Journal of Economic Behavior & Organization 102 (2014) 43–58

Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

Optimal prizes in dynamic elimination contests: Theory and experimental evidence☆

Rudi Stracke[a,*], Wolfgang Höchtl[b], Rudolf Kerschbamer[b], Uwe Sunde[a]

[a] University of Munich, Seminar for Population Economics, Geschwister-Scholl-Platz 1, D-80539 Muenchen, Germany
[b] University of Innsbruck, Universitätsstrasse 15, A-6020 Innsbruck, Austria

ARTICLE INFO

Article history:
Received 16 February 2012
Received in revised form 12 February 2014
Accepted 22 February 2014
Available online 12 March 2014

JEL classification:
C72
D72
J33

Keywords:
Dynamic contests
Multiple prizes
Risk aversion
Experiment
Over-provision

ABSTRACT

This paper investigates the implications of different prize structures on effort provision in dynamic (two-stage) elimination contests. Theoretical results show that, for risk-neutral participants, a structure with a single prize for the winner of the contest maximizes total effort, while a structure with two appropriately chosen prizes (a runner-up prize and a final prize) ensures incentive maintenance across stages. In contrast, a structure with two prizes may dominate a winner-takes-all contest in both dimensions if participants are risk-averse. Evidence from laboratory experiments is largely consistent with these predictions, suggesting that contest design should account for risk attitudes of participants.

1. Introduction

Contests are interactive decision situations in which agents compete by expending valuable resources to win a prize. Such situations appear in many different areas of economics and political economy – including the competition for bonus payments and promotions on internal labor markets, patent races in R&D, election campaigns, or military conflicts. Given the multiplicity of applications, it is not surprising that real world contests vary in several dimensions, for example, with respect to the number of participants, the number of prizes, or with respect to their structure. The effect of different modeling choices in these dimensions on behavior of contest participants has been studied extensively in theoretical work, which typically determines the optimal contest design with respect to a given optimality criterion under the simplifying assumption of common knowledge that all participants are rational and risk-neutral.[1] Two criteria are particularly prominent in the literature on optimal prizes in dynamic contests, namely the *maximization of aggregate incentives* (operationalized as the

☆ We thank the *Grundlagenforschungsfonds* of the University of St. Gallen for funding.
* Corresponding author. Tel.: +49 89 2180 1285; fax: +49 89 2180 17834.
E-mail addresses: rudi.stracke@lmu.de (R. Stracke), wolfgang.hoechtl@uibk.ac.at (W. Höchtl), rudolf.kerschbamer@uibk.ac.at (R. Kerschbamer), uwe.sunde@lmu.de (U. Sunde).
[1] See Konrad (2009) for a literature review.

http://dx.doi.org/10.1016/j.jebo.2014.02.018
0167-2681/

sum of efforts provided by all agents over all stages of the contest), and the *maintenance of incentives over all stages* of the contest (operationalized as constant individual efforts across stages). A common motivation for both objectives is that effort provision by contestants is valuable for the entity organizing the contest, henceforth called the contest designer. The maximization of aggregate incentives is a natural objective of the contest designer, in particular when efforts across stages are additively separable – see Sisak (2009) for an excellent survey of the literature addressing this criterion. Alternatively, complementarities between the efforts at different stages can imply that incentive maintenance across stages is the relevant criterion for the contest designer – the classical reference for this case is Rosen (1986) who argued that incentive maintenance is particularly important in corporate promotion contests.[2]

In this paper, we study the optimal design of a two-stage elimination contest with four homogeneous participants. Assuming that the overall prize money is fixed, we first show theoretically that a "winner-takes-all" structure with a single prize for the winner of the final round maximizes total effort under the standard assumption of common knowledge that all contestants are rational and risk-neutral. Then, we derive the prize structure that ensures incentive maintenance across stages under the same set of assumptions. As already shown by Rosen (1986), the optimal structure for the latter criterion turns out to be a format with two prizes, where the winner of the final receives most of the prize money, while a smaller part is assigned to the runner-up prize. Thus, the theoretical analysis based on risk neutrality suggests that there is a trade-off between the two optimality criteria 'maximization of aggregate efforts' and 'incentive maintenance across stages'. While we find that this trade-off does not depend on the number of stages or on the contest technology, the assumption that contestants are risk-neutral turns out to be crucial. In particular, under some simplifying assumptions (discussed in more detail in Section 2.3) we are able to show that the trade-off becomes weaker as the degree of (relative) risk aversion increases and might ultimately disappear for high degrees of risk aversion. We test these theoretical predictions using lab experiments. In particular, we implement two versions of the two-stage pairwise elimination contest that differ only in the structure of prizes: in the **single-prize** treatment (abbreviated as **SP**), we implement a "winner-takes-all" contest which allocates the entire prize money to the winner of the contest; this structure maximizes aggregate effort in the risk-neutral benchmark. In the **two-prizes** treatment (abbreviated as **TP**), the runner-up prize for the loser of the final is chosen such that incentives for effort exertion remain constant over both stages under risk neutrality. The experimental data reveal that effort choices by experimental subjects exceed their risk-neutral predictions substantially in both treatments. At the same time, the observed behavior is qualitatively in line with the theoretical predictions under risk neutrality: total effort is somewhat higher in **SP** than in **TP**, and incentives are maintained across stages in **TP**, but not in **SP**. The observed difference in total effort across treatments is smaller than predicted and statistically insignificant, however. To test whether risk attitudes account for the economically small and statistically insignificant difference across treatments in the total effort dimension, we disaggregate effort choices by risk attitudes of experimental subjects. We find strong evidence for the trade-off between total effort maximization and incentive maintenance for risk-neutral decision makers: for these subjects aggregate effort is significantly higher in **SP** than in **TP**, while only the **TP** treatment with a positive runner-up prize is capable to maintain incentives across stages. Moreover, the data suggest that the **TP** structure is better in both performance dimensions for strongly risk-averse subjects: incentives are only maintained in **TP**, and total effort is (weakly) higher in **TP** than in **SP**.

Our results matter for many contest schemes observed in reality. In sports competitions, for example, it is arguably not (only) the overall intensity across all playoff matches that matters for the interest of spectators (and therefore for the revenues a league is able to generate); it is also relevant how intensely players or teams fight in each game. For this application our results imply that a contest format with multiple prizes might well outperform a single-prize format in terms of overall intensity and in terms of maintaining intensity across plays if participants are sufficiently risk-averse. In the context of human resource management, our results suggest that a promotion contest might be able to kill two birds with one stone if employees are risk-averse. Since it is more likely to be the rule than the exception that employees are risk-averse, our results also provide a novel testable hypothesis for future empirical work. Both theoretical and experimental results suggests that firms who mainly care about total effort might want to choose an occupation-specific incentive profile for promotions. The optimal "steepness" of this profile is determined by the average risk-attitude of employees in a particular occupation, and there is convincing evidence that risk attitudes matter for sorting of employees into different occupations.[3]

Our results contribute to the recent experimental literature on optimal contest design. For a long time, the experimental literature focused on static contests.[4] In recent years, attention has shifted to dynamic settings and an increasing number of researchers now study dynamic contests with multiple stages. Our results are most closely related to Altmann et al. (2012) and Sheremeta (2010a), who both compare static (one-shot) and dynamic (two-stage) contests.[5] The paper by Altmann et al. (2012) considers a prize structure that predicts incentive maintenance across stages in the risk-neutral benchmark, and one of their main findings in the experiments is that effort provision in the first stage is much higher than in the second stage.

[2] Incentive maintenance is also an issue in sports tournaments where it is important to make all stages of the competition 'exciting' for the participants and observers.

[3] See, for instance, Bonin et al. (2007), Dohmen and Falk (2011), or Pollmann et al. (2012).

[4] Bull et al. (1987) is the seminal paper. See also Harbring and Irlenbusch (2003), Harbring and Lünser (2008) or Sheremeta (2011), as well as the references provided therein.

[5] Other related studies of dynamic contests are Parco et al. (2005), Amaldoss and Rapoport (2009), or Sheremeta (2010b).

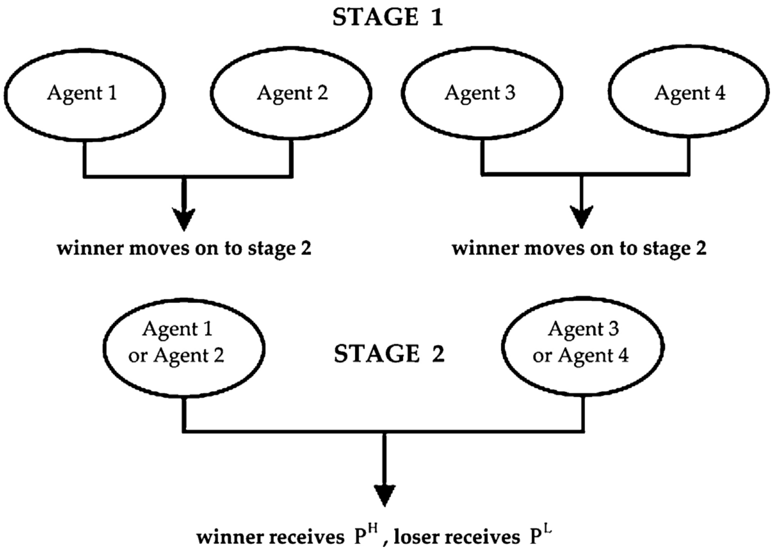

Fig. 1. Structure of the dynamic contest.

Sheremeta (2010a) investigates a single-prize two-stage contest format and compares it to an analogous one-stage contest interaction. Our paper combines the two approaches and analyzes a systematic variation of the prize structure in a dynamic contest. The research question we address is also related to recent work by Delfgaauw et al. (2014). They investigate whether a more convex prize spread affects relative effort exertion across different stages of a bonus tournament. Using data from a field experiment, they find that the effect of the prize structure on relative effort provision across stages is rather weak. The effect of the prize structure appears to be much stronger in our experimental data. A possible explanation for this difference in magnitude is that our prize-spread variation is more extreme, since we compare a "winner-takes-all" structure with a two-prizes setting, while Delfgaauw et al. (2014) investigate the effects of a more modest variation of prizes in a setting with two prizes. Finally, our paper also contributes to the theoretical literature on contest design by providing a formal proof that a "winner-takes-all" contest maximizes total effort in a Tullock (1980) lottery contest with pairwise elimination and multiple stages under risk-neutrality. The results show that the "winner-takes-all" principle, which Fu and Lu (2012) establish for the optimal multi-stage (pyramid) contest, also holds in a pairwise elimination structure. In addition, our theoretical findings reveal that the assumption of risk-neutral contestants is not innocuous for contest design.

The remainder of this paper is organized as follows: Section 2 provides a theoretical analysis of the relation between the two optimality criteria (total effort provision and effort maintenance) in a simple dynamic contest model, and a discussion of the robustness of the theoretical findings. Section 3 outlines the experimental design and derives the main testable hypotheses. The experimental results are presented and discussed in Section 4. Section 5 concludes.

2. Theoretical analysis

2.1. A simple dynamic elimination contest: risk-neutral benchmark

Set-up. Consider a two-stage pairwise elimination contest where four identical agents compete for two prizes. The sequence of events is illustrated in Fig. 1. In the first stage, two pairs of agents compete simultaneously for the right to move on to stage 2. Participation in stage 2 is valuable, since two prizes are awarded to the participants of this stage: the loser of the stage-2 interaction receives the prize P^L, while P^H is awarded to the winner, where $P^H > P^L \geq 0$. In each of the three interactions of this contest model, two agents independently choose their effort levels to maximize their expected payoff. In line with the literature on contest design, we start by assuming that all agents are risk-neutral and that this is common knowledge. We call this the risk-neutral benchmark.[6] The effort of agent i in stage $s \in \{1, 2\}$ is denoted $x_{si} \geq 0$. For each invested unit, agents incur constant marginal costs of one. The benefit of effort provision is that the probability to win an interaction is increasing in the amount invested into the contest. Thus, agents face a trade-off. Throughout, we assume that

[6] Below (in Section 2.3) we will discuss the impact of different levels of risk aversion on equilibrium outcomes.

the probability to win is given by a lottery contest success function á la Tullock (1980). That is, given investments x_{si} and x_{sj} by agents i and j in stage s, the probability that agent i wins against agent j in stage s equals

$$p_{si}(x_{si}, x_{sj}) = \begin{cases} \dfrac{x_{si}}{x_{si}+x_{sj}} & \text{if} \quad x_{si}+x_{sj} > 0 \\ \dfrac{1}{2} & \text{if} \quad x_{si}+x_{sj} = 0 \end{cases}.$$

Equilibrium. The equilibrium concept is Subgame Perfect Nash and determined by backward induction. Since agents are homogeneous, the identity of the agents who compete in stage 2 does not affect the solution. Therefore, without loss of generality, it is assumed that agents i and j interact in stage 2. Consider the optimization problem of agent i who takes the effort choice x_{2j} by j as given and chooses stage-2 effort x_{2i} to maximize her expected payoff $\Pi_{2i}(x_{2i}, x_{2j})$. Formally, the problem reads:

$$\begin{aligned} \max_{x_{2i}\geq 0} \Pi_{2i}(x_{2i}, x_{2j}) &= \frac{x_{2i}}{x_{2i}+x_{2j}} P^H + \left(1 - \frac{x_{2i}}{x_{2i}+x_{2j}}\right) P^L - x_{2i} \\ &= \frac{x_{2i}}{x_{2i}+x_{2j}} (P^H - P^L) + P^L - x_{2i}. \end{aligned}$$

The first-order condition $x_{2j}/((x_{2i}+x_{2j})^2)(P^H - P^L) - 1 = 0$ is both necessary and sufficient for optimality (Perez-Castrillo and Verdier, 1992). Using symmetry delivers equilibrium efforts

$$x_2^* := x_{2i}^* = x_{2j}^* = \frac{(P^H - P^L)}{4}. \tag{1}$$

Inserting equilibrium efforts in the objective functions gives the expected stage-2 equilibrium payoff

$$\Pi_2^* := \Pi_{2i}(x_{2i}^*, x_{2j}^*) = \Pi_{2j}(x_{2j}^*, x_{2i}^*) = P^L + \frac{(P^H - P^L)}{4}. \tag{2}$$

Reaching stage 2 has value Π_2^* for an agent that participates in stage 1. Agent k will take this value into account when choosing her stage-1 effort x_{1k}. Since all agents are identical by assumption, consider the interaction between two arbitrary agents k and l in stage 1. Agent k faces the optimization problem

$$\max_{x_{1k}\geq 0} \Pi_{1k}(x_{1k}, x_{1l}) = \frac{x_{1k}}{x_{1k}+x_{1l}} \Pi_2^* - x_{1k}.$$

As in the solution of stage 2 above, the first-order condition together with symmetry yields the equilibrium efforts on stage 1 as

$$x_1^* := x_{1k}^* = x_{1l}^* = \frac{P^L}{4} + \frac{(P^H - P^L)}{16}. \tag{3}$$

2.2. *Optimal structure of prizes*

Based on the simple two-stage contest, we subsequently derive the optimal structure of prizes for different optimality criteria. In particular, we consider two goals of the contest designer under the assumption that the overall prize money available is fixed: the maximization of total effort provided in the contest, and the maintenance of identical incentives for effort provision across stages. In particular, we are interested in the question whether both goals can be achieved at the same time, or whether they are mutually exclusive. This question is particularly important in the context of dynamic promotion contests. Intuitively, is seems reasonable that firms should try to maintain incentives across different levels of the hierarchy to avoid that employees slack off either early or late in their career, as suggested by Rosen (1986). At the same time, however, the overall effort provided across all stages and individuals is also important if the marginal effect of effort on profits is comparable across hierarchy levels.

Maximizing total effort. Since four agents provide effort in stage 1, while only two of them reach stage 2, total effort amounts to $\mathcal{E} = 4x_1^* + 2x_2^*$, where x_1^* is given by Eq. (3), and x_2^* by Eq. (1). Assuming that P units are available as total prize money (which implies $P^H = P - P^L$), total effort satisfies

$$\mathcal{E} = \frac{3P - 2P^L}{4}. \tag{4}$$

Since $\mathcal{E}$ is strictly decreasing in P^L, total effort is maximized in a "winner-takes-all" contest, i.e., for $P^L = 0$ and $P^H = P$. To see where this result comes from, consider the effect of an increase in the runner-up prize on total effort in the simple two-stage model. In particular, assume that P^H is reduced by ϵ to increase P^L by the same amount. This change reduces the spread between P^H and P^L for which agents compete in stage 2 by 2ϵ. Thus, it follows from (1) that each agent reduces her stage-2 effort by 0.5ϵ, such that aggregate stage-2 effort is now ϵ units lower. Can this loss in stage 2 be compensated for by additional effort in stage 1? Eq. (3) indicates that stage-1 participants compete both for the prize P^L (which they have for sure upon reaching stage 2) and for the continuation value of stage 2 (which offers the high prize P^H at the cost of additional effort).

Thus, the runner-up prize increase affects stage-1 effort through two channels: first, aggregate stage-1 effort increases one-to-one in P^L and thus by ϵ. Second, aggregate stage-1 effort increases at rate one-to-four in the spread $P^H - P^L$. Since this spread is reduced by 2ϵ, this channel reduces aggregate stage-1 effort by 0.5ϵ. The joint effect of both channels on aggregate stage-1 effort is 0.5ϵ. Since the effort loss in stage 2 amounts to ϵ, increasing the runner-up prize by ϵ will reduce total effort by 0.5ϵ.

Incentive maintenance across stages. Under the assumption that P units of prize money are available ($P^H = P - P^L$), stage-1 and stage-2 equilibrium efforts given in (1) and (3) read

$$x_1^* = \frac{P}{16} + \frac{P^L}{8} \quad \textit{and} \quad x_2^* = \frac{P}{4} - \frac{2P^L}{4}. \tag{5}$$

According to the expressions in (5), stage-2 effort x_2^* is strictly higher than stage-1 effort x_1^* in a winner-takes-all contest ($P^L = 0$). Moreover, stage-2 effort is decreasing in the runner-up prize P^L, while the opposite holds for stage-1 effort. Thus, incentive maintenance requires a strictly positive runner-up prize P^L. Equalizing the expressions for stage-1 and stage-2 effort given in (5) implies a runner-up prize of $P^L = 3P/10$, and a winner prize $P^H = 7P/10$. That is, the runner-up prize P^L is strictly positive, but smaller than the winner prize P^H.

2.3. Discussion and robustness

The previous analysis shows that there is a trade-off between the two goals total effort maximization and incentive maintenance across stages in the risk-neutral benchmark of the two-stage elimination contest: while aggregate effort is maximal with a single prize (equal to the total prize money) for the winner of the final, incentive maintenance across stages requires a strictly positive runner-up prize. In the following, we briefly discuss how the simplifying assumptions of our theoretical model affect this finding.

Contest technology. We consider the Tullock lottery contest technology with a discriminatory power of one in the benchmark of Section 2.1. In the generalized Tullock contest model – which allows for variations of the discriminatory power through the parameter $r \geq 0$ – the winning probability of agent i in a pairwise interaction between i and j in stage s is defined as

$$p_{si}(x_{si}, x_{sj}) = \begin{cases} \dfrac{x_{si}^r}{x_{si}^r + x_{sj}^r} & \text{if} \quad x_{si} + x_{sj} > 0 \\ \dfrac{1}{2} & \text{if} \quad x_{si} + x_{sj} = 0. \end{cases}$$

As is shown in Appendix A, the aforementioned trade-off is present in any pure strategy equilibrium, i.e., for any discriminatory power $0 < r < 2$. The intuition for the lower bound seems obvious: if effort has no effect on the probability to win (as is the case for $r = 0$), then agents will excert no effort in equilibrium and the prize structure become irrelevant for both goals. The intuition for the upper bound on r is that total effort approaches the overall amount of prize money as r approaches 2, which implies that the structure of prizes becomes again irrelevant for total effort.[7] Results by Krishna and Morgan (1998) suggests that the same holds for other contest technologies. They consider a Lazear and Rosen (1981) type model where random noise is additive and find that two prizes may maximize total effort in sequential elimination contests only if the noise parameter has very narrow bounds.

Number of stages. In the benchmark of Section 2.1, the elimination contest has two stages. Do our results depend on this detail? First, note that a runner-up prize, which is strictly smaller than the main prize, is necessary for incentive maintenance over all stages in a pairwise elimination structure with an arbitrary number of stages according to the seminal paper on promotion contests by Rosen (1986).[8] Moreover, we show in the appendix that a "winner-takes-all" prize structure maximizes total effort in any Tullock "lottery" contest with pairwise elimination and a finite number of agents and stages. This finding complements existing results on the total effort maximizing prize structure in dynamic contests by Fu and Lu (2012), who obtain the same result for a pyramid contest with sequential elimination that pools participants in each stage.

Risk attitudes. In line with much of the contest design literature we assumed that agents are risk-neutral in the benchmark model of Section 2.1. However, results by Krishna and Morgan (1998) for a contest technology with additive noise suggest that this assumption is not innocuous. They find that a positive runner-up prize may maximize total effort across stages in a four player contest with risk-averse agents, independent of the structure of the competition. Within the structure of a two-stage pairwise elimination contest the intuition for this finding is easily provided: a positive runner-up prize provides insurance for risk-averse agents against the worst possible outcome that occurs when costly effort is provided in both stages, and an agent loses in stage 2. In this sense, a runner-up prize reduces the riskiness of investments into the contest. As shown

[7] See Baye et al. (1994) for details.

[8] More precisely, Rosen (1986) shows that incentive maintenance requires a strictly positive reward for the winner of each stage. The reward remains constant across stages, except for the very last stage (where the winner receives a higher reward). For the two-stage elimination contest the prize structure that we consider (no prize for the winner of the first stage; a prize P^H for the winner and a prize P^L for the loser of the second stage) is strategically equivalent to a structure where the two winners of the first stage both receive the prize P^L, and where the winner of the second stage receives the additional reward $P^H - P^L$, while the loser of the second stage receives no additional reward.

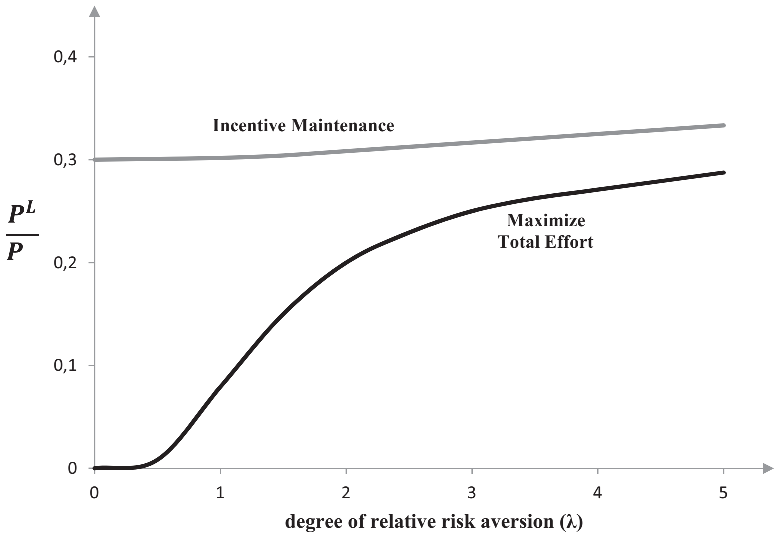

Fig. 2. Optimal runner-up prize with CRRA preferences.

by Konrad and Schlesinger (1997), this intuition does also hold in the Tullock contest model. Thus, a positive runner-up prize might be necessary to maximize total effort if agents are risk-averse, and the two goals considered here may then turn out to be compatible.

To address this issue, we modify the benchmark model of Section 2.1 by allowing for risk-averse agents. In particular, we assume that all agents have ‘constant relative risk aversion’ (CRRA) preferences of the form

$$U(X)=\begin{cases}\dfrac{X^{1-\lambda}}{1-\lambda}, & \text{if } \lambda\neq 1\\ \ln X, & \text{if } \lambda=1\end{cases}$$

where $\lambda\geq 0$ is the degree of relative risk aversion.[9] In this functional form a decision maker is risk-neutral for $\lambda=0$, and increasingly risk-averse as λ increases. Since a closed-form analytical solution is out of reach for this case, we solve the model numerically under the simplifying assumption that all agents have the same λ (and that this is common knowledge).

Fig. 2 plots the optimal share of total prize money P that a contest designer should allocate to the runner-up prize P^L as a function of λ in order to achieve the two goals, maximization of total effort and incentive maintenance. Consider the maximization of total effort first: in line with the formal analysis in the previous section, the figure reveals that a “winner-takes-all” contest with $P^L/P=0$ is optimal for risk-neutral agents ($\lambda=0$). This changes as λ increases, however. In particular, a positive runner-up prize tends to increase (rather than decrease) total effort if contestants are sufficiently risk-averse. Moreover, the share of the total prize money that is optimally allocated to the runner-up prize increases with λ. Next, consider the criterion of incentive maintenance across stages. The figure confirms our earlier result that a positive runner-up prize is necessary for incentive maintenance under risk-neutrality. Moreover, the share of the total prize money that is allocated to the runner-up prize is increasing in the degree of relative risk aversion. Intuitively, risk aversion reduces the value that contestants attach to a stage-2 participation, since the stage-2 outcome is uncertain. Consequently, a somewhat higher runner-up prize is necessary to offset the incentives of contestants to provide less effort in stage 1. The optimal share of the prize money awarded to the runner-up prize increases very modestly in λ, however, since the possibility to win the high prize P^H (the risky part) accounts for only 25% of the value that agents attach to a stage-2 participation under risk-neutrality. Thus, the runner-up prize that ensures incentive maintenance across stages must be only slightly higher if agents are risk-averse rather than risk-neutral, and that prize is still much smaller than the main prize.

Taken together, our results suggest that risk aversion weakens the trade-off between the maximization of total effort and incentive maintenance across stages. In particular, both objectives require a positive runner-up prize if contestants are risk-averse. As illustrated by Fig. 2, this implies that the difference between the optimal runner-up prizes becomes smaller as the degree of risk aversion increases.

[9] CRRA preferences constitute a natural benchmark. In addition, we considered a utility function that exhibits constant absolute risk aversion (CARA) and obtained qualitatively and quantitatively similar results. Details available upon request.

Table 1
Parameterization and predictions for risk-neutral benchmark.

	Single prize (**SP**)	Two prizes (**TP**)
Total effort ($\mathcal{E}$)	180	144
Stage-1 effort (x_1^*)	15	24
Stage-2 effort (x_2^*)	60	24
Prizes (P, P^L, P^H)	(240,0,240)	(240,72,168)

3. Design of the experiments

3.1. Experimental treatments, parameters, and testable hypotheses

We implement two treatments that correspond to two variants of the two-stage pairwise elimination contest in a lab experiment. The treatments differ only in the structure of prizes. In both treatments, the total prize money amounts to 240 units, which implies that $P^H + P^L = 240$. In the single-prize treatment (**SP**), we implement a "winner-takes-all" contest with $P^L = 0$ and $P^H = 240$. The runner-up prize is positive in the two-prizes treatment (**TP**), where $P^L = 72$ and $P^H = 168$. Table 1 shows the risk-neutral benchmark predictions for the two treatments. The table reveals that the prizes in the two treatments are chosen in such a way as to allow for a neat test of the trade-off between the maximization of aggregate effort and incentive maintenance across stages: due to the negative impact of the runner-up prize on aggregate effort, the total effort measure is predicted to be higher in **SP** than in **TP**. At the same time, the size of the runner-up prize in **TP** is chosen such as to ensure incentive maintenance across stages in the risk-neutral benchmark. As discussed in the previous section, the trade-off between the maximization of total effort and incentive maintenance becomes weaker as the degree of risk aversion increases. Thus, the **TP** treatment may turn out to dominate the **SP** treatment in both dimensions if agents are sufficiently risk averse. To address this question, we elicited risk attitudes of experimental subjects in both treatments. The subsequent testable hypotheses are based on the assumption of common knowledge that all agents are risk-neutral, which provides clear-cut predictions and a natural benchmark in light of the existing literature. Whether and to what extent risk aversion affects these hypotheses will be addressed below when the experimental results are presented.

In a first step, we examine the relation of the total effort measure across treatments **SP** and **TP**. Aggregate effort is predicted to be higher in **SP** than in **TP**, which delivers the first testable hypothesis:

Hypothesis 1 *(Total effort maximization).* Aggregate effort across all contestants and both stages is higher in **SP** than in **TP**:

$$\mathcal{E}^{\mathrm{SP}} > \mathcal{E}^{\mathrm{TP}}.$$

Apart from information on total effort provision, Table 1 provides the individual effort levels in each stage of both treatments. Effort choices are predicted to be the same across both stages in **TP**, but not in **SP**, which leads to Hypothesis 2.

Hypothesis 2 *(Incentive maintenance).* Individual efforts are identical across stages in **TP**, while individual efforts are lower in stage 1 than in stage 2 in **SP**:

(a) $x_1^{\mathrm{TP}} = x_2^{\mathrm{TP}}$
(b) $x_1^{\mathrm{SP}} < x_2^{\mathrm{SP}}$.

Table 1 also reveals that individual effort in stage 1 is higher in treatment **TP** than in **SP**, while the opposite holds for stage-2 effort. Intuitively, stage-1 effort is strictly increasing in the runner-up prize P^L, since a higher runner-up prize makes participation in stage 2 more valuable. Stage-2 effort is, however, decreasing in P^L. The reason is that each stage-2 participant receives the runner-up prize for sure, such that the two participants compete only for the residual prize $P^H - P^L$. We call this mechanism the "Runner-up Prize Effect" and test it in Hypothesis 3:

Hypothesis 3 *(Runner-up prize effect).* In stage 1, individual effort is higher in **TP** than in **SP**, while the opposite holds for stage-2 effort:

(a) $x_1^{\mathrm{SP}} < x_1^{\mathrm{TP}}$
(b) $x_2^{\mathrm{SP}} > x_2^{\mathrm{TP}}$.

Note that the strength of the "Runner-up Prize Effect" is at the heart of the result that a "winner-takes-all" prize structure maximizes total effort. Intuitively, the effort gain in early stages cannot compensate for the loss in later stages in the risk-neutral benchmark model, even though the number of participants is higher in early stages.

3.2. Experimental procedures

Protocol of an experimental session. We adopt a between-subject design; that is, our experimental subjects encountered either the **SP** or the **TP** treatment. The protocol of an experimental session was the same for both treatments: first, participants received some general information about the experimental session. Then, instructions for the respective treatment (either **SP** or **TP**) were distributed.[10] After each participant confirmed that he/she had read and understood the instructions, participants had to answer a set of control questions correctly. Only then did the first decision round start. Overall, each subject participated in 30 decision rounds. To rule out repeated game effects, subjects were randomly and anonymously rematched in each round. Matching groups corresponded to the entire session. After the main treatment, we first elicited risk preferences using a standard incentivized procedure, and then asked participants to fill out a questionnaire (voluntary and non-incentivized). Only thereafter participants were informed about their earnings in the experimental session. We ran a total of 10 computerized sessions (five for each treatment) with 20 participants in each session. The experiment was programmed in z-Tree (Fischbacher, 2007). All 200 participants were students from the University of Innsbruck, which were recruited using ORSEE (Greiner, 2004). We used the experimental currency "Taler", where 200 Taler corresponded to 1.00 Euro. Each session lasted approximately 70 minutes in total (including the distribution of instructions at the beginning and the payment at the end), and participants earned between 9 and 13 Euro (approximately 11 Euro on average).[11]

Implementation of the contest. Each participant played the same contest game 30 times, knowing that the identities of his/her opponents were randomly determined in each decision round. The role of investments into the contest (effort) was explained to subjects using a lottery analogy: participants were told that they could buy a discrete number of balls in each interaction.[12] The balls purchased by the subjects as well as those purchased by their respective opponents were then said to be placed in the same ballot box, out of which one ball would be randomly drawn subsequently. This set-up replicates the Tullock (1980) lottery contest technology studied in the theoretical analysis. Players had to buy (and pay for) their desired number of balls before they knew whether or not they won a pairwise interaction in the contest. For this purpose, each participant received an endowment of 240 Taler in each round. This endowment could be used to buy balls on both stages, i.e., a subject that reached stage 2 could use whatever remained of his/her endowment to buy balls for the stage-2 interaction. The part of the endowment that a participant did not use to buy balls was added to the payoffs for that round. Since the endowment was as high as the total prize money P, agents were not budget-constrained at any time.[13] Experimental subjects were told that the endowment could only be used in a given round, that is, that transfers across decision rounds were not possible. Therefore, the strategic interaction is the same in each of the 30 decision rounds. After each decision round, participants were informed about their own decision, the decision of their immediate opponent, and about their own payoff. This allows for an investigation of whether players learn when completing the task repeatedly. In order to minimize the potential impact of income effects participants were told that only four decision rounds (out of 30) would be randomly chosen and paid out at the end of the experiment.

Elicitation of risk attitudes. We used a choice list similar to the one employed by Dohmen et al. (2010) to elicit risk attitudes. Specifically, each subject was exposed to a series of 21 binary choices between a cash gamble and a safe payoff. The cash gamble remained the same in all 21 binary choices – it always gave either 400 Taler or 0 Taler, each with 50 percent probability. The safe payoff increased in steps of 10 Taler from 100 Taler in the first choice to 300 Taler in the last choice.[14] Given this design, a decision maker whose preferences satisfy ordering (completeness and transitivity) and strict monotonicity switches exactly once from the cash gamble to the safe payoff. The first choice scenario in which the subject decides in favor of the safe payoff serves as measure of risk attitude.

4. Experimental results

The main experimental results are summarized in Table 2. The table displays the risk-neutral benchmark predictions from the model in Section 2.1 as well as observed means for stage-1, stage-2, and total effort provision in both treatments. The data match all predicted relations qualitatively, even though the empirically observed efforts substantially exceed their theoretical counterparts. This finding of over-provision of effort is in line with much of the existing experimental literature and will be discussed at the end of this section.

[10] A translated version of the instructions is provided in the online appendix. The original (German) instructions are available from the authors upon request.

[11] In two (one) out of five sessions of the **SP** (**TP**) treatment, an additional experiment was conducted *after* the risk-elicitation part. This experiment was entirely unrelated to the contest experiment and subjects were not informed about what to expect in this third experiment. All they knew is that the session included a third part, rather than only two parts. These sessions where approximately 15 minutes longer, and payoffs in this additional experiment amounted to approximately 2.50 Euros, on average.

[12] The chosen prizes ensured that the discrete grid had no consequences for the equilibrium strategies in the risk-neutral benchmark of both treatments.

[13] This is also confirmed by the experimental data on effort.

[14] This procedure is cognitively simpler than the one employed by Holt and Laury (2002), where a subject is confronted with a series of choices between two binary lotteries that are both varied systematically. The instructions which experimental subjects received right before the risk-elicitation part are provided in the Appendix.

Table 2
Experimental results.

	SP		TP	
	Data	Benchmark	Data	Benchmark
Total effort ($\mathcal{E}$)	305.133	180	277.861	144
Stage-1 effort (x_1^*)	34.328	15	45.238	24
Stage-2 effort (x_2^*)	83.470	60	45.976	24

Note: Numbers in the columns "Data" denote averages over all rounds of the experimental sessions, measured in experimental currency ("Taler"). Total effort is computed as the average total effort over both stages by subject, multiplied by four; stage-1 (stage-2) effort is the average individual effort on that stage. The column "Benchmark" provides the theoretical prediction for the respective effort measure under risk-neutrality.

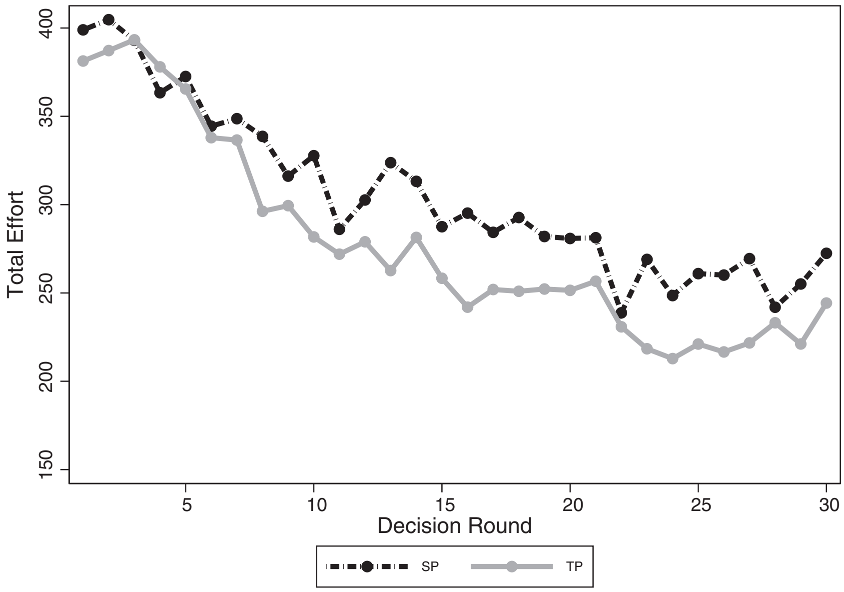

Fig. 3. Total effort by decision round and treatment.

4.1. *Experimental data vs. risk-neutral benchmark*

We proceed in the same order as in Section 3, starting with the comparison of total effort across treatments. In line with the risk-neutral benchmark prediction, total effort is higher in **SP** than in **TP** (305.133 compared to 277.861, see Table 2 for details). However, the difference appears to be smaller than predicted in relative terms. In particular, total effort is only 10% higher in **SP** than in **TP**, while theory predicts that it is 25% higher. A panel regression (with session random effects) of total effort across both stages on the treatment dummy indicates that the difference is statistically insignificant (p-value 0.13)[15]; the p-value for the non-parametric Mann–Whitney U-test (MWU-test) is 0.25.[16] Fig. 3 plots the evolution of total effort over time from rounds 1 to 30. Two findings emerge: first, total effort is decreasing across decision rounds in both treatments. Thus, it seems that participants realize in both treatments that their effort choices were initially too high. Even in later decision rounds, however, total effort is still well above the risk-neutral benchmark in both treatments. Second, total effort is very similar across treatments in early rounds of the experiment, but it seems that the treatment effect gains strength in later rounds of the experiment. Formal tests confirm both observations: there is a significant and negative time trend in both treatments, and the p-values for the treatment effect with respect to total effort are lower in the second half of the experiment (round 16–30); the p-value is still above 0.05 both for the regression with session random effects and the MWU-test, however. Thus, we summarize our findings with respect to total effort as follows:

Result 1 (Total effort maximization). Aggregate effort across all contestants and both stages of the contest is higher in **SP** than in **TP**, in line with the risk-neutral benchmark prediction. However, the difference is smaller than predicted and statistically insignificant.

[15] Controlling for a period trend or including period-dummies has no effect on the level of significance. Allowing for clustering at the session level delivers slightly more significant results, but might be problematic due to the small number of clusters.

[16] In line with the literature, we use session means as independent observations throughout this paper, i.e., we have five independent observations for each treatment and each effort measure.

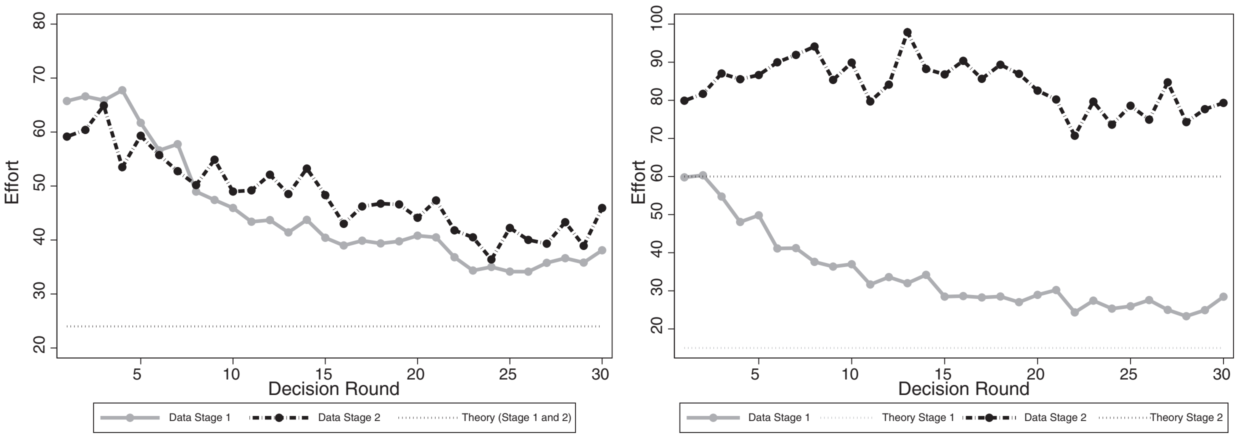

Fig. 4. Individual effort by treatment and decision round.

Hypothesis 2 is concerned with the maintenance of incentives across stages and states that individual efforts are identical across stages in **TP**, but lower in stage 1 than in stage 2 in **SP**. According to Table 2, this is exactly what is observed in the experiment: in **TP**, subjects invest approximately 45 units of effort in both stages, and we cannot reject the null of equality of individual session means. In contrast, we can clearly reject the equality of individual session means across stages in **SP**, where subjects invest much less in stage 1 than in stage 2.[17] Fig. 4 plots the stage-1 and stage-2 effort choices in both treatments over the different rounds of the experiment. The figure shows that both stage-1 and stage-2 efforts are decreasing across decision rounds in **TP**, while this holds only for stage-1 effort in **SP**. Even in the later decision rounds, however, effort choices remain well above the risk-neutral benchmark in both treatments. Moreover, the figure reveals that stage-1 effort choices are clearly lower than stage-2 efforts in any decision round in treatment **SP**, while effort choices in the two stages are similar throughout in treatment **TP**. Note, however, that the relation between stage-1 and stage-2 effort choices changes in treatment **TP**. Initially, stage-1 effort is slightly higher than stage-2 effort, while this relation is reversed after the first seven decision rounds. This observation might help to explain why we find incentive maintenance across stages in **TP**, while Altmann et al. (2012) in a similar experiment do not. Altmann et al. employ an experimental design where participants interact only once. In contrast, in our experiment the same contest is repeated 30 times with random rematching of experimental subjects.[18] If we only consider the first decision round, the data replicate the pattern observed by Altmann et al. (2012): in this round, subjects in **TP** choose, on average, an effort of 65.75 in stage 1, compared to 59.16 in stage 2.[19] However, this pattern disappears and is even reversed in later rounds, as discussed above. In fact, the equality of stage-1 and stage-2 efforts in **TP** can be rejected in some of the first seven decision rounds, while equality cannot be rejected in any subsequent round.

Result 2 (Incentive maintenance). In line with the risk-neutral benchmark prediction, we cannot reject incentive maintenance across stages in **TP**; incentive maintenance does not hold in **SP**, where effort choices are much lower in stage 1 than in stage 2.

Our third and last hypothesis addresses the opposing effects of the runner-up prize on effort choices in the two stages: theory predicts that a runner-up prize increases effort provision in stage 1 at the expense of stage-2 effort. Table 2 shows that this pattern is present in the experimental data: effort choices by experimental subjects in stage 1 are higher in **TP** than in **SP** (45.238 vs. 34.328), and the difference is highly significant ($p < 0.01$ for the regression with session random effects, and $p < 0.05$ for the MWU-test). In contrast, stage-2 effort is higher in **SP** than in **TP** (83.470 vs. 45.976), and again the difference is highly significant ($p < 0.01$ for both testing procedures). Fig. 5 illustrates that the observed pattern is present in each single round and remains constant across decision rounds. Thus, our findings are in line with Hypothesis 3:

Result 3 (Runner-up prize effect). The comparison of efforts in a given stage across treatments shows that the introduction of a runner-up prize has the effect predicted by the benchmark model: stage-1 effort is higher in **TP** than in **SP**, while stage-2 effort is higher in **SP** than in **TP**.

The opposing effects of the runner-up prize on the effort choices in the two stages are weaker than in the risk-neutral benchmark, however. Theory predicts that the runner-up prize increases stage-1 effort by 60%, while the increase amounts

[17] The respective p-values in treatment **TP** are 0.40 for a regression with session random effects, and 0.35 for the Wilcoxon signed-rank test. In **SP**, we obtain $p < 0.001$ for the regression with session random effects and $p < 0.05$ for the Wilcoxon signed-rank test.

[18] Another difference to their experimental design is that they use a 'difference' contest success function rather than the 'ratio' technology we employ. For a theoretical comparison of these technologies, see Hirshleifer (1989).

[19] This difference is significantly different from zero at the 5%-level for a regression with session random effects, but only at the 10%-level for the Wilcoxon signed-rank test.

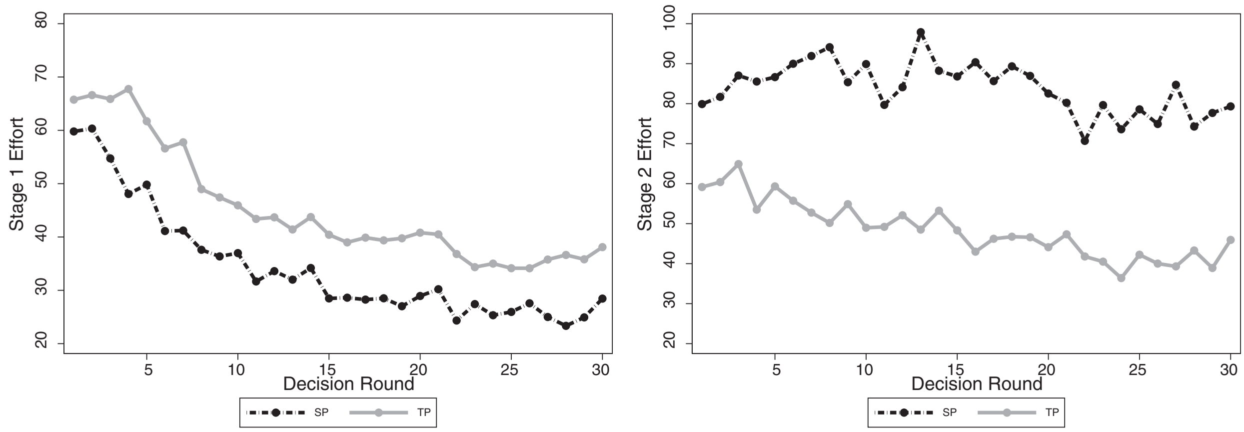

Fig. 5. Individual effort by stage, decision round, and treatment.

to little more than 30% in the experiment. Similarly, the runner-up prize should reduce stage-2 effort by 60%, but the reduction amounts to less than 45% in the experimental data. Thus, in the experimental data the runner-up prize works qualitatively as predicted by theory, but not in quantitative terms.

Summing up, we find some evidence for the trade-off between total effort maximization and incentive maintenance. Total effort seems to be somewhat higher in **SP** than in **TP** (Result 1), and incentive maintenance cannot be rejected in the **TP** treatment, but is rejected in the **SP** treatment (Result 2). However, the difference in total effort across treatments is statistically only weakly significant and economically less pronounced than theory predicts.

4.2. *Risk aversion and the runner-up prize*

The results in Section 2.3 suggests that a runner-up prize may increase (rather than decrease) total effort by risk-averse decision makers because such a prize provides insurance against the worst possible outcome in a two-stage elimination contest. Thus, risk attitudes of experimental subjects may be responsible for the finding of only moderate and insignificant differences in the total effort measures across treatments.

Hypotheses and risk aversion. Fig. 6 plots the differences between the various effort measures on which Hypotheses 1–3 are based as a function of the degree of relative risk aversion λ for the parameters used in the experiment. As in the numerical solution in Section 2.3, the curves in the figure are based on the assumption that all agents have the same CRRA preferences (characterized by λ) and that this is common knowledge. Consider first the effect of risk aversion on total effort: the locus labeled "Hypothesis 1" plots the difference between total effort in the two treatments. Since total effort is higher in

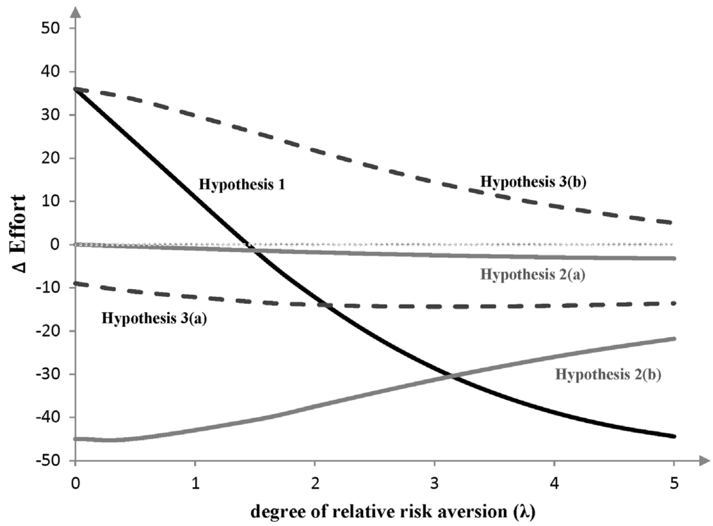

Fig. 6. The effect of risk aversion on effort differences. *Note*: The locus "Hypothesis 1" plots $\mathcal{E}(\mathbf{SP}) - \mathcal{E}(\mathbf{TP})$ as a function of λ; the loci "Hypotheses 2(a) and 2(b)" plot $x_1(\mathbf{TP}) - x_2(\mathbf{TP})$ and $x_1(\mathbf{SP}) - x_2(\mathbf{SP})$, respectively, as a function of λ; the loci "Hypothesis 3(a)" and "Hypothesis 3(b)" plot $x_1(\mathbf{SP}) - x_1(\mathbf{TP})$ and $x_2(\mathbf{SP}) - x_2(\mathbf{TP})$, respectively, as a function of λ.

Table 3
Testable hypotheses and risk-aversion.

Hypothesis ("Δ Effort")	1 $\mathcal{E}(SP) - \mathcal{E}(TP)$	2(a) $x_1(TP) - x_2(TP)$	2(b) $x_1(SP) - x_2(SP)$	3(a) $x_1(SP) - x_1(TP)$	3(b) $x_2(SP) - x_2(TP)$
Risk-neutral	33.486	0.911	−54.093	−11.037	43.967
Weakly RA	6.271	−1.309	−51.677	−16.794	33.575
Strongly RA	−22.767	−2.475	−46.546	−19.252	24.819

Note: Numbers denote averages over all rounds and individuals of the respective subgroup and are defined as in Table 2. Classification of subjects: 32 subjects are strongly risk-averse (switch to the safe option in row 6 or below); 33 are weakly risk-averse (switch to the safe option between rows 7 and 10); and 81 are risk-neutral (switch in row 11 or 12). Theoretical Prediction: Fig. 6 provides the effort differences across treatments or stages on which the hypotheses are based as a function of the degree of relative risk aversion λ. The expressions in line "Δ Effort" correspond to the loci plotted in Fig. 6.

SP than in **TP** under risk-neutrality, while the reverse relation holds for sufficiently risk-averse subjects, the treatment effect is predicted to be positive for risk-neutral subjects ($\lambda = 0$) and negative for sufficiently risk-averse ones. The locus labeled "Hypothesis 2(a)" plots the difference between stage-1 and stage-2 effort choices in treatment **TP**. It reveals that incentives are exactly maintained under risk-neutrality, while the difference between effort across stages becomes slightly negative as the degree of risk aversion increases; the effect of risk aversion on Hypothesis 2(a) is fairly weak, however, in particular relative the corresponding effect on Hypothesis 1.[20] The locus labeled "Hypothesis 2(b)" plots the difference between stage-1 and stage-2 effort choices in **SP** and reveals that incentives are not maintained in this treatment. However, the difference between stage-1 and stage-2 effort choices becomes slightly less negative as the degree of risk aversion increases. Finally, the figure plots the difference between stage-1 and stage-2 effort *across* treatments, referring to Hypotheses 3(a) and 3(b): the difference between stage-1 effort choices across treatments remains negative throughout and is predicted to increase with the degree of risk aversion in absolute terms; the difference between stage-2 effort choices across treatments is positive throughout and decreasing in the degree of risk aversion.

Results. Overall, the choices of 174 participants exhibit a unique switching point in the risk-preference elicitation procedure, while the remaining 26 subjects (13 in each treatment) have multiple switching points. In the sequel we restrict attention to participants with choices exhibiting a unique switching point.

Table 3 provides the differences between various effort measures on which our hypotheses are based for three different subgroups of experimental subjects: for risk neutral, for weakly risk averse, and for strongly risk averse subjects.[21]

The experimental data is qualitatively in line with the predictions derived from the CRRA solution: regarding Hypothesis 1, total effort is significantly higher in **SP** than in **TP** for risk-neutral subjects (p-value <0.05).[22] We cannot reject that total effort is the same in both treatments for any one of the two classes of risk-averse subjects; the numbers even suggest that total effort provision by strongly risk-averse subjects is higher in the treatment with a positive runner-up prize, as predicted by the numerical simulation. Regarding Hypotheses 2(a) and 2(b), we cannot reject incentive maintenance for any one of the three subgroups in **TP**, while incentive maintenance can be rejected for all subgroups in **SP**.[23] In line the with qualitative prediction from the numerical solution, however, the exact numbers in Table 3 suggest that stage-1 effort is (slightly) lower than stage-2 effort for risk-averse subjects in **TP**, and that the (absolute) difference between stage-1 and stage-2 effort choices in **SP** is decreasing in the degree of risk aversion. Finally, the results in Table 3 indicate that the qualitative effects of risk aversion on effort on the two stages are consistent with the predictions from Hypotheses 3(a) and 3(b). The negative effect of the runner-up prize on stage-1 effort is increasing in the degree of risk aversion in absolute terms. It is most pronounced and significantly different from zero for strongly risk-averse subjects (p-value 0.05), while the effect is smaller and statistically insignificant for the remaining two subgroups (p-values are 0.34 and 0.61, respectively). In contrast, the positive effect of the runner-up prize on stage-2 effort is most pronounced for risk-neutral subjects and decreasing in the degree of risk aversion; the effect is significant for risk-neutral and weakly risk-averse subjects (p-value < 0.01), but not for strongly risk-averse ones (p-value 0.43).

4.3. Discussion

Risk aversion and alternative explanations. The separate analysis of effort decisions by risk-neutral and risk-averse subjects suggests that the experimentally observed deviations from the risk-neutral benchmark hypotheses might indeed be driven by deviations from risk neutrality. While there is a clear trade-off between total effort maximization and incentive

[20] The only reason why risk aversion affects this hypothesis is the riskiness of the continuation value for which subjects compete in stage 1: the value of a participation in stage 2 is a lottery and therefore slightly less valuable for risk-averse subjects.

[21] Since a distinction between 'strongly' and 'weakly' risk-averse agents is somewhat arbitrary, we construct two groups of approximately equal size. The qualitative relation with respect to the information provided across subgroups is not affected if we increase or reduce the threshold, however. Details available upon request.

[22] This statement is based on a panel regression with robust standard errors and random effects on the individual level. Since the number of experimental subjects who participated in a particular session differs across subgroups, and average risk attitudes differ across sessions, we add session dummies for inference in this section.

[23] The respective p-values in treatment **TP** are 0.57 for risk-neutral, 0.81 for weakly risk-averse, and 0.70 for strongly risk-averse subjects; the respective p-value in treatment **SP** is below 0.01 for all subgroups.

maintenance for risk-neutral subjects, we do not find evidence for this trade-off for risk-averse subjects. Instead, the runner-up prize tends to kill two birds with one stone for this latter group of subjects: incentives are maintained across stages, and total effort is not lower. The numbers even suggest that total effort might be higher in a setting with runner-up prize. The experimental data deviate from predictions of the model in terms of the quantitative predictions about effort, however. According to the model, effort by risk-averse subjects is expected to be lower than effort by risk-neutral ones, independent of the treatment. While previous findings from contest experiments are broadly in line with this prediction (Millner and Pratt, 1991; Sheremeta, 2011), this is not what we observe. Instead, we find that risk-averse subjects provide slightly more effort than risk-neutral ones in both treatments, even though the difference is not statistically significant.[24] In contrast to these earlier studies, however, we consider a dynamic contest. Moreover, the pattern we observe is not necessarily surprising. First, the strategic uncertainty about behavior of the opponent that agents face due to random rematching introduces additional risk to which agents might respond differently.[25] Second, it was theoretically shown by Konrad and Schlesinger (1997) that the level effect of risk aversion on effort provision is indeterminate, even in a framework with homogeneous agents and no strategic uncertainty: on the one hand, risk-averse agents tend to invest less than risk-neutral ones, since effort provision in any contest can be seen as a risky investment with state dependent payoffs; on the other hand, risk-averse agents might invest more than risk-neutral ones to reduce the probability that the unfavorable outcome realizes; which effect dominates depends on the shape of the utility function.[26]

An alternative explanation for the data pattern displayed in Table 3 might be that loss aversion rather than risk aversion is responsible for the differences in the effort decisions of subjects in the three classes. Arguably more loss-averse agents make more 'conservative' choices in our risk-preference elicitation tasks and might therefore appear to be more risk averse. This story might yield an alternative explanation for the observation that those subjects that we have identified as risk-averse do not invest less than those subjects classified as risk-neutral. However, it turns out that this alternative story does not yield a plausible explanation for the observed pattern across treatments. Intuitively, a runner-up prize reduces the disutility of losing for loss-averse agents, and thus the incentives to avoid this situation by additional contest investments, leading to lower effort.[27] Thus, loss aversion makes counterfactual predictions, while the results for subgroups with different risk attitudes are qualitatively well in line with most predictions of the respective model. We therefore conclude that it is implausible that loss aversion explains the qualitative patterns in the experimental data, even though it is not possible to disentangle risk aversion from loss aversion with our data. However, it cannot be ruled out at this stage that other behavioral traits which are correlated with the concavity of the utility function affect our results. For instance, recent experimental evidence by Eriksen and Kvaloy (2012) that a more frequent evaluation of outcomes in tournaments induces higher effort and more risk-taking is consistent with the evidence in our experiments, where effort is higher in contests with a runner-up prize.

Over-provision of effort. As mentioned at the beginning of this section, we observe a substantial amount of effort over-provision relative to the risk-neutral benchmark prediction, with total effort in the experimental session being between 70% and 90% higher than predicted. This finding complements earlier evidence on over-provision in contest experiments – see Davis and Reilly (1998), Gneezy and Smorodinsky (2006), or Sheremeta (2010a), for instance.[28] Several explanations have been put forward in the literature to explain this phenomenon.[29] First, the endowment that experimental subjects receive at the beginning of each decision round may lead to over-provision if subjects perceive the endowment as 'play money' (Thaler and Johnson, 1990). In this case, subjects provide more effort due to this perception than they would without an endowment. Price and Sheremeta (2011) formally test the endowment effect hypothesis and find supporting evidence.[30] Moreover, observed effort choices in experiments without endowments are often much closer to the risk-neutral benchmark prediction (Altmann et al., 2012). In our experiments, we explicitly decided to use endowments to avoid the possibility of negative payoffs for the losers of a contest and the associated problem of limited liability in the experiment. Arguably, we could also have solved this issue through additional prizes for the losers, as in Altmann et al. (2012). Then, however, the contrast between a single- and a two-prizes treatment, which is central for our research question, would have been less clear-cut. A related explanation for over-provision – by Potters et al. (1998) – is that experimental subjects are prone to make mistakes in experimental settings. If this is the case, a higher endowment increases the chance to make mistakes. Sheremeta (2011) varies the endowment and finds evidence that is in line with this argument.[31] Finally, subjects may experience a 'joy of winning' in strategic interactions which amplifies the valuation of prizes awarded in contests. Since individual efforts are

[24] We regress the total effort variable on risk attitudes for the whole sample, and separately for each treatment. The p-value of the estimated coefficient on risk attitudes is above 0.10 in all specifications.

[25] The literature has only recently started to consider contest games where players have private information about their type. Testable predictions are hard to derive, however, since results crucially depend on beliefs (Ewerhart, 2010).

[26] Our elicitation procedure for risk attitudes does not allow for a separation of these two effects since it only delivers a measure of concavity of the utility of an experimental subject under the assumption of a given family of the utility function.

[27] A formal analysis is available from the authors upon request.

[28] Sheremeta (2010a) reports similar degrees of over-provision. In a treatment which is almost identical to **SP**, he finds that total effort is on average almost 90% higher than theory based on risk-neutrality predicts.

[29] See Sheremeta (2013) for an excellent survey of the literature addressing this issue.

[30] See also Price and Sheremeta (2014) for a more extensive discussion and potential explanations for the endowment effect.

[31] Sheremeta (2011) uses the concept of a quantal response equilibrium (QRE) by McKelvey and Palfrey (1995), which allows for mistakes of decision makers. He finds that a reduction of the endowment causes a proportional reduction of total effort, even if the endowment is not binding for equilibrium effort levels. Low (though non-binding) endowments even lead to under-provision of effort.

strictly increasing in the prizes at stake, non-monetary values of winning can rationalize over-provision of effort. Sheremeta (2010a) experimentally elicits a measure for the 'joy of winning' and finds that it is highly correlated with the amount of effort provided by individual subjects. This supports the hypothesis that the 'joy of winning' is at least partly responsible for over-provision.

It is important to note that these potential explanations for over-provision are unlikely to affect behavior differently in the two treatments contrasted here. This clearly holds for the 'play money' and the 'mistakes' effect, since both the endowment and the overall amount available for prizes are identical in the two treatments.[32] A 'joy of winning' that is independent of the prize money at stake has the potential to affect our hypothesis. It should disproportionably increase stage-1 effort in **SP**, since this is the only instance where winning does not lead to a prize. This is not what we observe, though.

5. Conclusion

This paper has investigated the impact of variations in the prize structure on effort decisions in dynamic contests. When comparing contests with two different prize-structures, a "winner-takes-all" contest with a single prize and a contest that also includes a positive runner-up prize, the standard model with risk-neutral participants predicts that total effort is higher in the single prize contest, while a positive runner-up prize is required for the maintenance of incentives (and efforts) across the different stages of the contest. Thus, in the risk-neutral benchmark, there is a trade-off for a contest designer who tries to achieve both goals simultaneously. The trade-off becomes less pronounced, however, when considering risk-averse contest participants. The intuitive reason for this finding is that prizes on intermediate stages of the contest reduce the riskiness of effort investments by providing insurance against being eliminated before prevailing in the final stage. As a consequence, effort in early stages is predicted to be higher in a contest with a runner-up prize.

Evidence from laboratory experiments is broadly in line with these predictions. In particular, total effort is higher in a single-prize treatment than in a treatment with a runner-up prize. Also, there are pronounced differences in the efforts on the different stages in the single-prize treatment, while effort is virtually identical across stages in the treatment with a runner-up prize. Closer inspection also reveals systematic differences in behavior depending on risk attitudes. When restricting attention to risk-neutral subjects, the evidence is consistent with a trade-off for a contest designer. The trade-off disappears for risk-averse participants where the contest with a runner-up prize produces both higher total effort as well as incentive maintenance across stages. Overall, our findings suggest that a systematic investigation of the role of risk attitudes for behavior in dynamic contests might be fruitful direction for future research.

Appendix A. Extensions of the simple dynamic elimination contest

A.1. Generalized Tullock contest technology

Assume that the probability for agent i in a pairwise interaction between i and j is defined as

$$p_{si}(x_{si}, x_{sj}) = \begin{cases} \dfrac{x_{si}^r}{x_{si}^r + x_{sj}^r} & \text{if} \quad x_{si} + x_{sj} > 0 \\ \dfrac{1}{2} & \text{if} \quad x_{si} + x_{sj} = 0. \end{cases}$$

Thus, the optimization problem in stage 2 reads

$$\max_{x_{2i} \geq 0} \Pi_{2i}(x_{2i}, x_{2j}) = \frac{x_{2i}^r}{x_{2i}^r + x_{2j}^r} P^H + \left(1 - \frac{x_{2i}^r}{x_{2i}^r + x_{2j}^r}\right) P^L - x_{2i} = \frac{x_{2i}^r}{x_{2i}^r + x_{2j}^r}(P^H - P^L) + P^L - x_{2i}.$$

The first-order condition is both necessary and sufficient if $0 < r < 2$ (Perez-Castrillo and Verdier, 1992). Using symmetry delivers equilibrium efforts

$$x_2^* := x_{2i}^* = x_{2j}^* = \frac{r(P^H - P^L)}{4}. \tag{6}$$

Inserting equilibrium efforts in the objective functions gives the expected stage-2 equilibrium payoff

$$\Pi_2^* \equiv \Pi_{2i}(x_{2i}^*, x_{2j}^*) = \Pi_{2j}(x_{2i}^*, x_{2j}^*) = \frac{(2-r)P^H + (2+r)P^L}{4}. \tag{7}$$

[32] The explanation based on errors would only be an issue if, e.g., the endowment were to bind more often in one than in the other treatment. While the share of experimental subjects who spend their entire endowment is slightly higher in **SP** than in **TP** (5.6% vs. 3.5%), the numbers are low in both treatments. Moreover, all qualitative findings remain unchanged if observations for which the endowment is binding are excluded. Details are available upon request.

Reaching stage 2 has value Π_2^* for an agent that participates in stage 1. Agent k will take this value into account when choosing her stage-1 effort x_{1k}. Since all agents are identical by assumption, consider the interaction between two arbitrary agents k and l in stage 1. Agent k faces the optimization problem

$$\max_{x_{1k}\geq 0}\Pi_{1k}(x_{1k},x_{1l})=\frac{x_{1k}^r}{x_{1k}^r+x_{1l}^r}\Pi_2^*-x_{1k}$$

As in the solution of stage 2 above, the first-order condition together with symmetry yields the equilibrium efforts on stage 1 as

$$x_1^*:=x_{1k}^*=x_{1l}^*=r\frac{(2-r)P^H+(2+r)P^L}{16}. \tag{8}$$

Under the assumption that $P=P^H+P^L$, (6) and (8) jointly determine total effort as

$$\mathcal{E}=r\frac{(4-r)P+(2r-4)P^L}{4}.$$

Thus, total effort is decreasing in the runner-up prize P^L in any interior equilibrium ($0<r<2$). Incentive maintenance across stages requires the equality of (6) and (8). Using $P=P^H+P^L$, it must hold that

$$r\frac{P-2P^L}{4}=r\frac{(2-r)P+2rP^L}{16}\Leftrightarrow P^L=\frac{2+r}{8+2r}P>0.$$

A.2. More than two stages

Consider the risk-neutral benchmark model discussed in the main text of the paper, but assume that $Y=2^N$ rather than just 4 contestants participate in a dynamic elimination contest with prizes P^H and P^L ($P=P^H+P^L$). In this case, equilibrium effort choices in the last two stages remain unchanged. Agents in the third-to-last stage of the contest compete in pairwise interactions for a continuation value

$$\Pi_1^*=\frac{P+2P^L}{16}.$$

Now, we can use the insight by Gradstein and Konrad (1999) that total effort provision by Y players in a sequential elimination contest equals the respective effort measure in a grand contest; total effort amount to $((Y-1)/Y)R$ in both cases, where R is the contested prize. In our setting, the contested prize equals Π_1^*, and we must consider four interactions between $Z=Y/4$ contestants in the third-to-last stage. Thus, total effort amounts to

$$\mathcal{E}=\frac{3P-2P^L}{4}+4\frac{Z-1}{Z}\frac{P+2P^L}{16}=\frac{(Y-1)P-2P^L}{Y}.$$

As in the four player case, total effort is strictly decreasing in the runner-up prize P^L, i.e., total effort is maximized in a "winner-takes-all" contest. It was already shown by Rosen (1986) that one positive runner-up prize is necessary for incentive maintenance across stages in each pairwise interaction of a sequential elimination contest. Thus, the trade-off we observe does not depend on the contest size.

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jebo.2014.02.018.

References

Altmann, S., Falk, A., Wibral, M., 2012. Promotions and incentives: the case of multi-stage elimination tournaments. J. Labor Econ. 30 (1), 149–174.

Amaldoss, W., Rapoport, A., 2009. Excessive expenditure in two-stage contests: theory and experimental evidence. In: Columbus, F. (Ed.), Game Theory: Strategies, Equilibria, and Theorems. Nova Science Publishers, Hauppauge, NY.

Baye, Michael, R., Kovenock, D., de Vries, C.G., 1994. The solution to the Tullock rent-seeking game when R>2: mixed-strategy equilibria and mean dissipation rates. Public Choice 81, 363–380.

Bonin, H., Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2007. Cross-sectional earnings risk and occupational sorting: the role of risk attitudes. Labour Econ. 14, 926–937.

Bull, C., Schotter, A., Weigelt, K., 1987. Tournaments and piece rates: an experimental study. J. Polit. Econ. 95, 1–33.

Davis, D., Reilly, R., 1998. Do many cooks always spoil the stew? An experimental analysis of rent-seeking and the role of a strategic buyer. Public Choice 95, 89–115.

Delfgaauw, J., Dur, R., Non, A., Verbeke, W., 2014. The effects of prize spread and noise in elimination tournaments: a natural field experiment. J. Labor Econ. (forthcoming).

Dohmen, T., Falk, A., 2011. Performance pay and multidimensional sorting: productivity, preferences, and gender. Am. Econ. Rev. 101, 556–590.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2010. Are risk aversion and impatience related to cognitive ability? Am. Econ. Rev. 100 (3), 1238–1260.

Eriksen, K., Kvaloy, O., 2012. Myopic Risk Taking in Tournaments, UiS Working Papers in Economics and Finance, 2012/13.

Ewerhart, C., 2010. Rent-Seeking Contests with Independent Private Values., University of Zurich Working Paper 490.

Fischbacher, U., 2007. Z-Tree, Zurich toolbox for readymade economic experiments. Exp. Econ. 10, 171–178.
Fu, Q., Lu, J., 2012. The optimal multi-stage contest. Econ. Theory 51 (2), 351–382.
Gneezy, U., Smorodinsky, R., 2006. All-pay auctions – an experimental study. J. Econ. Behav. Organ., 61 255–275.
Gradstein, M., Konrad, K.A., 1999. Orchestrating rent seeking contests. Econ. J. 109, 536–545.
Greiner, B., 2004. An online recruitment system for economic experiments. In: Kremer, K., Macho, V. (Eds.), Forschung und wissenschaftliches Rechnen 2003. GWDDG Bericht der 63. Gesellschaft für Wissenschaftliche Datenverarbeitung, Göttingen, pp. 79–93.
Harbring, C., Irlenbusch, B., 2003. An experimental study on tournament design. Labour Econ. 10 (4), 443–464.
Harbring, C., Lünser, G.K., 2008. On the competition of asymmetric agents. German Econ. Rev. 9, 373–395.
Hirshleifer, J., 1989. Conflict and rent-seeking success functions: ratio vs. difference models of relative success. Public Choice 63, 101–112.
Holt, C.A., Laury, S.K., 2002. Risk aversion and incentive effects. Am. Econ. Rev. 92, 1644–1655.
Konrad, K., 2009. Strategy and Dynamics in Contests. Oxford University Press, Oxford.
Konrad, K.A., Schlesinger, H., 1997. Risk aversion in rent-seeking and rent-augmenting games. Econ. J. 107, 1671–1683.
Krishna, V., Morgan, J., 1998. The winner-take-all principle in small tournaments. Adv. Appl. Microecon. 7, 61–74.
Lazear, E.P., Rosen, S., 1981. Rank-order tournaments as optimal labor contracts. J. Polit. Econ. 89, 841–864.
McKelvey, R., Palfrey, T., 1995. Quantal response equilibria for normal form games. Games Econ. Behav. 10, 6–38.
Millner, E.L., Pratt, M.D., 1991. Risk aversion and rent-seeking: an extension and some experimental evidence. Public Choice 69, 81–92.
Parco, J.E., Rapoport, A., Amaldoss, W., 2005. Two-stage contests with budget constraints: an experimental study. J. Math. Psychol. 49, 320–338.
Perez-Castrillo, J.D., Verdier, T., 1992. A general analysis of rent-seeking games. Public Choice 73, 335–350.
Pollmann, D., Dohmen, T., Palm, F., 2012. Robust Estimation of Wage Dispersion with Censored Data: An Application to Occupational Earnings Risk and Risk Attitudes, IZA Discussion Paper, 6447.
Potters, J., de Vries, C.G., van Winden, F., 1998. An experimental examination of rational rent-seeking. European J. Polit. Econ. 14, 783–800.
Price, C.R., Sheremeta, R.M., 2011. Endowment effects in contests. Econ. Lett. 111, 217–219.
Price, C.R., Sheremeta, R.M., 2014. Endowment origin, demographic effects and individual preferences in contests. J. Econ. Manage. Strategy (forthcoming).
Rosen, S., 1986. Prizes and incentives in elimination tournaments. Am. Econ. Rev. 76 (4), 701–715.
Sheremeta, R.M., 2010a. Experimental comparison of multi-stage and one-stage contests. Games Econ. Behav. 68, 731–747.
Sheremeta, R.M., 2010b. Expenditures and information disclosure in two-stage political contests. J. Confl. Resolut. 54, 771–798.
Sheremeta, R.M., 2011. Contest design: an experimental investigation. Econ. Inq. 49, 573–590.
Sheremeta, R.M., 2013. Overbidding and heterogeneous behavior in contest experiments. J. Econ. Surv. 27, 491–514.
Sisak, D., 2009. Multiple prize contests – the optimal allocation of prizes. J. Econ. Surv. 23 (1), 82–114.
Thaler, R.H., Johnson, E.C., 1990. Gambling with the house money and trying to break even: the effects of prior outcomes on risky choice. Manage. Sci. 36 (6), 643–660.
Tullock, G., 1980. Efficient rent-seeking. In: Buchanan, J.M., Tollison, R.D., Tullock, G. (Eds.), Toward a Theory of the Rent-Seeking Society. Texas A&M Press, College Station, pp. 97–112.