Workaholics and Drop Outs in Optimal Organizations*

Wieland Müller† Andrew Schotter‡
Tilburg University *New York University*

September 7, 2005

Abstract

This paper reports the results of experiments designed to test the theory of the optimal composition of prizes in contests. We find that while in the aggregate the behavior of our subjects is consistent with that predicted by the theory, such aggregate results mask an unexpected compositional effect on the individual level. While theory predicts that subject efforts are continuous and increasing functions of ability, the actual efforts of our laboratory subjects bifurcate. Low ability workers drop out and exert little or no effort while high ability subjects try too hard. This discontinuity, which is masked by aggregation, has significant consequences for behavior in organizations.

Keywords: contests, all-pay auctions, experiments.

JEL classification numbers: C92, D44, J31, D72, D82.

*We wish to thank Maher Said for valuable research assistance and seminar participants at Humboldt University for helpful comments. The first author acknowledges financial support from the German Science Foundation, DFG. Furthermore he thanks the Center for Experimental Social Science (CESS) at New York University for its hospitality.

†Department of Economics, Tilburg University, Warandelaan 2, Postbus 90153, 5000 LE Tilburg, The Netherlands, Fax: +31 13 466 3042, Email: *w.mueller@uvt.nl*.

‡Department of Economics, New York University, 269 Mercer Street, New York, N.Y. 10003-6687, USA, Fax 1 212 995 3932, Email: *andrew.schotter@nyu.edu*.

1 Introduction

Casual empiricism indicates that many organizations are characterized by a bifurcation of effort among workers. While one subset appears to not be able to stop themselves from working (the fast track) the other seem alienated and exerts no effort at all. One immediate reaction to this stylized fact is to assume that the incentive structure underlying these work patterns is sub-optimal and that if the organizer of the firm could only redesign the pay or promotion structure in the firm, such dichotomous behavior could be eliminated.[1]

In this paper we suggest that in hierarchic organizations where promotion possibilities are limited, such an effort bifurcation is the natural, if not the equilibrium, behavioral response to an optimally designed incentive structure. In other words, the problem in these organizations is not that the designer has failed to set the organizational prizes correctly but rather, given the optimal prize (promotion) structure, talented workers seem to enter into a rat race for promotions with each other while less talented ones, see the writing on the wall, and drop out knowing that their chances for promotion are limited.[2,3]

To illustrate our point we experimentally test a model proposed by Moldovanu and

[1] One interesting example of a poorly designed incentive structure is discussed by Rafael Tenorio (2000). In this paper it is argued that the compensation scheme used in professional boxing according to which a boxer's payment or purse for a given fight is entirely guaranteed, provide sub-optimal incentives which may (and sometimes does) result in improper preparation for the fight and, therefore, in an increased likelihood of a poor showing.

[2] The phenomena of workcaholism and dropping out is sometimes observed among junior faculty. While some hot prospects work hard (maybe even to the extent of getting burnt out, Amegashie, 2003) to get tenure, others see the writing on the wall and drop out of academia for fear not to be able to meet promotion standards despite high investments. The same may apply to young lawyers seeking partnerships in high-powered law firms and management consultants (see e.g. Akerlof, 1976). Prendergast (1999) suggests that such drop-out behavior can be seen in sports contests. Furthermore, if there are asymmetric budget constraints of parties involved in a legal dispute (which can be interpreted as a contest), the party having the higher budget can hire a better lawyer and can therefore increase its chances of winning. If this is realized by the other party, it might give up (drop out) immediately. Moreover, drop-out behavior in tournaments has been noted before by Schotter and Weigelt (1992) for workers with relatively high costs of effort. In that paper affirmative action laws were suggested as a policy intervention that could be used to rectify the situation.

[3] While in the model analyzed in this study dropping out is not part of the game's equilibrium, in other models it very well may be. For example, this is demonstrated in Benoit (1999). In his model, members of socioeconomically disadvantaged groups and members of other groups have to decide—after learning about their ability—whether or not to invest, say, in schooling and then take a test. Benoit finds that if there is no affirmative action, members of the disadvantaged group might not invest (i.e. drop out).

Sela (2001) entitled "Optimal Allocation of Prizes in Contests" (henceforth M-S). In this paper M-S derive the optimal set of prizes for an organization involved in motivating workers through an effort tournament. They investigate firms where workers have either linear, convex or concave cost-of-effort functions and where an organizational designer has a limited amount of money available for bonuses to be awarded to those workers whose outputs are highest. (Assume that output is linear in effort and non-stochastic so in essence effort is equivalent to output and both are observable). They demonstrate that for organizations where workers have linear or concave cost-of-effort functions, the optimal prize structure is one where the entire prize budget is allocated to one big prize while if costs are convex, it might be optimal to distribute the budget amongst several prizes. What is interesting is that in these contests the equilibrium effort functions are continuous functions of the abilities of the workers while in the lab we observe individual effort functions which appear to be discontinuous step functions where low ability workers drop out and exert zero effort while high ability workers over exert themselves leading to the bifurcation of efforts described above.

The ironic aspect of our experimental results is that despite this bifurcation of effort, on average the prize structures proposed by M-S elicit approximately the correct effort levels so that with respect to the mean one could say they work. Even more interesting is the fact that when we aggregate our data across laboratory work groups, efforts appear to be continuous so that the observed bifurcation of efforts is hard to detect on the aggregate level.

To illustrate this point, say a corporation has many plants. In addition, say that the corporation's headquarter sets the same incentive structure up in each plant. If one could observe the efforts and abilities of the workers in all of these plants (as we can in the lab) and aggregate them, it would appear that behavior is very consistent with the M-S theory (i.e., it would appear that the aggregate effort function had the right shape, was approximately continuous, and exhibited a mean effort level that was approximately equal to that predicted by the theory) while behavior on the individual and plant level would tell a very different story. Our discussion here concerns itself with higher moments of the distribution of effort on the plant level. While economists might feel that a risk neutral firm might only care about the mean effort levels of its workers and not the higher moments, we suspect that psychologists would be the first to indicate that an incentive system that creates an organization composed of a set of alienated drop-outs and a set of workaholics is bound, in the long run to be dysfunctional.

Our results have consequences far beyond those associated with eliciting efforts in a work organization. For example, let us say that colleges have money available for scholarships and want to distribute these funds amongst its applicants. In a society where some students are disadvantaged our experimental results would indicate that such students would drop out and exert no or suboptimal amount of effort in the competition for these prizes while others would enter a rat race. In fact, this phenomenon is what has worried so many people interested in education policy in the United States. The children of the privileged are obsessed with college admissions and eagerly enter the rat-race associated with it, while the under-privileged drop out.[4]

In this paper we will proceed as follows: In the next section we will present the M-S model and its results. In Section 3 we will describe our experimental design while in Section 4 we will present our results. Finally, in Section 5 we will offer some conclusions and discussion.

2 The Moldovanu-Sela Theory

2.1 Model Specification

In this section we lay out the model underlying our experiments and its predictions. In doing so we confine ourselves to the special cases relevant for our experiments. For more general results see Moldovanu and Sela (2001).

Assume that there exists an organization with $k \geq 3$ contestants competing in a contest in which two prizes can be awarded. The (commonly known) values of the prizes are $V_1 \geq V_2 \geq 0$ with $V_1 + V_2 = 1$. In the contest players simultaneously exert effort x_i thereby incurring cost $c_i\gamma(x_i)$. The function γ: $\mathbb{R}_+ \rightarrow \mathbb{R}_+$ is strictly increasing with $\gamma(0) = 0$ and $c_i > 0$ is an ability parameter. Notice that the lower ability c_i the more able is player (i.e., the lower is his or her costs) and vice versa.

It is assumed that the ability of player i is private information to i. Abilities are independently drawn from the interval $[m, 1]$, $m > 0$, according to the (commonly known) distribution function F with $F' > 0$. The contestant with the highest effort wins the prize V_1, the contestant with the second highest effort wins prize V_2 whereas all other contestants win nothing. Accordingly, the payoff of contestant i who has ability c_i and exerts effort x_i is either $V_j - c_i\gamma(x_i)$ if i wins prize j, or $-c_i\gamma(x_i)$ if i does not win a prize. Note, then, that this contest defines an all-pay auction where

[4] Drop-out behavior is not only observed in education but also in employment, and contracting opportunities. To work against this phenomenon, many countries, firms and universities implement affirmative action programs. For an overview of such programs in the US and their assessment see Holzer and Neumark (2000) as well as National Institute of Government Purchasing (1994). For an experimental look at this topic see Corns and Schotter (1999).

bidders make effort bids and pay the cost associated with their decision numbers whether they win or not. The contest designer determines the number of prizes and how to allocate the prize sum among the prizes in order to maximize the expected value of the sum of the efforts $\sum_{i=1}^{k} x_i$ given the contestants' equilibrium-effort functions.

Assuming that all contestants other than i make an effort according to the function b and assuming that this function is strictly monotonic and differentiable, player i's maximization problem is:

$$\max_{x} \left[V_1(1 - F(b^{-1}(x)))^{k-1} + V_2(k-1)F(b^{-1}(x))(1 - F(b^{-1}(x)))^{k-2} - c\gamma(x) \right]. \qquad (1)$$

Here the factor after V_1 is the probability that x is the highest among all efforts and the factor after V_2 is the probability that x is the second highest among all efforts.

In the experiments we chose $k = 4$, $m = 0.5$ and a uniform distribution of abilities, i.e., $F(c) = 2c - 1$, $c \in [0.5, 1]$.

2.2 Predictions and Prescriptions[5]

Linear cost functions: In case all contestants have linear costs, i.e., $\gamma(x) = x$ the optimal and symmetric effort function can be shown to be

$$b(c) = V_1 A(c) + V_2 B(c) \qquad (2)$$

with

$$A(c) = -36 + 48c - 12c^2 - 24 \ln c \quad \text{and} \quad B(c) = 84 - 120c + 36c^2 + 48 \ln c. \qquad (3)$$

Turning to the designer's problem, let $V_2 = \alpha$ and $V_1 = 1 - \alpha$, where $0 \leq \alpha \leq 1/2$ such that the second prize is smaller than the first. A contestant's equilibrium effort is therefore given by $b(c) = (1-\alpha)A(c) + \alpha B(c) = A(c) + \alpha\,(B(c) - A(c))$. Since each contestant's average effort is given by $\int_{0.5}^{1} [A(c) + \alpha\,(B(c) - A(c))]\, F'(c)dc$, the designer's problem reads

$$\max_{0 \leq \alpha \leq 1/2} 4 \int_{0.5}^{1} [A(c) + \alpha\,(B(c) - A(c))]\, F'(c)dc$$

or, equivalently,

$$\max_{0 \leq \alpha \leq 1/2} \alpha \int_{0.5}^{1} [B(c) - A(c)]\, F'(c)dc.$$

[5] See Moldovanu and Sela (2001) for a full derivation of these results.

Since the definite integral in this case is negative, the solution to this problem is $\alpha = 0$ such that it is optimal for the designer to award only one prize, i.e., $V_1 = 1$ and $V_2 = 0$.

Quadratic cost functions: In case all contestants have quadratic costs, i.e., $\gamma(x) = x^2$, the optimal and symmetric effort function can be shown to be

$$b(c) = \gamma^{-1}(V_1 A(c) + V_2 B(c)) = \sqrt{V_1 A(c) + V_2 B(c)} \tag{4}$$

where $A(c)$ and $B(c)$ are defined as in (3).

The designer's problem in this case reads

$$\max_{0 \leq \alpha \leq 1/2} 4 \int_{0.5}^{1} \gamma^{-1} \left(A(c) + \alpha \left(B(c) - A(c) \right) \right) F'(c) dc$$

and it turns out that in this case it is optimal to award two equal prizes, i.e. $V_1 = V_2 = 0.5$.

Hence, the prescriptions of the model are clear. When costs are linear the optimal prize structure is one where all the budget available for prizes in the organization are lumped together into one grand prize while when costs are quadratic two equally valuable prizes define the optimal prize structure.

3 Experimental Design and Procedures

In the experiments we rely on a classic 2-by-2 design: We implemented contests with either linear or quadratic costs and combine them with two different compositions of prizes, one that is optimal for that prize structure and one that is not. To be more precise, in treatment LC-1 all subjects have linear costs and there is only one positive-valued prize: $V_1 = 1$, $V_2 = 0$. As we have seen above, this prize composition is optimal from the designer's perspective if contestants have linear costs. In treatment QC-2 all subjects have quadratic costs and there are two equal prizes: $V_1 = 0.5$, $V_2 = 0.5$. This prize composition is optimal from the designer's perspective if contestants have quadratic costs. In treatment LC-2 all contestants have linear costs and the composition of prizes is the one that is optimal in the quadratic case. Finally, in treatment QC-1 all contestants have quadratic costs and the composition of prizes is the one that is optimal in the linear case. A summary of our four treatments is shown in Table 1.

The computerized[6] experiments were conducted in the experimental laboratory of the Eco-

[6] We used the software tool kit *z-Tree*, developed by Fischbacher (1999).

Treatment	Description	#Subjects	Period Endowm.	Max. Effort	#Periods
LC-1	linear costs	$6 \times 4 = 24$	0.22	1.96	50
	$V_1 = 1$, $V_2 = 0$				
LC-2	linear costs	$6 \times 4 = 24$	0.20	0.82	50
	$V_1 = V_2 = 0.5$				
QC-1	quadratic costs	$5 \times 4 = 20$	0.22	1.53	50
	$V_1 = 1$, $V_2 = 0$				
QC-2	quadratic costs	$5 \times 4 = 20$	0.20	0.99	50
	$V_1 = V_2 = 0.5$				

Table 1: Treatments.

nomics Department at New York University and the Center for Experimental Social Science. In each session fixed groups of four subjects were repeatedly matched to participate in a contest. Each of the experiments consisted of 50 periods. Payoffs were denoted in "points". At the beginning of each period each subject was assigned a "random number" indicating their type or ability, c_i. Each random number was an iid draw from the set of numbers {0.5, 0.51, ..., 1.00}. After subjects were informed about their individual random numbers, they simultaneously submitted "decision numbers". The set of admissible decision numbers was {0.01, 0.02, ..., *Maxeffort*} where *Maxeffort* was a number that was 20 per cent higher than the optimal effort of a contestant with ability $c = 0.5$ (the "best" ability possible) in a given treatment. In treatment LC-1, LC-2, QC-1, and QC-2 this number was respectively 1.96, 0.82, 1.53, and 0.99.[7] Subjects were informed that by choosing a decision number they would incur "decision costs." The form of the costs (depending on the treatment) was explained both verbally and in the form of a "decision cost calculator" that was accessible in each round. When fed with a trial decision number it showed the associated costs given the subject's random number in the current period. We implemented this cost calculator to help to avoid a bias due to the subjects' (possibly) limited computational capabilities.

After each member of a group had entered his or her decision number, the computer compared all of the decision numbers of the four members of a group. In one-prize contests, the player with the highest decision number received a "fixed payment" of one point whereas all other players received no additional payment. In two-prize contests, the two players with the two highest

[7] We expected no one would want to set a higher decision number so that this upper bound would not be binding on a subject. In fact we were almost right since there were only very few instances of subjects choosing the Maxeffort.

decision numbers received a "fixed payment" of 0.5 points whereas all other players received no additional payment. If in the one-prize contests, two or more group members chose the highest decision number, it was randomly decided which of these "tied" members received the prize of one point. In case of ties in the two-prize contests we proceeded in a similar fashion which was explained in the instructions. It was also explained and emphasized that decision costs would be subtracted no matter whether or not a subject had won. This implies that subjects could make losses. To cover those, subjects got a lump-sum fee of $5. (Given the exchange rate of 15 points = $1 in each treatment, all subjects started with an amount of 75 points in their experimental accounts.) Additionally, in each period subjects received an initial per-period endowment that was equal to their expected costs in equilibrium.[8] The specific numbers are shown in Table 1.

After each period, the feedback screen first informed a subject whether or not she had won an additional payment. Furthermore, the screen reiterated a subject's random number, decision number, decision costs, the difference between the payment in the previous period and the decision costs (excluding the initial endowment per period) and individual earnings in the previous period including the initial endowment per period. A last piece of information that was given to subjects depended on the number of prizes in a treatment and on whether or not a subject had won a prize. In one-prize contests, a subject that had not won a prize was informed about the random number of the winning subject. In two-prize contests, a subject that had won a prize was informed about the random number of the other winning subject whereas a subject that had not won a prize was informed about the random numbers of the two winning subjects.

In order to avoid income effects participants were informed that after the completion of the experiment ten out of the fifty periods would be randomly selected to count towards monetary earnings. That is, subjects were paid according to the sum of their individual earnings in these ten rounds. Finally, in order to make sure subjects had a good understanding of the decision problem and the procedures, we started each experiment with three trial periods that did not count towards monetary earning.

The experiments replicated the examples of contests described in section 2. The decision number corresponds to effort, the random number to a subject's ability, the decision costs to a subject's disutility of effort, and the payment corresponds to the prize(s).

Some remarks regarding our experimental design are in order. First, we avoided value-laden

[8] In equilibrium expected costs equal $\int_{.5}^{1} c(V_1 A(c) + V_2 B(c))dc$ where V_1 and V_2 depend on the treatment and $A(C)$ and $B(c)$ are given by (3). Note that expected costs in equilibrium do not depend on the form of the cost function.

terms in the instructions. Subjects were never called contestants or competitors. Similarly, other players were called "other group members." Also "prizes" were called "fixed payments. Second, each subject participated in only one treatment.

4 Results

In this section we will present the results of our experiments. We will do this by first presenting the aggregate results that, as we have noted in the introduction, appear to strongly support the theory. However, when this is done we will disaggregate our results and look at them more finely. Here we will demonstrate that these aggregate results mask the bifurcation phenomenon we have discussed above.

4.1 Aggregate Results

There is a sense in which an organizational designer need care only about aggregate or average results. Since he is designing the organization to maximize mean effort levels and revenues, these should be the variables he looks at. In addition, if he is risk neutral he need not worry how these means were composed.

In line with this way of thinking we first present the aggregate or mean results of our experiment and concentrate on effort behavior and revenue in the four treatments. Although we will also present summary statistics for all rounds of the experiment, we will concentrate on results in the second half of the experiment when subjects are more experienced. Also, unless we explicitly state otherwise, in all subsequent statistical tests of this section we take one session's average total effort in the second half of the experiment as one observation.

4.1.1 Effort Behavior

We will start our discussion by looking at the effort behavior of our subjects at the aggregate level. To do this consider Figure 1.

In Figure 1 we have four graphs, one for each of our four treatments. In each graph we present the equilibrium effort function (solid line) for the parameters defining that treatment as well as the average effort function (dashed line) being that non-linear function of the same form as the equilibrium function that best fits the scatter of mean efforts presented in the diagram. In other words, for any ability realization on the horizontal axis, the height of equilibrium effort function

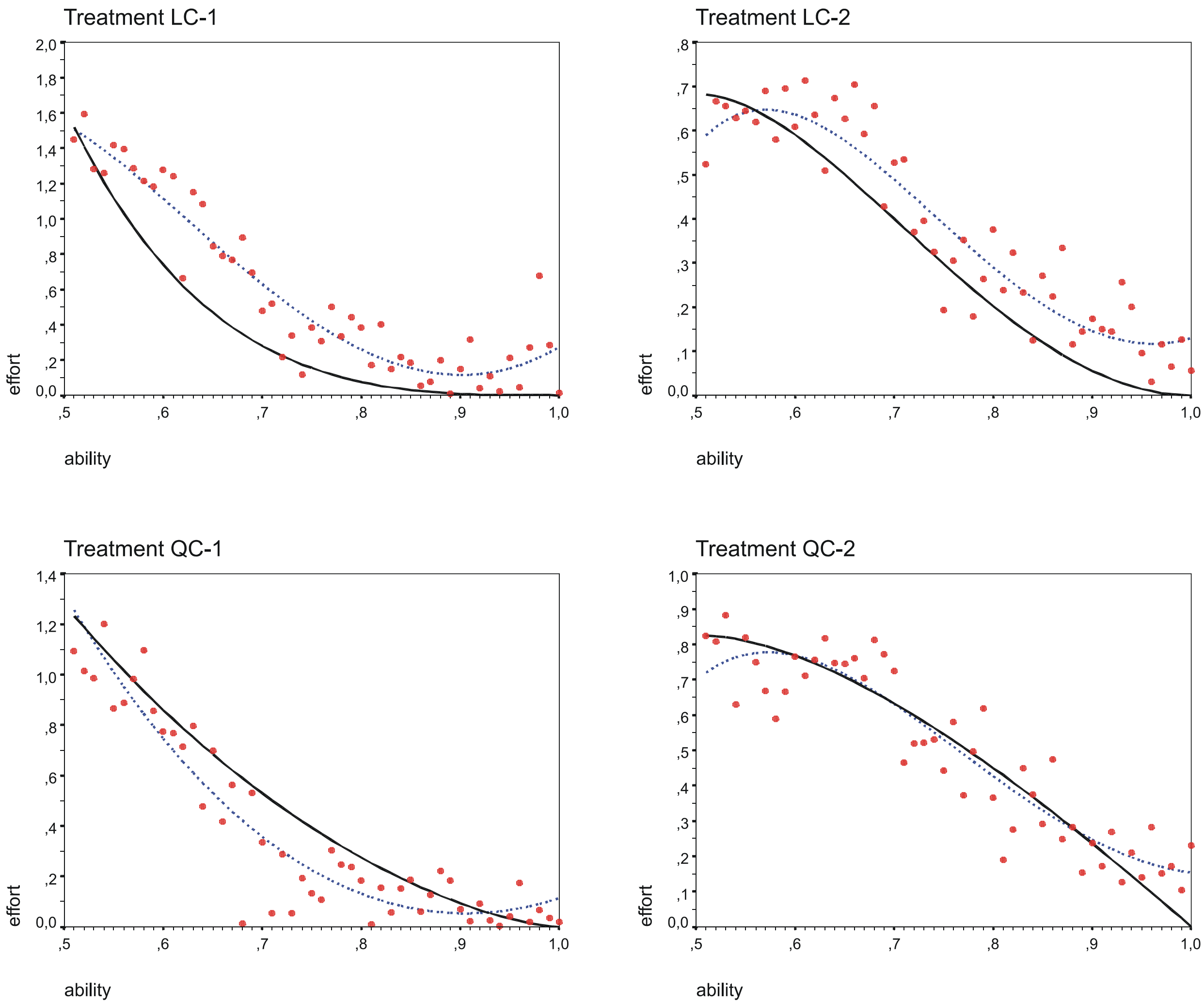

Figure 1: Average (•), optimal (solid line), and estimated (dashed line) bid functions in the second half of the experiment.

defines that effort which is a best response to that realization and the assumption that all others are using the same equilibrium function, while the height of the average effort function presents the conditional mean effort made at that realization. To show the pattern of efforts we also present the scatter of efforts representing the mean of the actual efforts placed when that ability was realized.

There are several things to note about Figure 1. First the average efforts made seem to track the shape of the equilibrium effort function quite well. Second, effort behavior appears to be continuous in that, on average there does not appear to be any large discontinuities in behavior. Finally, the levels of efforts appear to be consistent with the equilibrium effort function. This is particularly true for the QC-2 experiment where the equilibrium effort function appears to pass directly through the middle of the scatter of mean efforts. For the other treatments there appears to be overexertion in LC-1 and LC-2 and slight under-exertion in QC-1.

This behavior manifests itself in the average revenue data as well. Table 2 presents the mean revenue generated in each of our treatments along with the revenue that would have been generated by our subjects if, given their realizations, they had all submitted their equilibrium efforts. (For the column labeled "Sorting" see below.)

The revenue data presented in Table 2 are consistent with the observed effort behavior exhibited in Figure 1. While revenue levels were above those predicted by the equilibrium theory in the LC-1 and LC-2 treatments (with average observed revenue being about 65% higher than average equilibrium revenue in the LC-1 treatment (2.391 vs. 1.452) and 25% higher in the LC-2 treatment (1.452 vs. 1.164), in the QC-1 treatment they were below by about 18% (1.524 vs. 1.859). In the QC-2 treatment actual average revenues were remarkably on target (1.963 vs. 1.944). Applying a sign-test[9] to the data from the second half of the experiment we can reject the hypothesis that the median observed revenue is equal to the equilibrium level at the 1 per cent level in treatments LC-1, LC-2 and QC-1. For treatment QC-2, however, this hypothesis can not be rejected at any conventional significance level ($p = 0.858$, two-tailed).

Recall that theory predicts that in a linear-cost contest revenue is maximal if only one prize is awarded while, in our quadratic-cost contest, the designer maximizes total effort by awarding two equal prizes. Both of theses predictions are confirmed by our data. According to Table 2 and concentrating on results in the second half of the experiment, we see that whereas in treatment LC-1

[9] Consider the variable y_{jt} with $y_{jt} = 0$ in case the observed revenue in period t in session j is less than or equal to the equilibrium level and $y_{jt} = 1$ if observed revenue exceeds the equilibrium level. Then test whether or not the variable y_{jt} is binomial with 0.5 probability that $y_{jt} = 1$.

Treatment	Rounds	Average Revenue		Sorting
		optimal	observed	
LC-1	All	1.444	2.385	170/300
		(0.169)	(0.344)	(56.7%)
	Last 25	**1.452**	**2.391**	86/150
		(0.168)	(0.281)	(**57.3**%)
LC-2	All	1.236	1.602	149/300
		(0.050)	(0.172)	(49.7%)
	Last 25	**1.164**	**1.452**	70/150
		(0.061)	(0.312)	(**46.7**%)
QC-1	All	1.854	1.701	147/250
		(0.076)	(0.217)	(58.8%)
	Last 25	**1.859**	**1.524**	77/125
		(0.160)	(0.270)	(**61.6**%)
QC-2	All	1.966	2.045	121/250
		(0.090)	(0.308)	(48.4%)
	Last 25	**1.944**	**1.963**	65/125
		(0.093)	(0.364)	(**52.0**%)

Table 2: Observed revenue and efiiciency (Standard deviations based on group averages in parentheses)

average observed revenue is 2.391 it is only 1.452 in treatment LC-2. Taking one session's average total effort as one observation, a one-tailed Mann-Whitney U-test reveals that this difference is highly significant ($p = .001$). In the quadratic-cost contests, the average total effort of 1.963 in treatment QC-2 compares to an average of 1.524 in treatment QC-1. Again this difference is statistically significant ($p = 0.028$).

One might also ask, how reliable the different contests are in terms of producing the levels of average total efforts reported in Table 2. A look at standard deviations given in parentheses in Table 2 is revealing: One-prize contests are more stable than two-prize contests in the sense that standard deviations are lower in the first than in the latter (contests with linear costs: 0.281 vs. 0.312; contests with quadratic costs: 0.270 vs. 0.364).

Finally, one can ask whether our contests were efficient in sorting and promoting workers. For example, if there are one or two positions available for promotion (as in our experimental contests), the goal would be to select the worker with the highest ability or, respectively, the workers with the two highest abilities. This can be achieved with the contests studied in this paper since the equilibrium-effort functions are strictly monotonic with respect to ability. Thus if all subjects exert effort according to the equilibrium effort function, optimal sorting should occur. That is, in each round and each group of four contestants we would observe that the subject with the highest ability would exert the highest effort, the subject with the second highest ability would exert the second highest effort, and so on. Clearly, in an experimental setting optimal sorting in this strict sense cannot be expected throughout the entire experiment.[10] Instead we just ask, in how many cases it was true that the contestant with the highest ability won a one-prize contest respectively in how many cases it was true that the contestants with the two highest abilities were winners in a two-prize contest. The results are displayed in the fifth column of Table 2 labeled "Sorting". The entry in each cell gives the number of cases in which sorting worked and the number of all cases along with the percentage in parentheses. Concentrating on the results in the second half of the experiment, sorting in this weaker sense occurred in respectively 57.3%, 46.7%, 61.6% and 52.0% of the cases in treatment LC-1, LC-2, QC-1 and QC-2, respectively. Put differently, in about 40% (50%) of the cases in the one-prize (two-prize) contests, contestants not having the highest abilities won the contests. However, note that for example in the majority of cases in which sorting did not work in the two one-prize contests, the winner of the contest was the subject having

[10] In fact, optimal sorting in this strict sense is only observed in roughly ten per cent of all rounds in the second half of the experiment in the four different treatments.

the second highest ability who happened to exert an effort greater than the highest-ability subject. Hence it is the rat race that is responsible for inefficiencies.[11] Note finally that, non-surprisingly, the percentage of cases in which sorting does work is higher in one-prize contests than in two-prize contests.

In summary as predicted by the theory, we find that in case of linear costs a one-prize contest raises higher revenues than a two-prize contest whereas in case of quadratic costs a two-prize contest raises higher revenues than a one-prize contest. Furthermore, contests with linear costs seem to elicit excess efforts while those with quadratic costs elicit effort levels which are either too low or approximately equal to the equilibrium efforts. Finally, we observe that in only 50-60% of the cases (depending on the treatment), the contests are won by those subjects having the highest abilities.

4.2 Disaggregated Results

As we have seen above, if one were to look only at our data aggregated within treatments, one could come away with the impression that behavior was basically continuous and, on average, not far from that predicted by the theory. In this section of the paper we will attempt to disabuse you of those impressions by presenting a more disaggregated analysis of our data. We will do this in several steps. First we will present a small sample of individual effort functions presented just to give you a quick first impression of what typical effort behavior looked like. While this is not an exhaustive presentation of all effort functions, the subjects we select are by no means outliers so they should give you a good idea of what we are talking about. Second, we will present a set of histograms, one for each of our four treatments, which describe the efforts subjects made. These histograms will illustrate the fact that efforts tended to be bimodal. They were either heavily concentrated around zero (for those who dropped out) or scattered across high effort levels (for those entering the rat race) with relatively few effort levels chosen in the middle effort ranges. In other words, either people dropped out or they entered a rat race. Finally, we performed a model-selection test by contrasting, individual by individual, the goodness of fit of the best fitting step-wise linear effort function against the best fitting continuous function of the form specified by the equilibrium theory. Here we try to convince you that subject behavior can best be described by a step function characterized by an ability cut-off level c^* such that for all abilities below c^* (low costs) effort is very high while for abilities above c^* (high costs) efforts are low (or zero).

[11] There are, however, also cases in which contestants with the second to last ability or the least ability won a contest.

4.2.1 Individual effort functions

Figure 2 presents a set of 8 individual effort functions two each selected from our four treatments. While not all individuals exerted effort in this manner, in this section we will attempt to convince you that these effort functions are the rules and not the exception. More precisely, we will try to convince you that rather than effort being chosen in a smooth and continuous manner, typical behavior can be characterized by a discontinuous step function with a cut off effort level of c_i^* for individual i. While c_i^* varies from individual to individual, and while some individuals violate the rule, we still consider these 8 effort functions to be broadly representative of behavior.

Note how dramatic these effort functions are. For example, subject 4 in treatment LC-2 clearly exhibits a c_i^* of 0.70 and clearly drops out for all ability levels above it while subject 4 in treatment QC-1 drops out for all $c_i^* \geq 0.60$. Note, in addition, that when subjects exert positive effort they do so very often at levels far above those prescribed by the equilibrium effort function. These drop-out efforts and over exertions are precisely the bifurcations that were described in our introduction.

4.2.2 Effort Histograms

Perhaps a more efficient way to demonstrate the bifurcation of individual effort in these experiments is to present Figure 3 which describes the histograms of individual effort levels (on the right hand side) in our four treatments along with what we would expect these histograms to look like if, given the actual ability draws of our subjects, they had all made their equilibrium effort choices (on the left hand side).

To describe these histograms let us look first at the those of treatment LC-2 (second from the top in Figure 3). As we see in the left panel, if subjects had all used their equilibrium effort functions to select effort levels, given the ability realizations in the sessions, we would have expected to see a more or less uniform distribution of efforts. By contrast, the right panel presents what we actually saw which is quite different. Note that there are a huge numbers of efforts around the 0 effort level indicating a large amount of drop-out behavior as well as a larger number of effort levels above 0.60 indicating larger than expected efforts. The same pattern exists in all of the other Figures with an even more pronounced bifurcation in treatment QC-1 and QC-2.

From these histograms it should be clear that behavior in these experiments was bimodal. Either subjects dropped out or they exerted above expected effort levels which is consistent with

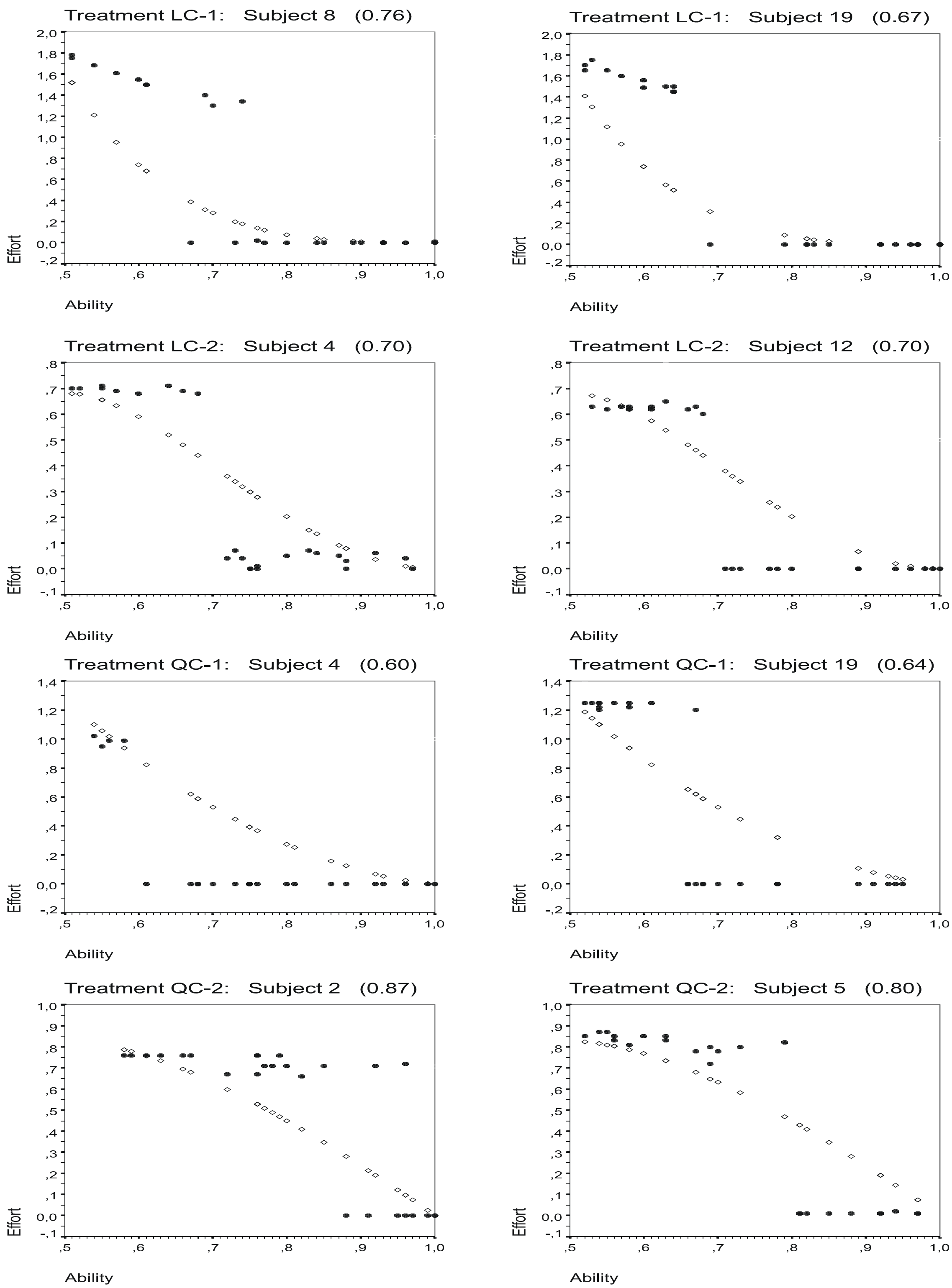

Figure 2: Examples of individual behavior (optimal $\diamond$; observed $\bullet$). Note: Cut-off levels c_i^* in parentheses (see Section 4.2.3).

our bifurcation hypothesis.

4.2.3 Step-Functions

Final support for our bifurcation hypothesis comes from the following model selection exercise. If we are correct in supposing that individual behavior was bimodal and exhibits either drop-out or over-exertion behavior, then we would expect that the best fitting model of individual effort would be a step-function characterized by a cut-off ability level, c_i^*, such that if subject i's observed ability, c_i, were above c_i^* then the subject would "drop out" and exert either zero or at least very low effort, while if c_i were below c_i^*, the individual would exert positive and substantial efforts. This model can be tested against the equilibrium model which posits a continuous effort function of the form specified by (2) and (4), or the best fitting continuous effort function of that general form.

To compare these models we first fit a simple switching regression model for each subject (using only data from the second half of the experiment) of the form,

$$b_{it} = \alpha_0 + \alpha_1 c_{it} + \alpha_2 D_{c_i^*} + \alpha_3 D_{c_i^*} c_{it} + \varepsilon_{it} \tag{5}$$

where b_{it} (c_{it}) is subject i's effort (ability) in period t, and $D_{c_i^*}$ is a dummy which is equal to 1 if $c_{it} > c_i^*$ and equal to 0 otherwise. The parameter $c_i^* \in \{.51, .52, ..., 1\}$ is the value of the ability at which the structural break in the subject's effort behavior occurs. Note that in case of $D_{c_i^*} = 0$ equation (5) reads $b_{it} = \alpha_0 + \alpha_1 c_{it} + \varepsilon_{it}$ whereas in case of $D_{c_i^*} = 1$ it reads $b_{it} = (\alpha_0 + \alpha_2) + (\alpha_1 + \alpha_3) c_{it} + \varepsilon_{it}$. Thus the graph of (5) consists of two line segments with intercepts α_0 before and $\alpha_0 + \alpha_2$ after the break and slopes α_1 before and $\alpha_1 + \alpha_3$ after the break, respectively. Note that in case of $-\alpha_2 \neq \alpha_3 c_i^*$ the graph in (5) has a discontinuity occurring at the point of structural break. The best-fitting breakpoint c_i^* and the respective coefficients in (5) were estimated from the data.

For this purpose, we estimated equation (5) for all possible points of structural break $c_i^* \in \{.51, .52, ..., 1\}$ and chose as the optimal breakpoint the one that maximizes the adjusted R^2. Using the corresponding estimates of the coefficients in (5), we then computed for each subject i and for each period $t \in \{26, 27, ..., 50\}$ the predicted effort $\left(b_i^t\right)^{pred}$and computed, subject by subject, the sum of the square deviation, SSD_i, defined as $SSD_i = \sum_{t=26}^{50} \left(\left(b_i^t\right)^{pred} - \left(b_i^t\right)^{obs} \right)^2$ where $\left(b_i^t\right)^{obs}$ is the observed effort of subject i in period t. The average SSD for each treatment is given in the second column in Table 3.

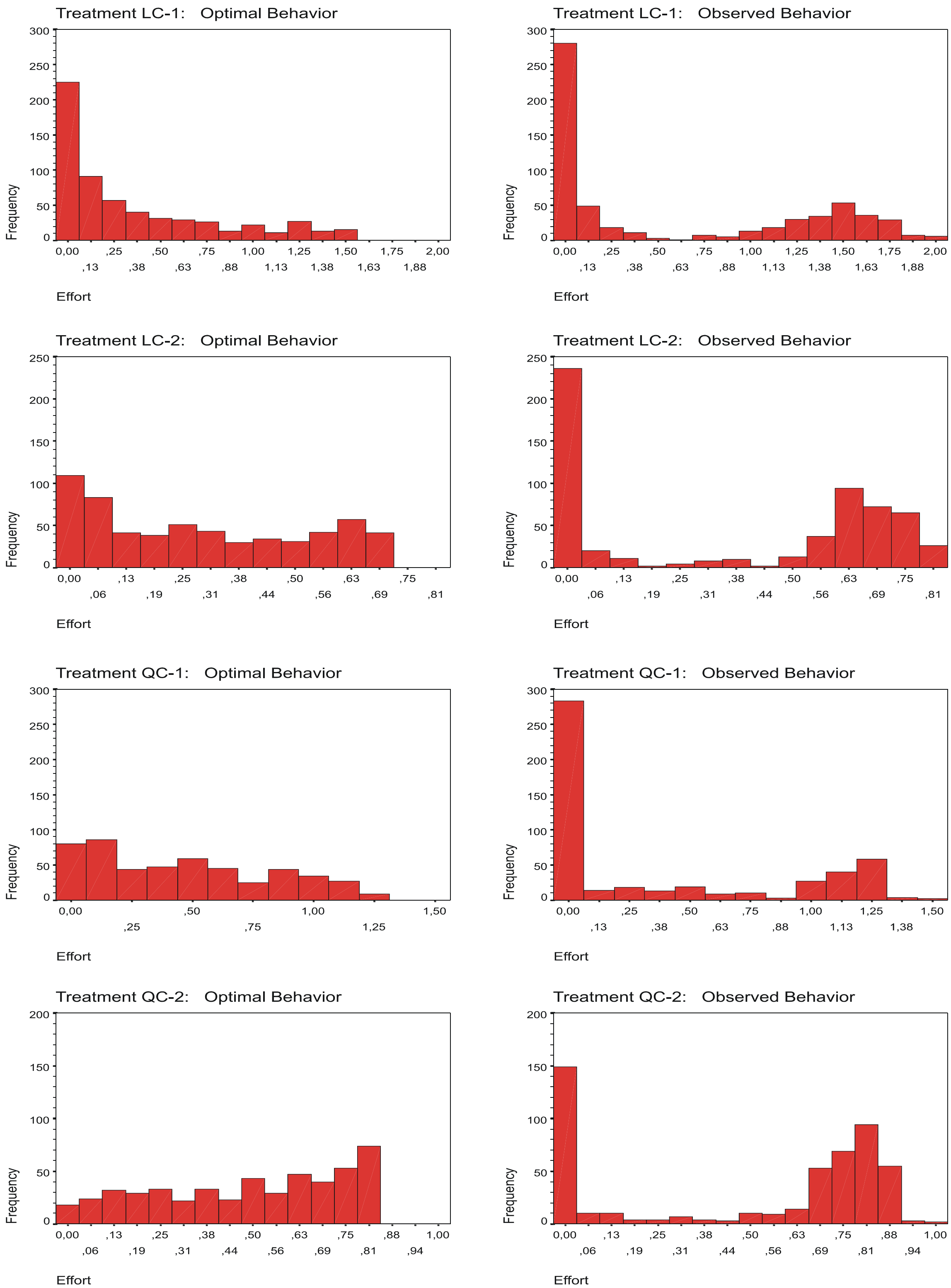

Figure 3: Histograms of individual effort choices

Treatment	Average sum of the square deviation (SSD) based on			p-level
	switching regr. model	equilibrium	"equilibrium form"	Wilcoxon test
LC-1	2.86	8.77	3.63	0.000
LC-2	0.29	1.82	0.48	0.000
QC-1	0.87	3.49	2.10	0.000
QC-2	0.59	2.08	1.05	0.000

Table 3: Overview: Sum of the square deviation (SSD)

We compared the resulting SSD_i's of this estimation to two others. The first was the SSD_i generated using, $\left(b_i^t\right)^{pred}$, the predictions of the equilibrium effort functions as given in (2) and (4). Second we compared our SSD_i's to those generated by estimating the best fitting effort function for each individual of the $form$ of the respective equilibrium-effort function in each of the four treatments as given in equations (2) and (4). For instance, using OLS regression we estimated for each subject in treatment LC-1 the model

$$b_{it} = \beta_0 + \beta_1 c_{it} + \beta_2 c_{it}^2 + \beta_3 \ln c_{it} + \varepsilon_{it} \tag{6}$$

where, again, b_{it} (c_{it}) is subject i's effort (ability) in period t. (Note that equation (6) has the form of the equilibrium effort function given in (2) with the exception that the coefficients are undetermined.) Likewise for treatment LC-2. Recall that the equilibrium-effort functions for the treatments with quadratic costs, i.e. treatments QC-1 and QC-2, are the square roots of the equilibrium effort-functions in the respective linear-costs treatments (compare equations (2) and (4)). In order to be able to use OLS regression for the estimation in these treatments, too, we proceed as follows. Consider e.g. treatment QC-1. Instead of estimating (4) we estimated the model

$$(b_{it})^2 = \beta_0 + \beta_1 c_{it} + \beta_2 c_{it}^2 + \beta_3 \ln c_{it} + \varepsilon_{it}$$

i.e., the squared equation. In order to compute the SSD_i for these cases we then used the radical of the predicted efforts.

The results of our exercise are given in Table 3 which presents the average SSD_i value for each treatment. Column 2 presents the results of our switching regression model while columns 3 and 4 present the results of our equilibrium and equilibrium-form models, respectively.

	LC-1	LC-2	QC-1	QC-2
(mean)	**(0.71)**	**(0.78)**	**(0.67)**	**(0.81)**
LC-1	—	—	—	—
LC-2	0.013	—	—	—
QC-1	0.174	0.000	—	—
QC-2	0.003	0.403	0.000	—

Table 4: Average cut-off levels and (two tailed) p-values of pairwise differences in the second half of the experiment

As can be seen, our simple switching regression models clearly outperform the prediction of both the equilibrium and equilibrium-form models. In fact, using a Wilcoxon test to compare the SSD_i's based on either the switching regression model and the estimates based on the equilibrium-form regressions indicates, that the former gives a highly significantly better fit than the latter.

Table 4 shows the average cut-off levels in each of the four treatments as well as (two tailed) p-values of pairwise Mann-Whitney U-tests. Recall that the switching regime consists of two line segments with a (possible) jump between the two segments. Note that the average cut-off points in the one-prize contests are lower than the average cut-off points in the two-prize contests (0.71 in LC-1 vs. 0.78 in LC-2; 0.67 in QC-1 vs. 0.81 in QC-2). As it turns out, these differences are also statistically highly significant. This means that subjects in the one-prize contests only start to exert serious effort when their ability parameters, the c_i's, are comparatively low. This implies that they exert low effort levels over a much larger interval of the domain of their effort function. Finally, note that the differences between cut-off levels in the two one-prize and the two two-prize contests are small and not significant.

4.2.4 Conclusions and Discussion

It should be clear from what we have described that behavior in laboratory organizations with what is claimed to be an optimal incentive structure, while eliciting average effort levels that were approximately what the theory expected, also determined behavior that we firmly expect to be dysfunctional in the long run. More precisely, in the real world, unlike in our experiment, subjects arrive at the organization with an ability that does not change from period to period. They are either high or low ability workers. What our results predict is that such organizations will evolve over time into a caste system in which a set of high ability workers will compete for the organizational prizes while the remaining low ability workers will be left behind. Once low ability workers drop out, as long as we assume they do not reenter the fight for organizational prizes, we would expect the high ability workers to also lower their efforts since they will eventually realize that there is less competition out there than they initially expected.[12] Hence, in the long run we might actually expect lower efforts than predicted by Moldovanu and Sela. Further, we can also expect such organizations to have higher than optimal turnover levels since any organization composed of over achievers and alienated drop outs can not be a pleasant place to work.

A second finding is that contests with linear costs, on average, elicit excess efforts while those with quadratic costs elicit effort levels which are either too low (in case of one prize) or approximately equal to the equilibrium efforts (in case of two prizes). It is instructive to compare this result with those reported in related studies. There is ample evidence that there is over-dissipation relative to the Nash equilibrium in rent-seeking contests (Davis and Reilly, 1998; Potters et al., 1998) as well as in single-unit all-pay auctions (Gneezy and Smorodinsky, 1999; Amann and Leininger, 1998). Barut et al. 2002 show that the over-dissipation result carries over to multiple-unit all-pay auctions. Also, "effort above the RNNE [risk-neutral Nash equilibrium] is the most common outcome in single unit first-price private value auctions." (Kagel 1995, p. 523). In sum, the evidence concerning over-dissipation in rent-seeking contests or over-exertion in (all-pay) auctions is overwhelming. In the light of this evidence, our over-exertion result in the contests with linear costs is hardly surprising. However, the under-exertion or close-to-optimal effort result in the contests with quadratic costs is. Furthermore, while over bidding in e.g. first-price auctions can be rationalized by assuming that subjects are risk-averse (see e.g. Cox et al. 1988), risk aversion cannot coherently explain the different patterns regarding revenue observed in our experiments since risk aversion would predict under-exertion relative to the risk-neutral Nash equilibrium independent of the kind of costs contestants have to bear.

Finally, our third finding is related to the efficiency of contests in organizations: Workers are to be promoted on the basis of their skills and abilities. Since workers usually differ with respect to their abilities and since the equilibrium-effort functions are strictly monotonic with respect to ability, the contests analyzed in this study theoretically serve the purpose of awarding promotion prizes to those workers having the highest abilities. We observe, however, that in only 50-60% of the cases our experimental contests are won by highest-ability subjects. The reasons driving this result are likely errors in decision making and idiosyncratic or random behavior of some of the

[12] For this line of reasoning see also e.g. Pema (2002) p.39.

subjects as revealed by the answers to a post-experimental questionnaire.[13]

In conclusion, what may appear to be an optimal organization (or contest) on paper may not be one in reality. The propensity of workers to drop out when they suspect that they do not have a sufficient chance of winning an organizational prize leads to a caste system within organizations which is many times demoralizing to all who work there.

References

[1] Akerlof, G. (1976): The economics of caste and of the rat race and other woeful tales, *Quarterly Journal of Economics* 90, 599-617.

[2] Amegashie, J.A. (2003): Burning out in sequential elimination contests, *Working Paper*, University of Guelph.

[3] Amann, E. and W. Leininger (1998): Auctions and Contests: An Experimental Investigation, Wirtschaftstheoretische Diskussionsbeiträge, Nr.98-06, Essen University.

[4] Barut, Y., D. Kovenock, and C. Noussair (2002): A comparison of multiple unit all-pay auction and winner pay auction under incomplete information, *International Economic Review* 43, 675-708.

[5] Benoit, J.-P. (1999): Color blind is not color neutral: testing differences and affirmative action, *Journal of Law, Economics & Organization* 15, 378 - 400.

[6] Corns, A. and A Schotter (1999): Can Affirmative Action Be Cost Effective? An Experimental Examination of Price-Preference Auctions, *American Economic Review* 89, 291-305.

[7] Cox, J. C., V. L. Smith, and J. M. Walker (1988): Theory and individual behavior of first-price auctions, *Journal of Risk and Uncertainty* 1, 61-99.

[8] Davis, D. and R. Reilly (1998): Do too many cooks always spoil the stew? An experimental analysis of rent seeking and the role of a strategic buyer, *Public Choice* 95, 89-115.

[13] A case in point is a subject in treatment LC-1 who stated: "[...] Then I was still frustrated that I hadn't ever had the highest number. Just for fun, I typed 1.96 and got the extra point, as well as a positive profit. From that point, I only picked extremely high numbers." Another subject in treatment QC-2, wrote "After a while it became clear that numbers, which did not win, would receive a high negative value. Therefore I started to choose values over 0.80 to receive the 0.50 payment, despite the fact that it reduced the profit."

[9] Fischbacher, U. (1999): Z-Tree, Zurich Toolbox for Readymade Economic Experiments, *Working paper 21*, Institute for Empirical Research in Economics, University of Zurich.

[10] Gneezy, U. and R. Smorodinsky (1999): All-pay auctions: An experimental study, mimeo, Department of Economics, University of Haifa.

[11] Holzer, H., and D. Neumark (2000): Assessing affirmative action, *Journal of Economic Literature* XXXVIII, 483-568.

[12] Kagel, J. H. (1995): Auctions: A survey of experimental research, in: *Handbook of experimental economics*, eds: J. H. Kagel and A. E. Roth, Princeton University Press, Princeton.

[13] Moldovanu, B. and A. Sela (2001): The optimal allocation of prizes in contests, *American Economic Review* 91, 542-558.

[14] National Institute of Government Purchasing (1994), 1993 procurement practices, Reston, VA: National Institute of Government Purchasing.

[15] Potters, J., C. de Vries and F. van Winden (1998): An experimental examination of rational rent-seeking, *European Journal of Political Economy* 14, 783-800.

[16] Pema, E. (2002): Internal labor markets revisited: Tournaments in Academia, *Working Paper*, Michigan State University.

[17] Prendergast, C. (1999): The provision of incentives in firms, *Journal of Economic Literature* XXXVII, 7-63.

[18] Schotter, A. and K. Weigelt (1992): Asymmetric Tournaments, Equal Opportunity Laws and Affirmative Action: Some Experimental Results, *Quarterly Journal of Economics* 107, 511-539.

[19] Tenorio, R. (2000): The economics of professional boxing contracts, *Journal of Sports Economics 1,* 363-384.

A Instructions for treatment LC-2

This is an experiment in decision-making. If you make good decisions you can earn a substantial amount of money, which will be paid to you when you leave. The currency in this decision problem is called Points.

All payoffs are denominated in this currency. At the end of the experiment your earnings in Points will be converted into real U.S. dollars at a rate indicated below.

As you read these instructions you will be in a room with a number of other subjects. Each subject has been randomly assigned an (electronic) ID number.

The experiment consists of 50 decision rounds. In each decision round you will be grouped with three other subjects by a random drawing of ID numbers. These three subjects will be called your "group members." Your group members will remain the same throughout the entire experiment. The identity of your group members will not be revealed to you.

The Decision Problem

In the experiment you will perform a simple task. At the beginning of each round the computer will first independently generate a random number for every group member. The random number will be one of the 51 numbers in the set {0.50, 0.51, , 1.00}. Each of these 51 numbers has an equally likely chance of being chosen. You will then be informed about the random number that was chosen for you. You will, however, not be informed about the random numbers that were chosen for the other group members. These random numbers will be important to you since they will determine your costs in the experiment as explained below. After informing you about your random number, the computer will ask all group members to simultaneously choose a Decision Number (which will be the only decision you have to make in a round.) This Decision Number must be chosen from the set of numbers {0, 0.01, 0.02, ..., 0.82}. Associated with each Decision Number are decision costs. These decision costs depend on your random number as well as on the Decision Number you chose. More precisely, the decision costs will be equal to the product of the random number and your Decision Number. For example, say you receive a random number of 0.6 and in the experiment choose a Decisions Number of 0.7. Then your cost would be 0.42 = 0.7 x 0.6. If instead your random number was 0.9 and you chose a Decision Number of 0.7, your decision costs would be 0.63 = 0.7 x 0.9. You can consider your random number to be the per-unit cost of choosing a Decision Number so the higher the random number the higher is that per unit cost. Note that the decision costs associated with the Decision Number 0 are equal to 0.

To help you calculate what the cost of any Decision Number will be given your random number, we have provided you with a calculator that is located on the left hand side of your decision screen. To find the decision cost associated with any Decision Number simply enter a Decision Number into the box and then push the button "compute". Your cost will then be shown to you at the top left corner of your screen.

When you are ready to make your final decision, please enter your Decision Number into the box on the right hand side of your screen and push the button "OK".

Calculation of Payoffs

Your payoff in each decision round will be computed as follows. First of all, in each round each participant will receive a flat payment of 0.20 Points no matter which number he or she and the other group members have chosen. Whether or not you receive an additional fixed payment will be determined in the following way. After every member of your group has entered his or her Decision Number, the computer will compare all of the Decision Numbers of the four members of your group. If your Decision Number is one of the two highest, you will receive the fixed payment of 0.5 Points otherwise you receive no additional fixed payment. If three or more group members chose the highest Decision Number, then the computer will randomly determine which two of these "tied" members receive the additional fixed payment of 0.5 Points. Those subjects with Decision Numbers that are not the highest two will receive nothing. From your fixed payment (of either 0.5 Points or 0 Points) you will have to subtract your decision cost. Hence, while choosing a high Decision Number increases the probability that you will win a positive fixed payment it also increases the cost of doing so. In addition, if your Decision Number is not one of the two highest of the group, you will receive no additional fixed payment and have to subtract your decision costs from your initial flat payment.

Your payoff in a given round is calculated as follows: First, as mentioned above, you receive a flat payment (FP) of 0.20 Points. In addition if you chose one of the two highest Decision Numbers you will be paid a fixed payment of 0.5 Points from which you will subtract your decision cost. If you do not choose one of the two highest Decision Numbers, you will receive a fixed payment of 0 and still have to subtract your decision costs. The resulting number is multiplied by 100 to yield your final Points payoff. This is then converted into dollars at the rate of 15 Points = $1. Thus, your final payoff in Points in a given round is: Payoff = 100*(Flat payment + Fixed payment (0 or 0.5) - Decision Cost).

Note: To make life easier for you so that you do not have to enter decimal Points, you will not be asked to enter a Decision Number from the set {0, 0.01, 0.02, ..., 0.82} but from the set {0, 1, 2, ..., 82}. The computer will then automatically divide the Decision Numbers of all group members by 100 before starting to evaluate them.

Example of Payoff Calculation

Suppose the following occurs: Group member 1 gets assigned random number 0.80 and chooses Decision Number 0.21 (21). Group member 2 gets assigned random number 0.55 and chooses Decision Number 0.17 (17). Group member 3 gets assigned random number 0.91 and chooses Decision Number 0.05 (5). Group member 4 gets assigned random number 0.77 and chooses Decision Number 0.33 (33).

Since group members 4 and 1 chose the highest two Decision Numbers they receive the Payment of 0.5 Points whereas all other group members receive no payment. Therefore, group member 4's earnings in this round would be 100*(0.20 + 0.5 - 0.77*0.33) = 44.59 Points whereas group member 1's earnings in this round would be 100*(0.20 + 0.5 - 0.80*0.21) = 53.2 Points. Group members 2, and 3 each receive no additional payment. Therefore group member 2 would earn 100*(0.20 + 0 - 0.55*0.17) = 10.65 Points, and, finally, group member 3 would earn 100*(0.20 + 0 - 0.91*0.05) = 15.45 Points.

Note again that the decision cost is a function of the random number and the Decision Number. Note also that your earnings in a round depend on the following: your random number, your Decision Number and your group members' Decision Numbers. Your earnings do not depend on your group members' random numbers.

Continuing Rounds

After round 1 is over, the same procedure will be repeated for round 2, and so on for 50 rounds. That is, in each round a random number will first be generated for you, then you will choose a Decision Number which will be compared to the Decision Numbers of the other members of your group, and the computer will calculate your earnings for the round.

After each round you will be informed about which payment you receive. In case you do receive a positive payment you will be informed about the random number of the other group member who also received a payment of 0.5 Points. In case you do not receive a positive payment (because your Decision Number was not one of the two highest among the Decision Numbers of all group members or because you were not randomly selected in case you and at least two other group members chose the highest Decision Number) you will be informed about the random numbers of the group members who received the payment of 0.5 Points.

Calculation of Final Monetary Payment

At the start of the experiment you get a one-off endowment of 75 Points. (This is the $5 show-up fee you were promised, see below.)

When round 50 is completed, the computer will randomly select 10 of the 50 rounds. Your final payoff in the experiment will be the sum of your individual earnings in Points for only these 10 rounds (plus your endowment). For each 15 Points you will be paid 1 $.

Trial Periods

At the beginning of the experiment there will be three trial periods that do not count towards payment of real money.