Journal of Economic Behavior & Organization 99 (2014) 155–167

Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

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

Bounded rationality and group size in Tullock contests: Experimental evidence☆,☆☆

Wooyoung Lim[a], Alexander Matros[b,c], Theodore L. Turocy[d,e,*]

[a] The Hong Kong University of Science and Technology, Hong Kong
[b] Moore School of Business, University of South Carolina, United States
[c] Lancaster University Management School, United Kingdom
[d] School of Economics, University of East Anglia, United Kingdom
[e] CBESS, University of East Anglia, United Kingdom

ARTICLE INFO

Article history:
Received 20 June 2012
Received in revised form 13 August 2013
Accepted 14 December 2013
Available online 14 January 2014

JEL classification:
C72
C92
D72

Keywords:
Contests
Rent-seeking
Laboratory experiments
Bounded rationality

ABSTRACT

We explore how models of boundedly rational decision-making in games can explain the overdissipation of rents in laboratory Tullock contest games. Using a new series of experiments in which group size is varied across sessions, we find that models based on logit choice organize the data well. In this setting, logit quantal response equilibrium (QRE) is a limit of a cognitive hierarchy (CH) model with logit best responses for appropriate parameters. While QRE captures the data well, the CH fits provide support for relaxing the equilibrium assumption. Both the QRE and CH models have parameters which capture boundedness of rationality. The maximum likelihood fits of both models yield parameters indicating rationality is more restricted as group size grows. Period-by-period adjustments of expenditures are more likely to be in the earnings-improving direction in smaller groups.

© 2014 The Authors. Published by Elsevier B.V. All rights reserved.

1. Introduction

Tullock (1980) introduced a simple model of competition, in which competitors irreversibly expend costly resources in the hope of obtaining a prize of fixed value. The winner of the prize is determined stochastically, with a competitor's chances of victory increasing as he expends more resources. Variations of the basic model can be applied to settings ranging from lobbying for political influence, to research and development races, to fund-raising lotteries. This broad applicability has

☆ This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

☆☆ We have benefited from comments by Kyung Hwan Baik, Maya Bar-Hillel, Subhashish Modak Chowdhury, David Cooper, Sergei Izmalkov, Peter Moffatt, Thomas Palfrey, Svetlana Pevnitskaya, Tim Salmon, Jorgen Weibull, and Huan Xie; seminar participants at University of South Carolina, Florida State University; University of Nevada, Reno; Higher School of Economics (Russia); New Economic School (Russia); and University of Nottingham; conference participants at the International Research Workshop "Rationality, Behavior and Experiments" (Russia) and the 20th International Conference on Game Theory, Stony Brook, as well as those of the co-editor and two anonymous referees. We gratefully acknowledge financial support from University of Pittsburgh for Central Research Development Fund award.

* Corresponding author at: School of Economics, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, United Kingdom. Tel.: +44 01603597179

E-mail addresses: T.Turocy@uea.ac.uk, ted.turocy@gmail.com (T.L. Turocy).

0167-2681/$ – see front matter © 2014 The Authors. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jebo.2013.12.010

supported a vibrant literature on studying these games in the laboratory. A majority of laboratory studies find that subjects on average exceed the risk-neutral Nash equilibrium predictions for resource expenditure. Morgan et al. (2012), in their Table 1, provide an excellent summary of the literature which illustrates the robustness of this result. Millner and Pratt (1989), Davis and Reilly et al. (1998), Fonseca (2009), Morgan et al. (2012), and Fallucchi et al. (2012), among others, observe higher-than-Nash average expenditure using a variety of designs.

A number of explanations have been advanced to attempt to account for the overdissipation of rents in these contests. Amaldoss and Rapoport (2009) and Sheremeta (2011) propose biases in judgment lead to aggressive play. Parco et al. (2005), Sheremeta (2010) and Sheremeta (2011) investigate the extent to which a non-monetary preference for winning can account for high expenditure levels. Mago et al. (2012) and Wärneryd (2012) develop the idea that higher expenditures form an evolutionarily stable behavior.

In the analysis in this paper, our focus will be on statistical models of boundedly rational behavior, supposing that play is noisy and that participants do not calculate or play best responses precisely. This approach has been considered in the past in Tullock contests by Sheremeta (2011) and Schmidt et al. (2013), as well as by Bullock and Rutstrom (2007) in a transfer-seeking game presented using a matrix frame.[1] Our model is founded on a logit-response assumption, which underlies a noisy cognitive hierarchy model in the spirit of Camerer et al. (2004), and a quantal response equilibrium (QRE) model (McKelvey and Palfrey, 1995). The model relaxes Nash equilibrium in two ways: (1) by permitting players to hold incorrect beliefs about the play of others, and (2) by assuming players may not choose best replies with probability one. Logit quantal response equilibrium, cognitive hierarchy with exact best responses for higher-order thinkers, and Nash equilibrium are all special cases of our model. In particular, in our estimation we do not need to impose the mutual-consistency assumption inherent in Nash equilibrium or logit QRE.

A criticism of statistical models of the sort we consider is that they are often employed in a post-hoc fashion, with less attention given to the ability of models to organize data across treatments. Haile et al. (2008) point out that quantal response-type models can capture essentially arbitrary distributions of play, unless further restrictions are imposed. In our model, there are two parameters, capturing the mean number of degrees of iterative reasoning, and the precision of best responses. Results in McKelvey and Palfrey (1995) and Rogers et al. (2009) show that when fitting parameters in models similar to ours across games, the resulting estimates can vary substantially.

In our experiment, the treatment variable is the number of players participating in the contest. We consider contests with two, four, and nine players, in an across-subjects design. The Nash equilibrium prediction is that the expenditure per player decreases as the number of players increases, while the total expenditure of all players increases, converging to full dissipation of the rent from below. In our model, individual expenditure is less sensitive to group size than Nash equilibrium predicts. For model parameterizations far from Nash equilibrium, total group expenditure rises more rapidly as the number of players increases, to levels well in excess of full dissipation of the value of the prize.

There are three previous studies which directly or indirectly consider the effect of group size on behavior in Tullock contests. Anderson and Stafford (2003) provide the most direct manipulation. In a one-shot contingent-choice design, participants are asked to formulate contest expenditures in each of six possible settings, which vary both in the number of players as well as heterogeneity of costs. They find that in general a larger number of opponents results in lower expenditures, although in their data average expenditure in five-player contests actually exceeds that in two-player and four-player games. Sheremeta (2011) investigates, among other treatments, whether total expenditure is larger in a grand contest involving four players, versus two sub-contests, involving two players each, each for a prize worth one-half as much. He finds that individual expenditure relative to the prize is lower for the four-player contests. Morgan et al. (2012) study contests where potential participants may choose to enter, or to sit out the contest and accept an outside option payoff. This generates contests with different numbers of players, depending on the entry decisions of the subjects. They also generally find that when the number of players is larger, individual expenditure falls.[2]

Our experiment is the first to consider the effect of group size in an across-subjects design with repeated trials and holding constant the size of the prize and endowment. Qualitatively consistent with our model, and generally in contrast to previous results, we find that the group size has little effect on average expenditure levels. However, we do find treatment effects in terms of the distribution of expenditures, with expenditures being more dispersed in larger groups. Because average individual expenditures do not respond to the group size, the result is that aggregate expenditures are significantly larger in larger groups, with nine-player groups spending on average almost three times the value of the prize.

Because in our experiment the earnings-maximizing best response depends only on the total expenditure of other players, and not on the number of players per se, our experiment allows us to ask whether the bounded rationality parameters of our model are stable within a class of games, and an experimental environment, where as much as possible is held constant. Previous results in such domains are mixed; Gronberg et al. (2012) report stable estimates of the logit QRE parameter across two treatments in a public goods game, whereas Sheremeta (2011) reports QRE parameter estimates which vary across

[1] In addition, Morgan et al. (2012) and Potters et al. (1998) display distributions of expenditures in Tullock contests which are qualitatively consistent with the predictions of our model. Those authors do not pursue modeling the heterogeneity in their data. Gneezy and Smorodinsky (2006) study the related all-pay auction and find some support for the predictions of QRE from Anderson et al. (1998a).

[2] In addition, there are several studies which consider the effects of the number of players in all-pay auctions and in tournaments, which share some characteristics of the Tullock contest game. See Dechenaux et al. (2012) for a survey of these results.

different implementations of contests. Our parameter estimates indicate that decisions are noisier, relative to the financial stakes in the experiment, in larger groups. Therefore, the overdissipation observed in large groups has two components; it arises partly because decisions are heterogeneous in the first place, and partly because the precision of best responses degrades in larger groups.

To investigate this last result more fully, we examine period-by-period changes in expenditures in relation to the predictions of learning direction theory (Selten and Buchta, 1994; Selten and Chmura, 2008). We find that, when the number of contestants is small, subjects do tend to adjust their expenditure in the direction of better responses, as learning direction theory would predict. With larger numbers of contestants, adjustment, when it occurs, is equally likely to be away from the earnings-maximizing choice as it is toward it. This observation is consistent with the interpretation of lower values for best-response precision in larger groups, in that both indicate that the expected payoffs of decisions have less explanatory power in organizing subject expenditure when more contestants are present.

The paper is organized as follows. In Section 2, we first describe the Tullock contest game, and provide an exposition of the logit cognitive hierarchy framework in this setting. In Section 3, we lay out the design of our experiment. Section 4 presents our data and analysis of the results. Section 5 concludes with a discussion.

2. Theory

2.1. The game

We study a single-prize contest following Tullock (1980). There is a commonly known number of players N, each of whom has an endowment ω. The value of the prize to be awarded is V, and is the same for all contestants. Each contestant $i = 1, \ldots, N$ simultaneously chooses an amount $x_i \in [0, \omega]$ to spend. The chance of contestant i winning the prize given a profile of expenditures $\{x_j\}_{j=1}^{N}$ is $x_i/(\sum_{j=1}^{N} x_j)$. The prize is not awarded if $x_1 = \ldots = x_N = 0$.

For a given contestant i, write the sum of other contestants' choices as $X_{-i} = \sum_{j \neq i} x_j$. The expected monetary payoff to contestant i is

$$u_i(x_1, \ldots, x_N) = \omega - x_i + \frac{x_i}{x_i + X_{-i}} \cdot V. \tag{1}$$

If players are risk-neutral, the game has a unique Nash equilibrium, assuming ω is large enough, in which

$$x_1^\star = \ldots = x_N^\star = x^\star(N) = \frac{N-1}{N^2} \cdot V. \tag{2}$$

The total expenditure in Nash equilibrium is $(N - 1)/N \cdot V$, which is less than the amount of the prize.

2.2. Statistical models of noisy play

2.2.1. Logit best responses

In laboratory experiments studying Tullock contests, expenditure levels are observed across the entire strategy space, a phenomenon Chowdhury et al. (2013) refer to as "overspreading." In addition, Sheremeta (2010), Morgan et al. (2012), and Schmidt et al. (2013) all demonstrate that the size of the strategy space shapes behavior, in that restricting the endowment to be less than the value of the prize lowers expenditures, even when the endowment remains large relative to the value of the prize. To capture these regularities, the building block of our theoretical analysis is a random-utility model using the logit specification. Let π denote a mixed strategy over the feasible expenditures. Denote by $u(x ; \pi)$ the expected payoff to a player if he chooses expenditure level x, assuming all other players play according to the mixed strategy π.[3] The player observes not the expected payoff itself, but rather the expected payoff plus some additive noise term, i.e., $\tilde{u}(x; \pi) = u(x; \pi) + \varepsilon_x$, where the noise term is independently drawn for each expenditure from the extreme value distribution with precision parameter $\lambda \geq 0$. Then, the player's choice probability for each expenditure x is given by[4]

$$\hat{\pi}(x) = \frac{\exp(\lambda u(x; \pi))}{\int_y \exp(\lambda u(y; \pi))}. \tag{3}$$

For $\lambda = 0$, this reduces to uniform randomization; in the limit as $\lambda \to \infty$, best responses are chosen with probability one.

Applying this to the Tullock contest game, let $P_\pi(X)$ be the probability that the sum of the other $N - 1$ players' expenditures is equal to X, assuming that mixed strategy π is played. Then, the expected payoff to spending x is

$$u(x; \pi) = \omega - x + V \int \frac{x}{x + X} \, dP_\pi(X), \tag{4}$$

[3] Here and subsequently we consider the case of symmetric play, and therefore omit player-specific subscripts.

[4] The standard derivation of the logit choice rule requires the set of choices be discrete. For convenience, we follow Anderson et al. (1998b,a) and others in working with a continuous strategy space in parts of the exposition. The application of the model is done in a discrete space, with sums replacing integrals as appropriate.

for $x > 0$, and $u(0\,;\pi) = \omega$. Differentiating with respect to x gives

$$\frac{\partial u(x;\pi)}{\partial x} = -1 + V \int \frac{X}{(x+X)^2} dP_{\pi}(X) \tag{5}$$

The logit decision distribution implies

$$\hat{\pi}(x) = K \exp(\lambda u(x;\pi)) \tag{6}$$

with K being the normalization constant to ensure integration to unity. Differentiating with respect to x,

$$\frac{d\hat{\pi}(x)}{dx} = K\lambda \frac{\partial u(x;\pi)}{\partial x} \exp(\lambda u(x;\pi)) \tag{7}$$

$$= \lambda \pi(x) \frac{\partial u(x;\pi)}{\partial x} \tag{8}$$

$$= \lambda \pi(x) \left[-1 + V \int \frac{X}{(x+X)^2} dP_{\pi}(X) \right]. \tag{9}$$

Proposition 1. *For any conjectured distribution of expenditures by other players, for any $\lambda > 0$, the logit distribution of expenditures is single-peaked with a unique modal choice.*

Proof. The integral in (9) is strictly decreasing in x. Therefore, the term in square brackets is strictly decreasing in x, and therefore can equal zero for at most one value of x. Because $\lambda > 0$ and $\pi(x) > 0$, (9) can equal zero for at most one value of x. If there exists $x > 0$ which makes (9) zero, that is the modal choice which maximizes the density π. When (9) is negative for $x = 0$, then $x = 0$ is the modal choice. □

The introduction of a noisy best response relaxes one feature of Nash equilibrium. To proceed further, we must formulate the beliefs to which a player is (noisily) responding.

2.2.2. Cognitive hierarchy

In a Nash equilibrium, players' beliefs correctly anticipate the play of others. This assumption of correct beliefs is carried through to its statistical extension, quantal response equilibrium. Such fixed-point assumptions are analytically convenient, but implausible as a procedural model of how players think strategically in games. This has led to interest in models which dispense with such mutual consistency assumptions.

One active family of such models are "level-k" models (Stahl and Wilson, 1995). In standard level-k models, some fraction of players (at level 0) are assumed to play the game naïvely, usually modeled as choosing a strategy at random. A player at level 1 chooses a best response to the level 0 strategy; a player at level 2 chooses a best response to the level 1 strategy, and so on. In this most commonly used version, the notion of precise best response is retained, while removing the assumption of fixed-point reasoning. We generalize this by permitting noisy (logit) responses at each level, in a version of the cognitive hierarchy of Camerer et al. (2004).

Suppose there is a population of players. Each player is of one type $k \in \mathbb{Z}_+$, where k is the number of steps of reasoning the player undertakes. The proportion of players of type k in the population is $f(k)$. The behavior rules for each type are built up iteratively. Type $k = 0$ randomizes uniformly over all actions. For types $k > 0$, players are assumed to be overconfident, believing that all other players are using strictly fewer steps of reasoning. Therefore, their beliefs about the types of other players are truncated; player k's belief about the proportion of type $h < k$ in the population is $g_k(h) = f(h)/(\sum_{t=0}^{k-1} f(t))$ for $h < k$, and $g_k(h) = 0$ for $h \geq k$. Players have correct beliefs about the play of lower types.

In addition to different steps of strategic sophistication, we also permit players to make noisy responses. Specifically, for all types $k > 0$, we assume that players of that type play the logit best response to the distribution of play of lower types, with precision parameter λ. Finally, we follow Camerer et al. (2004) by assuming that the distribution f of types follows a Poisson distribution with mean τ. With this assumption, we can then compute the prediction of the model by taking a mixture of the individual expenditure distributions, weighted by the Poisson distribution probabilities.

Let $CH(\tau, \lambda)$ denote the logit-Poisson cognitive hierarchy prediction with τ expected levels of reasoning, and logit precision parameter λ.

2.2.3. Quantal response equilibrium

Another way to close the model is to assume a fixed point, i.e., that (3) is satisfied with $\hat{\pi} = \pi$. This leads to the logit quantal response equilibrium (QRE) concept of McKelvey and Palfrey (1995). For $\lambda = 0$, logit QRE generates uniform randomization across all expenditures; as $\lambda \to \infty$, the set of logit QRE converges to a subset of the Nash equilibria of the game. We will write $QRE(\lambda)$ to denote the logit QRE prediction with parameter λ.

The use of the logit rule for specifying choice probabilities means that in a majority of applications, closed-form expressions for logit QRE are not known.[5] Applications of QRE therefore rely on numerical computation to establish predictions. Turocy (2005) demonstrates an efficient, practical algorithm for computing a branch of the set of logit QRE; a reference implementation of this algorithm is available in Gambit (McKelvey et al., 2013).

2.2.4. Quantitative predictions

We now give an overview of the quantitative predictions of the QRE and CH models in the setting of our experiment, in which $\omega = 1200$ and $V = 1000$. Fig. 1 plots the distribution of individual expenditures for the group sizes $N = 2$, $N = 4$, and $N = 9$ used in our experiment, with $\lambda = \{0.5, 1.0\}$, and $\tau = \{1.0, 2.0, 5.0, \infty\}$. For each group size, the Nash equilibrium prediction is plotted as a vertical line for reference. The similarity of the $\tau = 5.0$ and $\tau = \infty$ plots illustrates that CH and QRE make essentially indistinguishable predictions, for these values of λ, already when there are on average five levels of reasoning being used.

Rogers et al. (2009) have shown that QRE and cognitive hierarchy models with deterministic best responses can each be thought of as special cases of a more general concept they call truncated heterogeneous quantal response equilibrium. Fig. 1 relates the predictions of logit QRE and logit CH directly. The logit CH model limits to logit QRE in any game if λ is sufficiently small.[6] The payoff structure of the game determines how large λ can be without destroying this convergence. Below we use the Tullock contest game with a discretized strategy space for our model fits. When $N = 2$, convergence holds for all λ, even for finely discretized strategy spaces. For $N = 4$ and $N = 9$, convergence holds for all λ when the discretization of the strategy space is not too fine. Direct calculation verifies that the discretization we use in our estimations is sufficiently coarse such that CH limits to QRE for all λ.[7]

Relative to the QRE baseline, the CH model accommodates distributions of expenditures with thicker right tails and with modes at lower levels of expenditure. Intuitively, the best response to uniform randomization, the step-0 behavior, is to expend very small (possibly zero) amounts.[8] Therefore, the step-1 logit response will have its mode at or near zero. When τ is small, a large proportion of the population is assumed to be step-0 or step-1, so these characteristics of those types' decision distributions are therefore reflected in the aggregate prediction.

Fig. 2 plots how average individual and group expenditure changes in QRE as a function of λ and the group size.[9] Average individual expenditures start at one-half the endowment and converge to the Nash equilibrium prediction from above in this setting, as observed previously by Sheremeta (2010). With larger groups, this convergence in individual expenditure is sufficiently slow that at the group level, total expenditure can exceed the value of the prize, often by substantial amounts.

The calculations illustrate that if λ is independent of the number of players, then the presence of heterogeneity alone could account for a significant increase in overdissipation when the number of players increases. However, there is no reason to assume λ is independent of the number of players, and indeed we estimate it separately for each group size. The models therefore allow us to distinguish how much the change in overdissipation can be attributed to the fact that best responses are imprecise in the first place, as opposed to changes in the precision of best responses as the number of other players is varied.

3. Experimental design

We report on a total of 9 experimental sessions, conducted at the Pittsburgh Experimental Economics Laboratory (PEEL) using subjects recruited from the participant pool maintained by the laboratory. The group size was varied across sessions, with three sessions conducted with each group size $N = 2$, $N = 4$, and $N = 9$. Cohort sizes ranged from 12 to 22 subjects.

At the beginning of each session, the instructions (see Appendix A) were read aloud. For comparability with the literature, the instructions followed the standard convention of using lottery terminology in presenting the game to the subjects. After the instructions were read and clarifying questions answered, subjects completed a questionnaire to check their understanding. In particular, the comprehension questionnaire verified each subject's ability to compute the expected earnings arising from given combinations of the subject's own expenditure and expenditures by other subjects in the same group.

In each session there were 10 iterations of the game. In each period, subjects were randomly rematched into groups of size N. Each subject was given an endowment $\omega = 1200$ tokens in each period. Subjects simultaneously selected a number of tokens between 0 and 1200 to spend in competition for a prize worth $V = 1000$ tokens. Feedback on each round was given in two stages. After the token expenditures were selected, but before realizing the random outcome, each subject viewed a

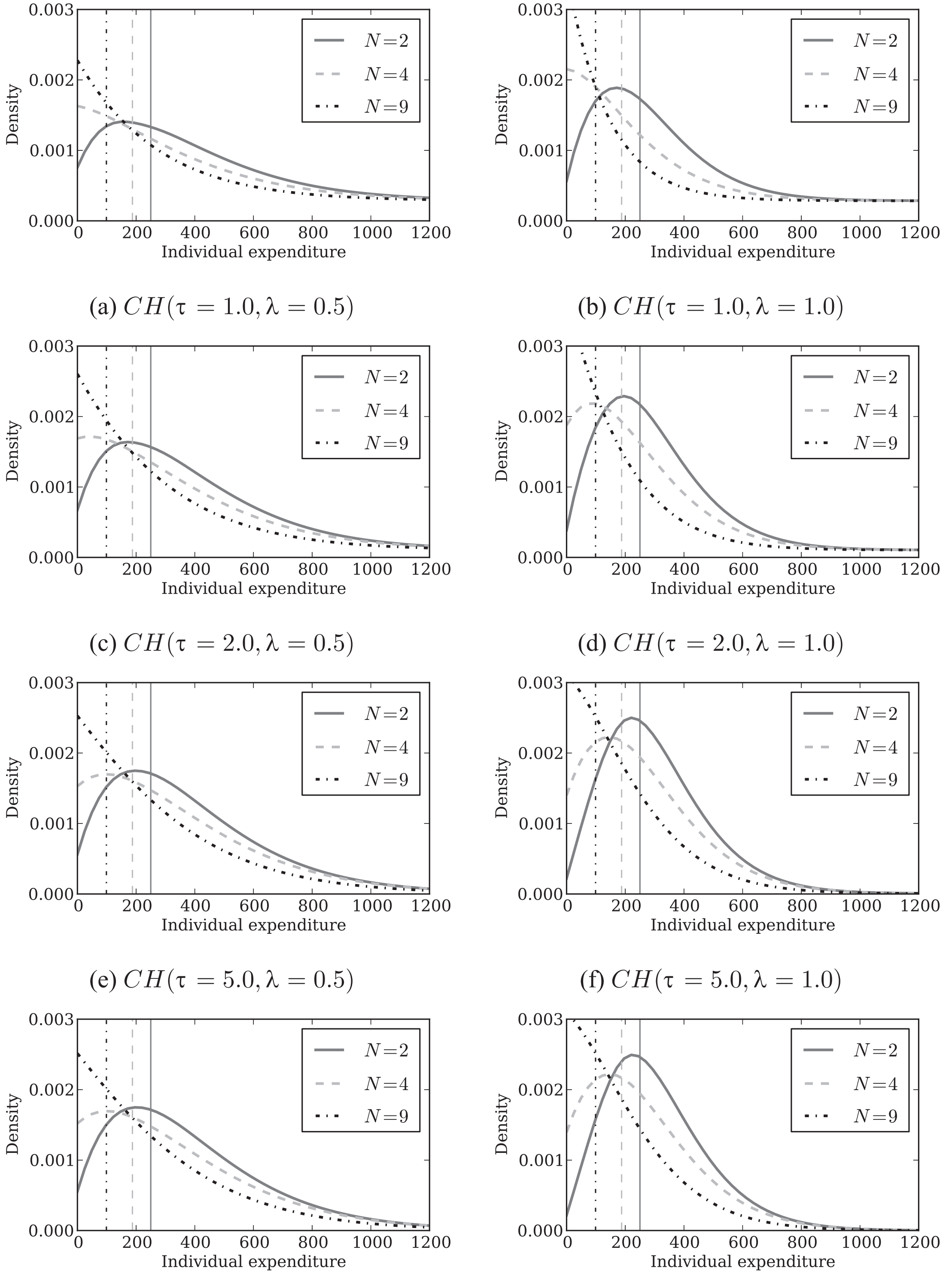
[5] Two exceptions are Anderson et al. (1998a), who obtain closed-form expressions for the QRE of the related all-pay auction, and Anderson et al. (1998b), who provide closed-form QRE for some classes of public goods games with linear or quadratic payoffs. The technology used in those papers to establish the closed-form solutions is defeated by the dependence of the integral in (9) on the full distribution of the sum of other players' expenditures, which appears in the denominator of the winning probability.

[6] This is implied by the result in McKelvey and Palfrey (1995) which shows that the logit mapping defined by (3) is a contraction when λ is small. Therefore, iterative application of the mapping from any initial distribution of play converges to the unique QRE for that λ.

[7] When the strategy space is a sufficiently fine grid, and λ and N are large enough, $CH(\tau, \lambda)$ instead exhibits cyclical behavior as τ is increased. Intuitively, this occurs because the Nash equilibrium is unstable under perturbed best reply dynamics when N is larger.

[8] This is because our endowment is at least as big as the value of the prize. If the endowment were sufficiently small this would no longer necessarily be true. This is why, as Sheremeta (2010) notes, QRE can capture the effect of the size of the strategy space on the aggressiveness of expenditures.

[9] The predictions of the CH model are similar and are therefore omitted.



(a) $CH(\tau = 1.0, \lambda = 0.5)$

(b) $CH(\tau = 1.0, \lambda = 1.0)$

(c) $CH(\tau = 2.0, \lambda = 0.5)$

(d) $CH(\tau = 2.0, \lambda = 1.0)$

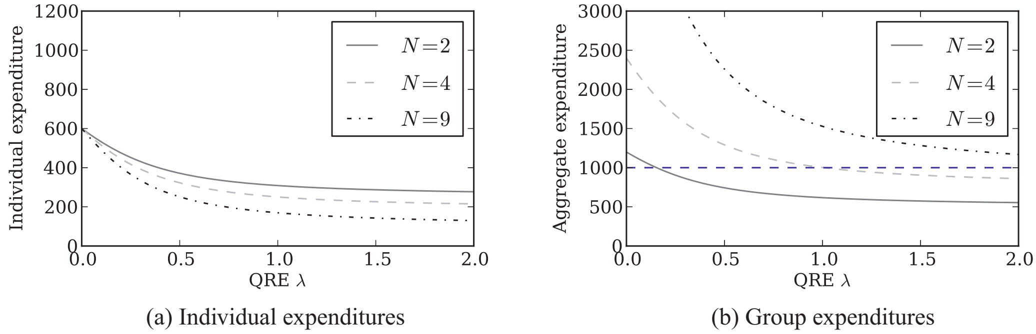
(e) $CH(\tau = 5.0, \lambda = 0.5)$

(f) $CH(\tau = 5.0, \lambda = 1.0)$

(g) $\lim_{\tau \to \infty} CH(\tau, \lambda = 0.5) = QRE(\lambda = 0.5)$

(h) $\lim_{\tau \to \infty} CH(\tau, \lambda = 1.0) = QRE(\lambda = 1.0)$

Fig. 1. Comparison of CH and QRE model predictions. The vertical lines indicate the Nash equilibrium expenditure for each group size *N*.



(a) Individual expenditures (b) Group expenditures

Fig. 2. Individual and group expenditures predicted by QRE, as a function of λ and the group size. The horizontal dashed line in total expenditures indicates full dissipation of the surplus.

Table 1
Mean expenditures per session, by treatment.

Treatment	Mean expenditure by session			Mean	Nash
$N=2$	351	339	287	331	250
$N=4$	277	377	254	300	188
$N=9$	380	251	348	326	99

screen reporting his/her own chosen expenditure, the sum of the expenditures of all other subjects in the same group, and the probability with which (s)he would win the contest. Then, a subsequent screen displayed the realization of the (random) lottery outcome for that subject, and the total income in the round, computed as the endowment, minus the expenditure, plus any income from winning the contest. No subject IDs were reported in the computer interface, nor did subjects have access to any history of play besides their own experience.

At the end of the 10 periods, one of the 10 periods was selected at random, and subjects' earnings for this portion of the experimental session were determined by the selected period with the exchange rate 100 tokens = $1.

After the 10 periods of Tullock contest games, the experimental session continued with 40 rounds of unrelated games. The overall length of each session was about two hours, with the contest portion comprising under an hour on average. In addition to their decision-contingent earnings, subjects received a $5 show-up fee.

4. Results

Result 1. Average expenditure exceeds the Nash equilibrium prediction for all group sizes. The average expenditure does not differ significantly across group sizes. There is weak evidence for learning over time only when $N=9$.

Support. Table 1 provides mean expenditures by session for each treatment. In each session, mean expenditures exceed the Nash prediction, replicating the standard result in the case where the endowment is at least as large as the value of the prize. Notably, the expenditures do not appear to be sensitive to the group size.

Fig. 3 plots the mean expenditure by period, aggregated across sessions for each group size. Even though expenditures exceed Nash predictions by large margins, there is little visual evidence of a significant convergence toward Nash equilibrium. There is no clear ordering of group sizes at any stage of the experiment.

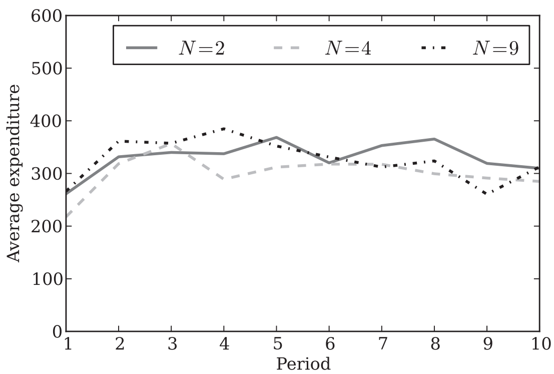


Fig. 3. Mean expenditure by period, aggregated across sessions, for each group size.

Table 2
Panel estimation of group-size effects on the amount of expenditure and overdissipation. Treatment $N=4$ is taken as a baseline. Model includes a random effects error structure, with the individual subject as the random effect; standard errors are clustered at the session level. Robust standard errors are reported. Key hypothesis tests regarding the coefficients are reported, with their p-values for the two-sided alternative.

Dependent variable: x_{it}		Coefficient	Std err	Test	p-Value
β_0	Baseline $N=4$	293.14	31.98	$\beta_0 = 188$	0.001
$\beta_{N=2}$	Dummy for $N=2$	21.68	37.21	$\beta_0 + \beta_{N=2} = 250$	0.0007
				$\beta_{N=2} = 0$	0.560
$\beta_{N=9}$	Dummy for $N=9$	56.21	38.55	$\beta_0 + \beta_{N=9} = 99$	<0.0001
				$\beta_{N=9} = 0$	0.145
β_t	Period number baseline	1.33	8.56	$\beta_t = 0$	0.876
$\beta_{t\times N=2}$	Period interaction $N=2$	1.54	9.15	$\beta_t + \beta_{t\times N=2} = 0$	0.375
$\beta_{t\times N=9}$	Period interaction $N=9$	-5.54	8.83	$\beta_t + \beta_{t\times N=9} = 0$	0.053

To formalize these observations, we conduct a regression, with individual expenditure as the dependent variable. We use the $N=4$ treatment as the baseline, as this is the most common group size to date in the literature on Tullock contest experiments, and define dummy variables for the $N=2$ and $N=9$ treatments. We also include the period number as a time trend, and interact this with the treatment dummies. Following e.g. Sheremeta (2010), we use a per-subject random effect, and cluster standard errors at the level of the session. Results of this estimation are reported in Table 2.

The table includes the key hypothesis tests and corresponding p-values. We test whether average expenditures are equal to Nash for that treatment; these hypotheses are rejected with p-values of 0.001 or less. We also test formally whether the treatment dummy variables are significantly different from zero; we cannot reject these hypotheses at the 10% level, and therefore cannot conclude that mean individual expenditure is sensitive to the number of players.

Turning to adjustment over time, we test for a non-zero time trend in each treatment. We cannot reject the null hypothesis of no trend in the $N=2$ or $N=4$ treatments at the 10% significance level, but there is weak support for a decreasing trend in the $N=9$ treatment (p-value 0.053). Nevertheless, average play remains far from Nash even after 10 iterations. The regression predicts an average expenditure of 345 in the first period versus 307 in the tenth period, still far from the Nash prediction of 99. The absence of a statistically identifiable time trend may be a function of our 10-period design; time trends reported in other studies (e.g. Sheremeta, 2010) are small enough to require many periods to have a statistically and economically significant effect. □

Because the average individual expenditure does not depend in a significant way on the group size, it follows that total expenditure increases rapidly as the group size increases. For group sizes $N=4$ and $N=9$, we find that total expenditure exceeds the value of the prize. For example, the average expenditure for a randomly constituted nine-person group in the $N=9$ treatment is \$29.35, for a prize worth only \$10.

Although there is little evidence that the number of players affects average individual expenditures systematically, there are differences in the distribution of expenditures as a function of group size.

Result 2. Average expenditures do not change significantly across treatments, but distributions of expenditures do. In larger groups, the distribution is more dispersed. The logit cognitive hierarchy model captures the qualitative features of the distribution of expenditures.

Support. Fig. 4 plots the kernel-smoothed distribution of individual expenditures for each group size. These distributions are qualitatively consistent with the predictions of the CH and QRE models in Fig. 1, in that the peak of each distribution falls below the Nash prediction. In addition, the frequency with which players effectively choose not to enter the contest, by selecting a zero or very small expenditure, is increasing in the group size, consistent with CH and QRE. In larger groups, the modal expenditure is lower and the distribution is flatter, with a more substantial right tail. These two effects roughly balance each other, leaving average expenditure approximately unchanged.

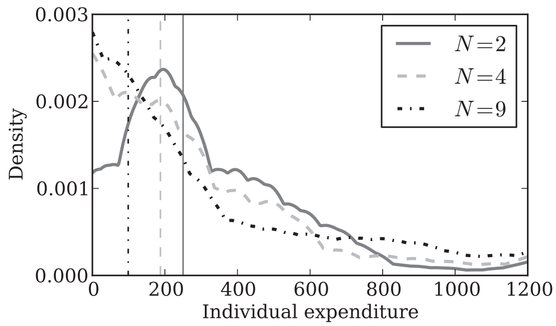


Fig. 4. Kernel density comparison of distribution of expenditures by group size, all 10 periods. Kernel density estimates computed by STATA kdens package, accounting for boundary corrections, using Silverman bandwidth selection.

Table 3
Summary of fitted values of CH and QRE models, using all data from all periods. Q measures quality of fit, with $Q = 0$ corresponding to the uniform density being the best fit, and $Q = 1$ the case where the model predicts the observed frequencies perfectly.

N	Uniform	Maximum	Cognitive hierarchy			QRE		
	$\log L$	$\log L$	$(\lambda^{CH}, \tau^{CH})$	$(\log L)^{CH}$	Q^{CH}	λ^{QRE}	$(\log L)^{QRE}$	Q^{QRE}
2	−1282.47	−1047.15	(0.779, 2.52)	−1063.19	0.932	0.642	−1067.40	0.914
4	−1333.77	−1119.63	(0.772, 2.18)	−1132.79	0.938	0.568	−1139.55	0.907
9	−1385.07	−1204.70	(0.430, 1.52)	−1223.28	0.897	0.318	−1227.28	0.875
Sum	−4001.31	−3371.48	–	−3419.26	–	–	−3434.23	–
Pooled	–	–	(0.618, 2.01)	−3450.50	0.875	0.471	−3462.96	0.855

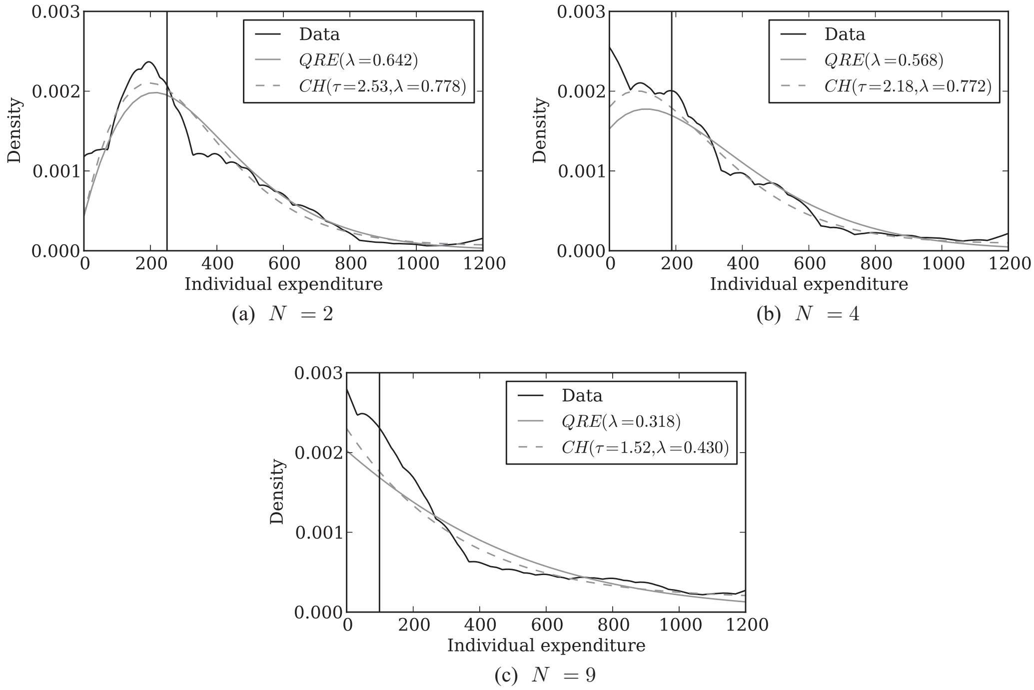


Fig. 5. QRE and CH model best fits compared to distribution of subject expenditures. In each figure the vertical line indicates the Nash equilibrium prediction.

We follow the standard set by McKelvey and Palfrey (1995) and Rogers et al. (2009) in estimating the values of the parameters λ and τ from the experimental data.[10] We discretize the game with 13 expenditure levels $k = 0, 100, \ldots, 1100, 1200$.[11] For each k, individual expenditures in $[k - 50, k + 50)$ are binned and treated as the choice in the discretized game. In the absence of any evidence of a significant time trend, we pool all 10 periods.[12] Table 3 summarizes the estimated parameter values and log-likelihoods. Fig. 5 compares the distribution functions of the empirical data, and the distribution arising from the best-fit parameters for each model.

Our estimates for the mean numbers of levels of reasoning τ lie between approximately 1.5 and 2.5, which is broadly consistent with the stylized fact in the level-k literature that most players employ between one and three levels of reasoning, depending on the game. We estimate higher values of λ for the cognitive hierarchy models than with the QRE restriction. This is qualitatively consistent with the results of Golman (2011a,b), who shows that imposing a common value of λ in a QRE when agents in fact have heterogeneous precision parameters leads to a downward bias; that is, QRE estimates make players out to be less precise in their responses than they actually are. Rogers et al. (2009) show that a (noiseless) cognitive hierarchy model is observationally equivalent to a QRE model where players differ in their response precisions.

We construct a measure to capture the quality of the obtained fits. Both QRE and CH generate uniform randomization when $\lambda = 0$; therefore, any fit necessarily results in a log-likelihood at least as good as that of the uniform distribution. In the

[10] Our approach to fitting the model parameters is the same as that in Rogers et al. (2009). Our "CH" model differs from theirs in that ours has a stochastic best response at east step, whereas their step-1 and higher players best response precisely.

[11] A substantial majority of our observations occur exactly on multiples of 100, so this discretization of the game parallels what many participants did implicitly.

[12] The main results are not sensitive to using all 10 periods; similar conclusions are obtained looking only at the first 5 or last 5 periods.

Table 4
Evaluation of quality of fit of models from a strategic perspective.

N	Mean expenditure			Best response			BR payoff		
	Data	*CH*	*QRE*	Data	*CH*	*QRE*	Data	*CH*	*QRE*
2	331	336	345	211	213	212	1457	1445	1436
4	300	284	307	121	129	108	1234	1233	1227
9	326	325	321	0	0	0	1200	1200	1200

other extreme, the best possible log-likelihood would result if a model predicted the empirical distribution exactly. Let $\ln L_u$ be the log-likelihood associated with the uniform distribution, and $\ln L_m$ the log-likelihood associated with a perfect fit. We define a pseudo-R^2 measure Q as

$$Q = \frac{\ln L - \ln L_u}{\ln L_m - \ln L_u}. \tag{10}$$

We report values for this measure for all fits in Table 3. Fit quality is high by this measure for all treatments, formalizing the quality of the fit observable from inspection of Fig. 5. □

Result 3. The null hypothesis that parameters are constant across group sizes can be rejected for both the CH and QRE models. QRE can be rejected statistically in favor of CH, but the practical benefit of the improved fit is small.

Support. First we consider whether the parameter estimates for each model are independent of the group size. We fit each model with the restriction that the parameter(s) remain the same for all group sizes. The results are reported in the row labeled "Pooled" in Table 3. For CH, we reject the null hypothesis of equality of parameters using a likelihood ratio test (log-likelihood −3450.50 for the pooled model versus −3419.26 for the non-pooled, with 4 degrees of freedom; p-value $\approx 10^{-13}$). For QRE, we also reject the null of equality of parameters (log-likelihood −3462.96 for the pooled versus −3434.23 for the non-pooled, with 2 degrees of freedom; p-value $\approx 10^{-13}$).[13]

Because QRE is a special case of CH for the discretization we have chosen, we test the null hypothesis of QRE against the alternative of CH, for each group size. We can statistically reject QRE in favor of CH using a likelihood ratio test, with p-values 0.004 for $N = 2$, 0.0002 for $N = 4$, and 0.005 for $N = 9$.

Does the improved fit from using the CH model have a practical benefit? Table 4 compares the empirical data, the best-fit CH model, and the best-fit QRE model, in three quantities relevant to a potential player in the game. Both the CH and QRE predict the mean individual expenditure closely. The best responses against the best-fit CH and QRE distributions are also very close to the best response against the empirical data. That is to say, if a player conjectured play to be according to the best-fit CH or QRE distribution and best responded to that distribution, their play would be very close to the best response if they knew the empirical distribution exactly.

The financial consequences of this optimization error are captured in the last group of columns. This group presents the expected earnings of a player playing the best response derived from the CH or QRE distribution, respectively, against the empirical distribution. For $N = 2$ and $N = 4$, playing the best response to CH gives only slightly higher earnings than the best response to QRE; for $N = 9$, the play is sufficiently aggressive that the best response is zero expenditure. The practical benefit of the CH model, measured in payoff terms, is negligible relative to QRE. □

Mathematically, λ expresses the relative influence that expected monetary payoffs have on choices, as opposed to the effects of other, unobserved random influences. Lower values of λ correspond to logit choice distributions where the unobserved payoff shocks have greater variance, and therefore greater influence on observed choices. Because λ is denominated in payoff terms, the estimates are already capturing any differences in the expected payoff consequences of "errors" as N is varied. Therefore, the fact that λ estimates decrease in N for both models cannot be attributed to changes in the optimization premium due to group size. In addition, the feedback structure across treatments was the same, providing total expenditure by others in the group; in the Tullock contest, what matters for determining best responses is the anticipated total expenditures and not per se the number of other participants.

The logit choice component of the model is statistical rather than behavioral in nature. There is no underlying theory or procedural description as to how λ arises in practice. We therefore look at behavior at a more micro level to examine the underpinnings of our estimates of λ more closely. One interpretation of smaller values of λ is that either subjects are less able to identify how to adjust behavior to improve expected earnings, or are less interested in doing so. We can ask whether there is any evidence showing that subjects tend to adapt in directions which improve expected earnings. The learning direction theory of Selten and Buchta (1994) is a simple heuristic model in which agents adjust their choices in response to recent experience. Learning direction theory asserts that, given an action profile $\{a^t\}$ played in period t, subject i will choose an action a_i^{t+1} in the subsequent period which is in the direction toward the best response against the previous-period choices of others, $\{a_{-i}^t\}$.

13 Here we are comparing the respective best-fit models.

Table 5
Summary statistics on directions of period-by-period adjustment relative to best response.

	$N=2$	$N=4$	$N=9$
Total observations	450	468	486
No change	176	177	179
% with no change	39.1	37.8	36.8
Toward best response	178	169	151
Away from best response	96	122	156
% changes toward best response	64.9	58.1	49.2

Table 6
Logistic panel regressions capturing determinants of decisions whether to adjust and the direction of adjustment. A per-subject random effect is included.

Dependent variable: $I(x_{i,t} \neq x_{i,t-1})$	β	SE	Odds
(a) Determinants of whether expenditure is changed from previous period			
$I_{N=4}$	−0.329	0.421	0.720
$I_{N=9}$	−0.732*	0.421	0.481
Period number	−0.167***	0.029	0.846
Ex-post optimization error	0.0024***	0.0004	1.002
Won in previous period	−1.601***	0.368	0.201
Observations	1404		
χ^2	108.76***		

Dependent variable: adjustment toward BR	β	SE	odds
(b) Determinants of whether expenditure is adjusted towards best response, conditional on an adjustment being made			
$I_{N=4}$	−0.491**	0.213	0.612
$I_{N=9}$	−1.075***	0.231	0.341
Period number	0.028	0.029	1.028
Ex-post optimization error	0.0021***	0.0004	1.002
Won in previous period	−0.027	0.198	0.973
Observations	872		
χ^2	38.24***		

* 10% significance level.
** 5% significance level.
*** 1% significance level.

Result 4. Participants adjust decisions in the direction of the best response less consistently as the group size N increases.

Support. Table 5 summarizes, for each group size, how often subjects adjust expenditures in the direction of the myopic best reply to the expenditures by the other players in the previous period. For all group sizes, in approximately 40% of instances the subject chose the same expenditure as the previous period. Among instances in which the subject elected to adjust their expenditures, there is an effect of group size. When facing only one other player, subjects making an adjustment to their expenditures did so in the direction of the best response 64.9% of the time. This proportion decreases as the group size increases, with only 49.2% of the adjustments with $N=9$ going in the direction of the best response.

We formalize these observations by performing two logistic panel regressions, with a random effect by participant. The first regression examines determinants of whether a participant changes their expenditure level from period $t-1$ to period t. Let x_{it} be the expenditure of participant i in period t. The dependent variable is 1 if and only if $x_{it} \neq x_{i,t-1}$. We include as independent variables dummy variables for the group size (with $N=2$ taken as the baseline and omitted), the period number, whether the participant won in period $t-1$, and a measure of the ex-post optimization error from period $t-1$. This latter measure is computed as follows. Let $o_{i,t-1}$ be the sum of other participants' expenditures in participant i's group in period $t-1$. Given this, we compute $x^\star(o_{i,t-1})$, the (myopic) best response to the behavior of the other participants in the previous period. The ex-post optimization error is then $|x_{i,t-1} - x^\star(o_{i,t-1})|$.

Table 6a reports the results of this regression. We can conclude that there is a treatment effect when comparing $N=2$ and $N=9$; a participant is significantly less likely to change his expenditure in groups with $N=9$ versus $N=2$. This contrasts with the aggregate frequencies reported in Table 5, which shows adjustment frequencies to be roughly constant by treatment. The intuition for this is that in the $N=9$ treatment, the average ex-post optimization error is much larger. In the regression we find the size of the ex-post optimization error is strongly significant in determining whether the participant makes an adjustment in the subsequent period. The two effects therefore operate in opposite directions: adjustment is less likely conditional on the optimization error in larger groups, but the optimization error is on average much larger. In this regression, we also find evidence that there is more adjustment in early periods than later ones, and that participants are significantly more likely to adjust after losing in the previous period.

In the second regression, we focus on the instances in which the participant did change their behavior from the previous period. We conduct a logistic panel regression with the dependent variable being 1 if and only if the participant changes their expenditure in the direction (in the strategy space) of the best response. We use the same set of regressors as in the first model. Table 6b presents the parameter estimates. We find that the group size has a significant effect; adjustment is significantly less likely to be in the direction of the best response when the groups are larger. The size of the optimization error is again significant; the larger the error, the more likely the direction of change is toward the best response. The period number and whether the participant won in the previous period are not significant. □

5. Conclusion

In our experiment, we replicate the result that average individual expenditure is significantly above the Nash equilibrium prediction in Tullock contests, where the endowment is set at least as large as the value of the prize. Our series of experiments isolates the effect of group size, holding fixed the value of the prize and endowment, with an across-subjects design. Our 10-period sessions allow for some opportunity for subjects to learn and adjust from previous experience. We find that individual expenditures are not very responsive on average to group size, but that distributions of individual expenditures do depend on group size. Because individual expenditures do not decline in larger groups, total expenditure exceeds the value of the prize for four-player and nine-player groups.

Our logit cognitive hierarchy model allows us to relax independently the fixed-point and best-response assumptions of Nash equilibrium, nesting both cognitive hierarchy and logit QRE as special cases. Viewed through the lens of this model, our data imply that for larger groups, the rationality parameters in our model are farther from perfect rationality, or, equivalently, that the expected financial earnings in the game are less capable of organizing behavior. Therefore, the excess expenditure in larger groups arises for two reasons: because there is heterogeneity in decisions in the first place, and because decisions are less clearly tied to financial incentives in the larger group settings.

Our result on average individual expenditure contrasts with the preceding literature, as well as the prediction of Nash equilibrium. The most directly comparable prior study is that of Anderson and Stafford (2003). Their design was within-subjects and one-shot. Each participant was given a set of six scenarios and asked to formulate expenditures for each of them. After all decisions were made, one of the scenarios was chosen at random, and the decisions of the participants were played out. One possible reason for the contrast in results is that their instrument more directly suggests that the expenditure decision should depend on the number of other participants, as participants needed to contemplate scenarios with various numbers of opponents. In our across-subjects design, participants did know the number of opponents they faced, but were not asked to contemplate other group sizes. Similarly, participants in the experiment of Morgan et al. (2012) had experience in contests with different numbers of participants over the course of a session.

Nevertheless, while the cue on the number of players in the contest was less direct in our setting, this information “should” have fed through over time, with players better-responding to aggressive play by reducing expenditure. A novelty in our results is our linkage between lower rationality parameters in our logit cognitive hierarchy model and period-by-period adjustments. Subjects become less likely to change their expenditure levels in larger groups, and furthermore, conditional on making an adjustment, that adjustment is less likely to be in the direction of the better response. This suggests that our subjects found the nine-player group to be a significantly more challenging learning environment.

The implications of the theoretical and experimental results on noisy behavior in contests depend on the domain of application of the model, and the objectives of the contest designer. In the case of political rent-seeking or research and development races, if expenditures represent efforts which are fully dissipated in the contest and have no spillover, scrap, or other socially redeeming value, then our results are negative; it would be socially optimal to avoid the existence of the contest altogether, since the net surplus would be negative. However, if the contest is being used, for example, in a firm as one method to encourage worker effort, then these models predict contests might be significantly more effective in accomplishing this objective than Nash equilibrium would indicate.

Following many previous studies, we presented the contest model in our instructions using a lottery or raffle frame. The use of lotteries as fundraising vehicles is attested as far back as the construction of the Great Wall of China. Large-scale state-run lotteries exist, and are quite profitable, in many countries. Lotteries on this scale offer small chances of winning enormous, life-changing prizes, and therefore a “chance to buy hope.” (Clotfelter and Cook, 1989, 1990; Cook and Clotfelter, 1993) Lotteries also occur widely on much more modest scales as fundraisers for schools, churches, community groups, and the like. Laboratory experiments in the lottery frame abstract away from massive prizes, small probabilities, and the possibility of charitable motives. In this frame, models of noisy boundedly rational behavior predict lotteries can be profitable even in small groups, due solely to modest amounts of noise in decision-making.[14]

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jebo.2013.12.010.

[14] Because of our use of the lottery frame in the experimental procedures, it is perhaps in the case of such fundraising raffles that our experimental results have the most direct external validity.

References

Columbus, F. (Ed.), 2009. Game Theory: Strategies, Equilibria, and Theorems. Nova Science Publishers, Hauppauge, NY.

Anderson, L.R., Stafford, S.L., 2003. An experimental analysis of rent seeking under varying competitive conditions. Public Choice 115, 199–216.

Anderson, S.P., Goeree, J.K., Holt, C.A., 1998a. The all-pay auction: equilibrium with bounded rationality. Journal of Political Economy 106, 828–853.

Anderson, S.P., Goeree, J.K., Holt, C.A., 1998b. A theoretical analysis of altruism and decision error in public goods games. Journal of Public Economics 70, 297–323.

Bullock, D.S., Rutstrom, E.E., 2007. Policy making and rent-dissipation: an experimental test. Experimental Economics 10, 21–36.

Camerer, C.F., Ho, T.-H., Chong, J.-K., 2004. A cognitive hierarchy model of games. Quarterly Journal of Economics 119, 861–898.

Chowdhury, S.M., Sheremeta, R.M., Turocy, T.L., 2013. Overdissipation and overspreading in rent-seeking experiments: cost structure and prize allocation rules, Working paper.

Clotfelter, C., Cook, P., 1989. Selling Hope: State Lotteries in America. Harvard University Press, Cambridge, MA.

Clotfelter, C., Cook, P., 1990. On the economics of state lotteries. Journal of Economic Perspectives 4, 105–119.

Cook, P., Clotfelter, C., 1993. The peculiar scale economies of lotto. American Economic Review 83, 634–643.

Davis, D.D., Reilly, R.J., 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.

Dechenaux, E., Kovenock, D.J., Sheremeta, R.M., 2012. A survey of experimental research on contests, all-pay auctions, and tournaments. Chapman University, Working paper.

Fallucchi, F., Renner, E., Sefton, M., 2012. Information feedback and contest structure in rent-seeking games. University of Nottingham, CeDEx Working paper.

Fonseca, M.A., 2009. An experimental investigation of asymmetric contests. International Journal of Industrial Organization 27, 582–591.

Gneezy, U., Smorodinsky, R., 2006. All-pay auctions: an experimental study. Journal of Economic Behavior and Organization 61, 255–275.

Golman, R., 2011a. Homogeneity bias in models of discrete choice with bounded rationality. Journal of Economic Behavior & Organization 82, 1–11.

Golman, R., 2011b. Quantal response equilibria with heterogeneous agents. Journal of Economic Theory 146, 2013–2028.

Gronberg, T.J., Andrew, L.R., Turocy, T.L., Van Huyck, J.B., 2012. Are tax-financed contributions to a public good completely crowded-out? Experimental evidence. Journal of Public Economics 96, 596–603.

Haile, P., Hortacsu, A., Kosenok, G., 2008. On the empirical content of quantal response equilibrium. American Economic Review 91, 180–200.

Mago, S.D., Savikhin, A.C., Sheremeta, R.M., 2012. Facing your opponents: social identification and information feedback in contests, ESI Working paper.

McKelvey, R.D., McLennan, A.M., Turocy, T.L., 2013. Gambit: software tools for game theory. http://www.gambit-project.org

McKelvey, R.D., Palfrey, T.R., 1995. Quantal response equilibrium for normal form games. Games and Economic Behavior 10, 6–38.

Millner, L.E., Pratt, M.D., 1989. An experimental investigation of efficient rent-seeking. Public Choice 62, 139–151.

Morgan, J., Orzen, H., Sefton, M., 2012. Endogenous entry in contests. Economic Theory 50, 867–879.

Parco, J.E., Rapoport, A., Amaldoss, W., 2005. Two-stage contests with budget constraints: an experimental study. Journal of Mathematical Psychology 49, 320–338.

Potters, J.C., de Vries, C.G., van Winden, F., 1998. An experimental examination of rational rent seeking. European Journal of Political Economy 14, 783–800.

Rogers, B.W., Palfrey, T.R., Camerer, C.F., 2009. Heterogeneous quantal response equilibrium and cognitive hierarchies. Journal of Economic Theory 144, 1440–1467.

Schmidt, D., Sheremeta, R.M., Shupp, R.S., Walker, J., 2013. Resource allocation contests: experimental evidence. Indiana University, Working paper.

Selten, R., Buchta, J., 1994. Experimental sealed-bid first price auctions with directly observed bid functions. In: Budescu, D., Erev, I., Zwick, R. (Eds.), Games and Human Behavior: Essays in the Honor of Amnon Rapoport. Lawrenz Associates, Mahwah, NJ.

Selten, R., Chmura, T., 2008. Stationary concepts for experimental 2×2-games. American Economic Review 98, 938–966.

Sheremeta, R.M., 2010. Experimental comparison of multi-stage and one-stage contests. Games and Economic Behavior 68, 731–747.

Sheremeta, R.M., 2011. Contest design: an experimental investigation. Economic Inquiry 49, 573–590, http://dx.doi.org/10.1111/j.1465-7295.2009.00274.x.

Stahl, D.O., Wilson, P.W., 1995. On players' models of other players: theory and experimental evidence. Games and Economic Behavior 10, 213–2154.

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 University Press, College Station, TX, pp. 97–112.

Turocy, T.L., 2005. A dynamic homotopy interpretation of the logistic quantal response equilibrium correspondence. Games and Economic Behavior 51, 243–263.

Wärneryd, K., 2012. The evolution of preferences for conflict. Economics Letters 116, 102–104.