged between −€0.15 and €18.04.

[7] For 10 subjects in our main treatments, we could not calculate the certainty equivalent since these subjects switched between the lottery and the safe option more than once. Results reported below are robust to including these subjects either with their minimum or maximum switching point.

[8] A translation of the experimental instructions for the TS treatment can be found in the online appendix. The cognitive reflection test and the Barratt Impulsiveness Scale were only administered to a subset of participants.

Table 2
Summary of Behavior in All Treatments

	Treatment and Stage						
	OS	TS		THS			OSL
	$k = 1$ (1)	$k = 1$ (2)	$k = 2$ (3)	$k = 1$ (4)	$k = 2$ (5)	$k = 3$ (6)	$k = 1$ (7)
Average effort	71.7	84.8	76.7	79.2	83.1	74.6	77.3
Median effort	75	84	79.5	78.5	81	77	78
min(effort)	0	40	14	40	59	8	1
max(effort)	125	125	125	117	125	116	125
SD	29.1	19.7	26.5	16.2	17.3	23.9	25.8
N	60	64	32	64	32	16	100
e^*	74	74	74	74	74	74	74

Note that our experimental procedures differ from some of the existing tournament experiments in that we implement a one-shot interaction structure. Previous experiments have typically used repeated interactions. The advantage of the latter is that it allows for learning, which is potentially important given the strategic decision environment in tournaments. A potential downside, however, is that repeated game structures question the validity of static equilibrium predictions. Since we are explicitly interested in testing theoretical predictions, we decided to use a one-shot design. This has the additional advantage that stakes in the one-shot interaction are relatively high.

IV. Results

In this section we first test whether the results for the one-stage tournament replicate earlier findings from similar tournaments. We then study the dynamic aspect of multistage tournaments by comparing the one-stage tournament OS to the strategically equivalent first stage of the two-stage tournament TS. In a third step, we address the question whether the TS treatment is incentive maintaining by analyzing behavior across stages of the tournament. We then report findings from two additional control treatments. The first asks whether our results for the two-stage tournament extend to elimination tournaments with more than two stages. In the second, we analyze a one-stage tournament with a more refined hierarchy. Finally, we explore how personal characteristics are related to effort provision.

A. Behavior in the One-Stage Tournament

Table 2 summarizes effort decisions in the OS treatment (col. 1). Two points are worth noting: first, efforts are on average very close to the theoretical predictions. While the average effort of 71.7 is slightly below the Nash prediction of 74, median effort is marginally above this level.

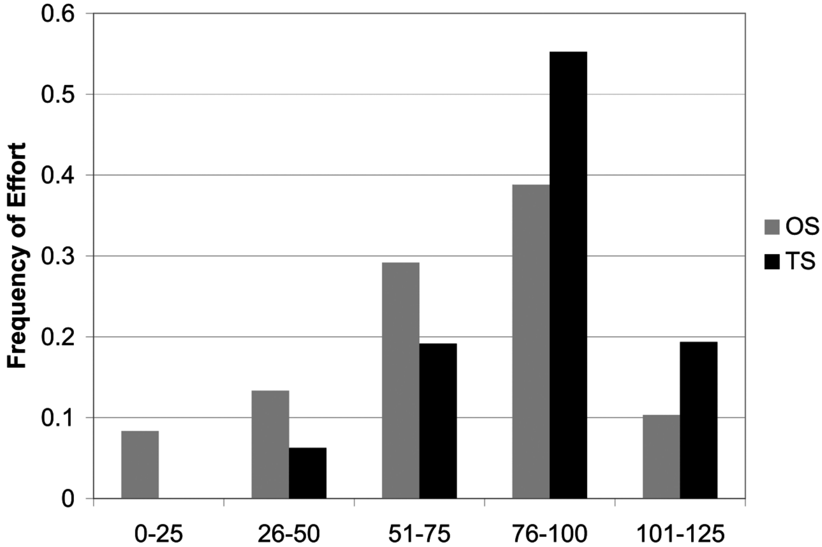

Fig. 1.—Frequency of effort choices in the OS treatment and the first stage of the TS treatment.

Second, there is substantial heterogeneity in subjects' behavior (see fig. 1).

Both observations are in line with previous findings from symmetric one-stage tournaments (e.g., Bull et al. 1987; Eriksson, Teyssier, and Villeval 2009). In particular, our results replicate those of Orrison et al. (2004), who observe an average effort of 73.3 for an almost identical tournament which was repeated 20 times using lower stakes in a given round. The similarity of our results to those of Orrison et al. (2004) show that one of the most important findings in the experimental literature on symmetric promotion tournaments—average effort being close to Nash predictions—seems to be robust with respect to using one-shot versus repeated interactions and with respect to increased stake sizes.

Result 1. Average behavior in the one-stage tournament is close to the predictions of the symmetric Nash equilibrium.

B. Testing Behavioral Equivalence

Our one-stage tournament and the first stage of the two-stage tournament TS are strategically equivalent in the sense that the wage w^{OS}_{med} in the one-stage tournament includes the equilibrium option value of participating in the second stage of the two-stage tournament. A comparison of $e^{OS}_{i,1}$ and $e^{TS}_{i,1}$ therefore serves as a test of how subjects in the two-stage tournament perceive this option value. Columns 1 and 2 of table 2 show that first-stage behavior differs considerably between the two treatments (see also fig. 1). Average effort in the first stage of the TS treatment is 84.8, with a median of 84. Thus, subjects behave much more competitively in the multistage tournament, exerting efforts that are almost 20% higher than those of their counterparts in the OS treatment. A Mann-Whitney U-test confirms that the treatment difference is statistically significant ($|z| = 2.536$, $p = .011$, two sided).

Comparing effort levels in the two treatments to the theoretical predictions derived in Section III indicates that it is excess effort in TS rather than "too low" effort in OS that drives the treatment difference. The null hypothesis that efforts are equal to Nash predictions is rejected in TS (t-test, $p < .001$) but cannot be rejected in OS ($p = .534$).

Note that the derivation of equilibrium predictions is based on the assumption of risk-neutral agents. We check the validity of this assumption with our incentivized measure of subjects' risk attitudes. It turns out that our experimental subject pool is close to risk neutrality. The median subject in both treatments is risk neutral, and the certainty equivalent of more than 50% of subjects lies in a range of ±€0.25 around the risk-neutral certainty equivalent. We investigate the relationship between risk and effort provision in more detail in Section IV.F.

Effort choices in the TS treatment suggest that subjects are not naive in the sense that they ignore the second stage. Quite the contrary, the two-stage elimination tournament seems to trigger especially competitive behavior in the first stage. The difference in behavior between OS and TS is not just driven by some subjects choosing extreme effort levels in the TS treatment. A closer look at the distributions of first-stage efforts in fig. 1 reveals instead that the whole effort distribution is shifted to the right in the TS treatment. As a consequence, efforts are less dispersed (Levene's test, $W = 5.79, p = .018$). The effort distribution also illustrates that exerting excess effort is quite widespread in TS: 77% of subjects choose efforts above the equilibrium effort level of 74. This compares to only 52% in the OS treatment. The strong difference between treatments is also reflected at the lower tail of the distribution. While the lowest effort level exerted by a subject in the TS treatment is 40, 20% of subjects in the OS treatment choose efforts at or below this level.[9]

[9] To examine whether excess effort in the first stage of TS is a result of the specific wage structure in this treatment, we conducted a further experiment where participants ($N = 32$) faced a two-stage tournament with a lower intermediate wage, $w_{\text{med}}^{\text{TSC}} = €9.33$. This more convex wage structure in the "two-stage convex" (TSC) treatment implies that the equilibrium effort level in the first stage is 42 instead of 74 in the TS treatment. Subjects' behavior in the TSC treatment underlines that excess effort obtains although incentives to provide effort in the first stage are weaker in this treatment. On average, participants in TSC choose an

Result 2. Efforts in the first stage of TS are significantly higher than in the OS treatment. This difference is driven by excess effort in TS.

C. Testing Incentive Maintenance

In the next step, we turn to behavior in the second stage of the TS treatment. Remember that parameters were chosen such that equilibrium effort levels are the same in both stages of TS. We know already that efforts are above the equilibrium prediction in the first stage. In this sense, we can reject the hypothesis of *equilibrium* effort choices in both stages of TS. It remains to be seen, however, whether effort levels are the same in both stages of the tournament, or whether they differ. In particular, does the two-stage character of TS induce above-equilibrium effort also in the second stage, or do players reduce efforts relative to their first-stage performance?

On the aggregate level, effort decreases when comparing behavior across stages in the TS treatment. Average effort in the first stage is 84.8; in the final it goes down to 76.7. In fact, we cannot reject the null hypothesis that efforts in the second stage of the TS treatment are equal to the equilibrium prediction of 74 (t-test, $p = .565$). Note that the average effort in the first stage includes the efforts of those who did not make it to the second stage. Since—by design of the promotion competition—the latter usually exerted lower effort, the decrease from stage 1 to stage 2 is even stronger if we consider only finalists' behavior. Their mean effort in the first stage is 93.1, implying that on average finalists decrease their effort by about 16 points (see also fig. 2). A Wilcoxon signed rank test confirms that the decrease in finalists' effort is statistically significant ($|z| = 2.315, p = .021$).

It is also informative to compare behavior in the second stage of TS to behavior in the OS treatment. After all, conditional on having reached the second stage, the finalists in TS are taking part in a one-stage tournament with exactly the same wage spread as in the OS treatment. Interestingly, the observed behavior in the second stage of TS is not significantly different from behavior in the OS treatment (Mann-Whitney U-test, $|z| = .771, p = .441$).

In sum, comparing effort choices across the two stages of the TS treatment indicates that—relative to the excess effort observed in the first stage—subjects reduce their performance in the second stage. However, average effort in the final stage does not fall below the Nash equilibrium of this stage. In this sense, the two-stage tournament does maintain incentives to provide effort in the second stage.

---

effort of 82.4 points in the first stage (median = 83, SD = 24.6). A t-test confirms that effort choices in TSC are significantly above the equilibrium prediction of 42 ($p < .001$).

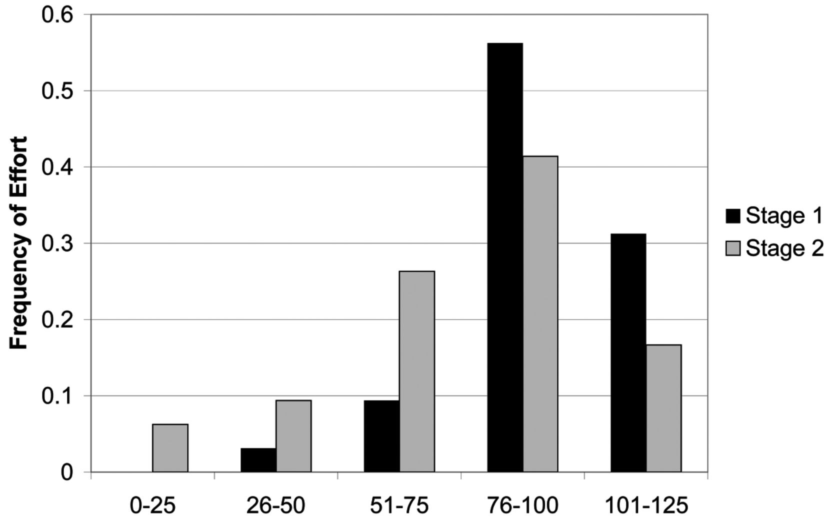

FIG. 2.—Frequency of effort choices in the TS treatment: first-stage versus second-stage efforts of finalists in the TS treatment.

Result 3. In the second stage of TS, efforts are lower than in the first stage of this treatment. Average behavior in the second stage is close to the predictions of the symmetric Nash equilibrium.

D. Going beyond Two Stages

The findings in our two main treatments, and in particular the excess effort observed in the first stage of the TS treatment, raise interesting questions. Is excess effort specific to the first stage of multistage tournaments or to the prefinal stage, that is, the stage where contestants compete for entering the final? Or do multistage tournaments provoke excess effort more generally, except for the final stage? These alternatives can only be distinguished in tournaments with more than two stages. Adding a further stage could also lead contestants to smooth efforts across stages, thereby muting the tendency to exert excess effort early in the tournament.

To examine these possibilities more closely, we implement an additional treatment where participants compete in an elimination tournament with three stages (THS treatment). In the first stage of THS, eight contestants compete to be among the four participants in a second stage. The four participants in the second stage in turn compete for entering the third stage. Finally, the two participants with the highest second-stage outputs compete again in a third stage. Parameters are chosen such that the THS treatment shares all essential features of our two-stage tournament. Output and cost functions are identical to those in the TS treatment. Most importantly, wages in THS are chosen such that the tournament is also incentive maintaining in theory, with an equilibrium effort level of 74 in all three stages (see col. 3 of table 1). The THS treatment therefore embeds all relevant design aspects of TS in a richer strategic context. It allows us to assess whether excess effort is specific to a particular stage of a multistage tournament or whether above-equilibrium effort in early stages of elimination tournaments is a more general phenomenon.

A total of 64 subjects participated in the THS treatment. Columns 4–6 of table 2 summarize behavior in all three stages of the treatment. Most importantly, we observe above-equilibrium effort levels both in the first and second stage of the THS treatment. In the first stage, subjects on average choose an effort of 79.2 points; in the second stage, the corresponding value is 83.1. In the final stage of THS, effort is again relatively close to the equilibrium, with an average value of 74.6. The null hypothesis that efforts are equal to Nash predictions is rejected both for the first (t-test, $p = .013$) and second stage ($p = .006$) of the THS treatment but cannot be rejected for the final stage ($p = .918$).

Comparing behavior in THS to the TS treatment indicates that efforts in the last two stages of the three-stage tournament are very similar to the corresponding values in our two-stage treatment. In the final stage, the value of 74.6 for the THS treatment compares to 76.7 in TS; in the prefinal stage, efforts are 83.1 in THS and 84.8 in TS. In both cases, the difference in behavior is not significant (Mann-Whitney U-tests, $|z| = .37, p = .710$, and $|z| = .81, p = .419$, respectively). Turning to the first stage of THS, we find that the tendency to provide excess effort in this stage seems to be somewhat less pronounced than in the two-stage tournament. In fact, while still above equilibrium, the average effort of 79.2 in THS lies below the first-stage effort of 84.8 in TS (Mann-Whitney U-test, $|z| = 1.85$, $p = .065$). Overall, our findings from the three-stage tournament suggest that above-equilibrium effort provision seems to be generally relevant in early stages of multistage tournaments, but that the tendency to provide excess effort is particularly pronounced in the prefinal stage of a tournament.

Result 4. The three-stage treatment THS replicates the findings from the two-stage tournament in a richer context: subjects exert above-equilibrium effort both in the first and second stage of the three-stage tournament. In the final stage of THS, average effort is again close to Nash predictions.

E. The Relevance of Strategic Complexity and Clearly Defined Hierarchies

In the previous sections we documented our main finding: subjects tend to provide excess effort in early stages of multistage elimination tournaments. In what follows, we explore potential reasons for this deviation from theoretical predictions. We concentrate on two factors that are inherent to elimination tournaments with multiple stages—higher strategic complexity, on the one hand, and a more refined definition of hierarchical positions on the other.

The higher strategic complexity of the decision environment makes elimination tournaments with multiple stages cognitively more demanding. When deciding on first-stage effort in the TS treatment, participants have to look ahead and be able to evaluate contingencies in the second stage. Excess effort could be the result of systematic cognitive mistakes or misperceptions that let subjects overestimate the option value of participating in later stages. For example, subjects could hold wrong beliefs about others' behavior in the second stage.[10] More generally, they could overestimate their chances of winning the tournament upon reaching the second stage. Subjects may also fail to fully account for the cost associated with achieving the higher wages in future stages, that is, they might at least partly neglect the effort cost in the second stage.

The more refined hierarchy, which is another natural characteristic of multistage elimination tournaments, provides a second potential source of excess effort in the TS treatment. In our experiments, we control for the direct and indirect economic value of being promoted. That is, we exactly know the option value of future internal promotion, and we can set the external "signaling" value of a promotion that could generate career concerns to zero (see, e.g., Gibbons and Murphy 1992; Holmström 1999). However, subjects might attach additional nonmonetary value to holding a certain position in the hierarchy per se. For instance, they might derive direct utility from being the unique overall winner of the tournament or from achieving the final. This extra utility could, for instance, stem from a "joy of winning" when being promoted (Parco, Rapoport, and Amaldoss 2005; Kräkel 2008) or from social comparison and status preferences (Moldovanu, Sela, and Shi 2007; Dohmen, Falk, Fliessbach, et al. 2011; Kuhnen and Tymula 2011).

To examine the relevance of the two factors we conduct an additional control experiment—the one-stage lottery treatment (OSL). This treat-

[10] However, in order to rationalize excess effort to the degree which we observe, beliefs about second-stage effort would need to deviate substantially from actual second-stage behavior and from stated beliefs. The median subject in TS exerts an effort of 84 in the first stage. This subject would have to expect second-stage efforts of 55 if wrong beliefs about what happens in the second stage were the only reason for overestimating the option value. This value is, however, almost 30% lower than actual effort choices in the second stage. In addition, subjects in TS on average stated second-stage effort expectations of 81. It is therefore unlikely that wrong beliefs about second-stage behavior are a driving force of excess effort in the first stage.

ment is constructed to display the more refined hierarchy of the TS treatment, while at the same time entailing the lower complexity of a one-stage tournament. In the OSL treatment, four people compete for the option to participate in a lottery by exerting effort once. Output and cost functions are identical to those in our two main treatments. The two players with the lowest outputs do not participate in the lottery and earn a wage of €5.73. Between the two participants with the highest outputs a coin toss determines who receives a wage of €17.57 and who receives a wage of €9.68. The wages and odds in the lottery of OSL are chosen such that they exogenously implement the symmetric equilibrium of the second stage in TS. As a consequence, the OSL treatment is strategically equivalent to the first stage of our two main treatments, with an equilibrium effort level of 74 (see col. 4 of table 1).

If deviations from equilibrium behavior in TS are driven solely by the higher complexity in this treatment, we should observe that efforts in OSL are equal to those in the OS treatment, since the latter two treatments share the lower complexity of one-stage tournaments. In contrast, if positional concerns and the more refined hierarchy are the main driving force for excess effort in TS, we expect that behavior in OSL is equal to the TS treatment, since both treatments entail the same hierarchical positions. Finally, if excess effort in the two-stage treatment is driven by a mixture of both factors, we should observe that efforts in OSL fall in between those of OS and TS.

A total of 100 subjects participated in the OSL treatment. Column 7 of table 2 summarizes the outcomes. Average effort in the OSL treatment is 77.3, with a median of 78. While these values lie somewhat above the equilibrium prediction of 74, the difference turns out to be insignificant (t-test, $p = .199$). Comparing the values obtained in the OSL treatment to those of our two main treatments shows that efforts in OSL lie in between those of OS and the first stage of TS. The difference in efforts to the OS treatment is not statistically significant, whereas the difference to the TS treatment is (weakly) significant (Mann-Whitney U-tests, $|z| = .94$, $p = .349$, and $|z| = 1.81$, $p = .070$, respectively). Taken together, the results suggest that a mixture of both factors might account for excess efforts in early stages of multistage tournaments. The fact that efforts in OSL are lower than in the TS treatment indicates that excess effort in the latter may partly be driven by cognitive factors and an incapacity to correctly perceive option values in multistage tournaments. On the other hand, efforts in OSL are on average still more than five points above those in OS, suggesting that positional concerns or joy of winning could also play a role in explaining the phenomenon of excess effort provision in multistage tournaments.

An alternative explanation for why efforts in OSL lie in between those of TS and OS is that participants do not derive extra utility from winning a competition, but from competing per se. In early stages of multistage tournaments, such preferences for competition would generate an additional option value for participants, since—by being promoted—they gain the right to enjoy further competition in future stages. In the final stage of multistage tournaments as well as in one-stage tournaments, this additional option value is reduced to zero, which could explain why we do not observe excess effort in these stages. Utility from competition might also be reduced in the OSL treatment, where subjects can only gain the option to participate in a lottery rather than the option to actively participate in a second-stage competition.

Result 5. Efforts in the one-stage lottery treatment OSL lie in between the values of OS and the first stage of TS.

F. Assessing Heterogeneity in Effort Provision

In both our main treatments, we observe substantial heterogeneity in efforts across individuals. In this section we ask whether heterogeneity in effort provision can be related to differences in participants' personal characteristics. This is important in light of the fact that a substantial amount of unexplained variation in efforts characterizes many empirical studies on tournaments (e.g., Bull et al. 1987; Falk, Fehr, and Huffman 2008). Studying the influence of personal characteristics on effort provision is directly related to the question of who is selected for promotion in a tournament-based incentive scheme. Promotions generally intend to select the most able employees into higher levels of the hierarchy. We abstract from such selection according to productive abilities (see Sec. II). However, if persons with certain individual characteristics systematically exert more effort than others, these are also more likely to be promoted, which may or may not be intended from an organizational perspective. Personal characteristics could also be related to our finding of excess effort in the first stage of the TS treatment, if the multistage structure in this treatment provoked especially competitive behavior of specific subgroups of contestants.

We focus on individual measures of competitiveness, attitudes toward risk, overconfidence, cognitive abilities, and forward-looking behavior which are elicited as follows. We use a 20-item questionnaire, the competitiveness index introduced by Smither and Houston (1992), to measure competitive inclination. Our incentivized measure for individual risk attitudes is the certainty equivalent for a lottery elicited in a series of binary lottery choices (see Sec. III). To measure overconfidence, subjects first had to answer 20 trivia questions. Before receiving feedback on their actual performance, subjects estimated how many questions they had answered correctly. This estimate was incentivized by rewarding a correct guess with €2. We use the difference between the estimated number of correct

Table 3
Relationship between Standardized Personality Measures and First-Stage Effort in OS and TS

	Dependent Variable: First-Stage Effort				
	Competitiveness (1)	Risk (2)	Overconfidence (3)	CRT (4)	Hit 15 (5)
P	.571	1.664	2.156	−3.185	2.683
	(3.133)	(3.474)	(4.844)	(4.512)	(5.353)
P × TS	3.287	1.435	−2.098	1.636	.269
	(4.500)	(4.708)	(6.549)	(6.372)	(6.640)
TS	12.658***	11.907**	19.599***	6.356	6.767
	(4.482)	(4.672)	(6.469)	(6.333)	(6.295)
Constant	71.719***	72.702***	69.628***	73.918***	73.623***
	(3.220)	(3.361)	(4.579)	(4.626)	(4.597)
N	124	114	64	60	60
R^2	.078	.066	.134	.031	.034

NOTE.—OLS coefficients (standard errors in parentheses). Column 1: *P* refers to the measure of competitiveness; col. 2: *P* = risk attitudes; col. 3: *P* = overconfidence; col. 4: *P* = cognitive reflection score; col. 5: *P* = Hit 15 score. *P* × TS refers to the interaction between the respective personality measure and the treatment dummy TS.
* $p < .10$.
** $p < .05$.
*** $p < .01$.

answers and actually correct answers as our proxy for overconfidence. We also administered the cognitive reflection test, which assesses individuals' ability to suppress spontaneous but wrong answers in favor of the less intuitive correct answer (Frederick 2005). As a measure for subjects' ability to think ahead, we had participants play the "Hit 15 points" game (Burks et al. 2008). This game requires reasoning backward from the game's goal, which is to reach 15 total points. Starting from varying initial numbers below 15, players take turns at adding between 1 and 3 points to a "point basket." Subjects play this game four times against a computer player who always best responds. They earn €1 for each round in which they add the fifteenth point to the basket.

To assess the relevance of personal characteristics for effort provision in general and for differences in behavior between treatments, we regress first-stage efforts (OS and TS pooled) on each of the personal characteristics in turn, including a treatment dummy for the TS treatment and an interaction term between the respective personal characteristic, *P*, and the treatment dummy.[11] Since all our personality measures use different scales, we standardize the measures to facilitate comparison of effects both within and across treatments. Table 3 depicts our results. The treatment dummy TS is positive and sizable in all specifications, indicating that the "mean" subject in each personality dimension exerts higher effort in TS than in

[11] For instance, *P* in column 1 of table 3 refers to our measure of competitiveness, and *P* × TS to the interaction of competitiveness with the TS treatment. In column 2, *P* refers to risk attitudes, in column 3 to overconfidence, and so on.

OS. Overall, we find only limited evidence of systematic relations between effort provision and the personal characteristics under consideration. The baseline coefficients for all personal characteristics as well as the interaction terms of characteristics with the TS dummy are individually and jointly insignificant in all specifications.

However, one has to take into account that these estimates are based on relatively small samples. At least in one of our measures, Hit 15, the low association with effort provision could also be attributable to a lack of variation in the personality measure. In our sample, 57% of subjects solve all four games correctly; on average, subjects solve 3.4 games. These values illustrate the limits of analyzing personality characteristics in a sample of subjects with relatively homogeneous backgrounds. The results may well look different in a more heterogeneous sample.[12] Moreover, our coefficients could be biased downward due to measurement error. To relax this concern, we make use of multiple measures on a given characteristic for the personality dimensions in which we have additional information from our postexperimental questionnaire (risk attitudes as well as cognitive and planning abilities). However, the relatively weak and insignificant associations between effort provision and personal characteristics still obtain if we include the principal components of multiple measures of risk, or forward-looking behavior, instead of the individual measures.

Some of the coefficients reported in table 3 are at least suggestive in that the effort-personality associations go in intuitive directions. For instance, in both treatments an increase in competitive inclination tends to increase effort. Higher levels of cognitive reflection tend to reduce effort and, consequently, the tendency to provide excess effort. The qualitative differences in slopes between OS and TS also suggest that—in addition to an overall positive treatment effect at the means—some personal characteristics make contestants behave particularly aggressively in the two-stage tournament. For instance, a one-standard-deviation increase in competitive inclination increases effort by 0.6 points in OS, whereas the corresponding value is 3.9 in the TS treatment. Although the difference in coefficients is not statistically significant, a qualitatively stronger influence of competitiveness in TS would be consistent with the notion that some subjects derive direct utility from competing. The extra option value generated by such a preference for competition should be higher in the first stage of TS than in the OS treatment.

Overall, our analysis of heterogeneity in behavior indicates that above-equilibrium effort provision in the TS treatment is a widespread phe-

[12] For instance, the original study of Burks et al. (2008) reports a mean of 2.4 and a higher variance in the number of solved tasks for a sample of truck drivers. Only 25% of their subjects solve all four games correctly.

nomenon that is not confined to specific subgroups of our sample.[13] At the same time, our data also suggest some interesting differences in the degree to which certain groups of participants increase their effort in TS relative to the OS treatment.

Result 6. Effort choices seem only weakly related to individual differences in personal characteristics, and subjects tend to exert excess effort in TS irrespective of their personal characteristics.

V. Conclusions

Promotions in most hierarchical organizations take the form of multistage elimination tournaments. In this paper we have studied behavior in such tournaments with controlled laboratory experiments. Our results demonstrate the importance of carefully analyzing the incentive effects of promotions in multilevel hierarchies. They show that the basic logic of incentive provision in multistage elimination tournaments works. People do take future promotion possibilities into account when deciding on current work effort. However, our experiments also suggest that behavior in multistage tournaments systematically deviates from behavior in one-stage tournaments in important ways. In particular, subjects tend to exert excess effort in the prefinal stages of our elimination tournaments. By contrast, we do not observe this phenomenon in a strategically equivalent one-stage tournament and in the final stages of our multistage tournaments.

A number of previous papers have used one-stage tournaments to study various aspects of promotion-based incentive schemes, such as the effects of different wage spreads (Bull et al. 1987; Harbring and Irlenbusch 2003) or sabotage activities (Falk et al. 2008; Harbring and Irlenbusch 2008). These papers have generally found supportive evidence for the tournament model. In view of our findings it is not clear to what extent these previous results translate to multistage setups. For instance, more competitive behavior in early stages might also lead to an increase in sabotage activities in multistage tournaments.

Further research is also called for regarding the role of contestants' ability in elimination tournaments. It would be interesting to analyze, for instance, whether differences in ability and the tendency to exert excess effort are complements or substitutes. Analyzing multistage tournaments with asymmetric contestants is an obvious next step given that selection

[13] This can further be illustrated if we analyze behavior of subgroups within the TS treatment. If we split our sample at the median of a given personal characteristic, average efforts are above equilibrium for all considered subgroups in TS. For instance, subjects with below-median competitiveness exert an effort of 81.3 whereas the corresponding value for those with above-median competitiveness is 87.7.

of more able agents into higher positions is an important function of promotion-based incentive schemes, and given that the efficiency of selection could be distorted if excess effort provision is especially pronounced for low-ability agents (see also Bull et al. 1987). Irrespective of what particular motivation leads a firm or organization to use promotions in a given context, one needs to bear in mind that promotions do have incentive effects and that people react to these incentives not always as predicted by theory.

References

Ameriks, John, Andrew Caplin, and John Leahy. 2003. Wealth accumulation and the propensity to plan. *Quarterly Journal of Economics* 118: 1007–47.

Baker, George P., Michael Gibbs, and Bengt Holmström. 1994. The internal economics of the firm: Evidence from personnel data. *Quarterly Journal of Economics* 109:881–919.

Baker, George P., Michael C. Jensen, and Kevin J. Murphy. 1988. Compensation and incentives: Practice vs. theory. *Journal of Finance* 43:593–616.

Belzil, Christian, and Michael L. Bognanno. 2008. Promotions, demotions, halo effects, and the earnings dynamics of American executives. *Journal of Labor Economics* 26:287–310.

Bognanno, Michael L. 2001. Corporate tournaments. *Journal of Labor Economics* 19:290–315.

Bull, Clive, Andrew Schotter, and Keith Weigelt. 1987. Tournaments and piece rates: An experimental study. *Journal of Political Economy* 95:1–33.

Burks, Stephen V., Jeffrey Carpenter, Lorenz Götte, Kristen Monaco, Kay Porter, and Aldo Rustichini. 2008. Using behavioral economic field experiments at a large motor carrier: The context and design of the Truckers and Turnover Project. In *The analysis of firms and employees: Quantitative and qualitative approaches*, ed. Stefan Bender, Julia Lane, Kathryn L. Shaw, Fredrik Andersson, and Till von Wachter. Chicago: University of Chicago Press.

Chevalier, Judith, and Austan Goolsbee. 2009. Are durable goods consumers forward-looking? Evidence from college textbooks. *Quarterly Journal of Economics* 124:1853–84.

Dohmen, Thomas, Armin Falk, Klaus Fliessbach, Uwe Sunde, and Bernd Weber. 2011. Relative versus absolute income, joy of winning, and gender: Brain imaging evidence. *Journal of Public Economics* 95:279–85.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner. 2011. Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association* 9:522–50.

Ehrenberg, Ronald G., and Michael L. Bognanno. 1990. Do tournaments have incentive effects? *Journal of Political Economy* 98:1307–24.

Eriksson, Tor. 1999. Executive compensation and tournament theory: Empirical tests on Danish data. *Journal of Labor Economics* 17:262–80.

Eriksson, Tor, Sabrina Teyssier, and Marie-Claire Villeval. 2009. Self-selection and the efficiency of tournaments. *Economic Inquiry* 47:530–48.

Fairburn, James A., and James M. Malcomson. 2001. Performance, promotion, and the Peter Principle. *Review of Economic Studies* 68:45–66.

Falk, Armin, Ernst Fehr, and David Huffman. 2008. The power and limits of tournament incentives. Unpublished manuscript, Department of Economics, University of Bonn.

Falk, Armin, and James J. Heckman. 2009. Lab experiments are a major source of knowledge in the social sciences. *Science* 326:535–38.

Farber, Henry S. 1999. Mobility and stability: The dynamics of job changes in labor markets. In *The handbook of labor economics*, ed. Orley C. Ashenfelter and David Card. 3 vols. Amsterdam: North-Holland Elsevier.

Fischbacher, Urs. 2007. z-Tree-Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10:171–78.

Frederick, Shane. 2005. Cognitive reflection and decision making. *Journal of Economic Perspectives* 19:25–42.

Gibbons, Robert, and Kevin J. Murphy. 1992. Optimal incentive contracts in the presence of career concerns: Theory and evidence. *Journal of Political Economy* 100:468–505.

Gibbons, Robert, and Michael Waldman. 1999. Careers in organizations: Theory and evidence. In *The handbook of labor economics*, ed. Orley C. Ashenfelter and David Card. 3 vols. Amsterdam: North-Holland Elsevier.

Greiner, Ben. 2003. An online recruitment system for economic experiments. In *Forschung und wissenschaftliches Rechnen 2003: GWDG Bericht 63*, ed. Kurt Kremer and Volker Macho. Göttingen: Gesellschaft für wissenschaftliche Datenverarbeitung.

Harbring, Christine, and Bernd Irlenbusch. 2003. An experimental study on tournament design. *Labour Economics* 10:443–64.

———. 2008. How many winners are good to have? On tournaments with sabotage. *Journal of Economic Behavior and Organization* 65:682–702.

Holmström, Bengt. 1999. Managerial incentive problems: A dynamic perspective. *Review of Economic Studies* 66:169–82.

Houser, Daniel, Michael Keane, and Kevin McCabe. 2004. Behavior in a dynamic decision problem: An analysis of experimental evidence using a Bayesian type classification algorithm. *Econometrica* 72:781–822.

Knoeber, Charles R., and Walter N. Thurman. 1994. Testing the theory of tournaments: An empirical analysis of broiler production. *Journal of Labor Economics* 12:155–79.

Kräkel, Matthias. 2008. Emotions in tournaments. *Journal of Economic Behavior and Organization* 67:204–14.

Kuhnen, Camelia M., and Agnieszka Tymula. 2011. Feedback, self-esteem and performance in organizations. *Management Science* (forthcoming).

Lazear, Edward P., and Michael Gibbs. 2008. *Personnel economics in practice*. 2nd ed. Hoboken, NJ: Wiley.

Lazear, Edward P., and Sherwin Rosen. 1981. Rank-order tournaments as optimum labor contracts. *Journal of Political Economy* 89:841–64.

Moldovanu, Benny, Aner Sela, and Xianwen Shi. 2007. Contests for status. *Journal of Political Economy* 115:338–63.

Orrison, Alannah, Andrew Schotter, and Keith Weigelt. 2004. Multiperson tournaments: An experimental investigation. *Management Science* 50: 268–79.

Parco, James E., Amnon Rapoport, and Wilfried Amaldoss. 2005. Two-stage contests with budget constraints: An experimental study. *Journal of Mathematical Psychology* 49:320–38.

Patton, Jim H., Matthew S. Stanford, and Ernest S. Barratt. 1995. Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology* 51:768–74.

Prendergast, Canice, and Robert H. Topel. 1996. Favoritism in organizations. *Journal of Political Economy* 104:958–78.

Preuss, Ulrich W., Dan Rujescu, Ina Giegling, Stefan Watzke, Gabriele Koller, Thomas Zetzsche, Eva Maria Meisenzahl, Michael Soyka, and Hans-Jürgen Möller. 2008. Psychometrische Evaluation der deutschsprachigen Version der Barratt-Impulsiveness-Skala. *Der Nervenarzt* 79:305–19.

Rosen, Sherwin. 1986. Prizes and incentives in elimination tournaments. *American Economic Review* 76:701–15.

Smither, Robert D., and John M. Houston. 1992. The nature of competitiveness: The development and validation of the Competitiveness Index. *Educational and Psychological Measurement* 52:407–18.

Waldman, Michael. 1984. Job assignments, signalling, and efficiency. *RAND Journal of Economics* 15:255–67.

Zábojník, Ján, and Dan Bernhardt. 2001. Corporate tournaments, human capital acquisition, and the firm size-wage relation. *Review of Economic Studies* 68:693–716.