



Working Papers

www.cesifo.org/wp

Intention-Based Fairness Preferences in Two-Player Contests

Magnus Hoffmann
Martin Kolmar

CESIFO WORKING PAPER NO. 4104
CATEGORY 2: PUBLIC CHOICE
FEBRUARY 2013

An electronic version of the paper may be downloaded

- *from the SSRN website:* www.SSRN.com
- *from the RePEc website:* www.RePEc.org
- *from the CESifo website:* www.CESifo-group.org/wp

Intention-Based Fairness Preferences in Two-Player Contests

Abstract

We analyze equilibria of two-player contests where players have intention-based preferences. We find that players invest more effort compared to the case with selfish preferences and are even willing to exert effort when the monetary value of the prize converges to zero. As a consequence, overdissipation occurs if the value of the prize is sufficiently small.

JEL-Code: D030, D720, D740.

Keywords: contests, intention-based fairness preferences.

Magnus Hoffmann
School of Economics and Management
University of Technology Berlin
Straße des 17. Juni 135
10623 Berlin
Germany
magnus.hoffmann@tu-berlin.de

Martin Kolmar
Institute of Economics
University of St. Gallen
Varnbühlstrasse 19
9000 St. Gallen
Switzerland
martin.kolmar@unisg.ch

We would like to thank Georgios Katsenos and Hendrik Rommeswinkel for helpful suggestions. All errors are ours.

1. Introduction

Empirical tests of contest & tournament models with selfish players show that the predictive power of these models is relatively poor. Subjects in experiments tend to spend more than predicted, and overdissipation is frequently observed. Subjects are even willing to spent effort if the value of the prize is zero (see Dechenaux et al. (2012) for an excellent survey). Therefore, a growing literature analyzes contest behavior with non-standard preferences (see for instance Grund and Sliwka (2005), Fonseca (2009), and Lim (2010)). All these models, however, remain consequentialist in the sense that fairness exclusively depends on the outcomes or consequences of individual behavior. With the exception of Gill and Stone (2010) who develop a model of desert in tournaments that can be given a procedural interpretation, models combining consequentialist and procedural aspects of fairness played no role in the theoretical literature on contests so far.

We analyze how intention-based fairness preferences, as introduced by Rabin (1993), influence behavior in Tullock-type contests. Rabin's model is especially attractive because recent findings suggest that purely consequentialist models of social preferences are too narrow to explain human behavior and that individuals take intentions into consideration when making economic decisions (Konow (2003)). It is therefore interesting to see how the Rabin model stands the test of explaining behavior in contests.

2. The model

Consider a situation in which each of two players $i \in \{1, 2\}$ simultaneously exerts effort $x_i \in \mathbb{R}_+$ in order to win a prize of monetary value $V > 0$. The contest success function (CSF), $p_i : x_i \times x_j \rightarrow [0, 1]$, determines for any given value of the vector $\mathbf{x} = (x_i, x_j)$ player i 's probability of winning the prize (or player i 's share of the prize). It takes the form of a lottery contest (see Tullock (1980)):

$$p_i(\mathbf{x}) = \begin{cases} \frac{x_i}{x_i + x_j} & \text{for } \mathbf{x} \neq \mathbf{0}, \\ \frac{1}{2}, & \text{for } \mathbf{x} = \mathbf{0}. \end{cases} \quad (1)$$

The (expected) material payoff of individual i can then be defined as:

$$v_i(\mathbf{x}) = p_i(\mathbf{x})V - cx_i, \quad (2)$$

with $c > 0$ representing the marginal costs of effort.

Two features of the model by Rabin (1993) stand out. First, a player's payoff increases (decreases) if she believes that the other player intends to treat her kindly (unkindly). Second, if a player believes that the other player treats her kindly (unkindly), she will reciprocate and treat the opponent kindly (unkindly) as well. Intentions are inferred from actions according to second-order beliefs about effort choices where $x_i \in \mathbb{R}_+$ denotes player i 's effort choice, $y_j \in \mathbb{R}_+$ i 's belief about j 's effort choice, and $z_i \in \mathbb{R}_+$ i 's belief about j 's belief about i 's effort choice.

In order to quantify the utility from imputed intentions Rabin uses two *kindness* functions: $f_i(x_i, y_j)$ and $\tilde{f}_j(y_j, z_i)$. The first measures how kind i is to j , the second measures how kind i believes j to be. Define

(i) $v_j^h(y_j) \equiv \max_{x_i} v_j(y_j, x_i) = V - y_j$ as the maximum, (ii) $v_j^{min}(y_j) \equiv \min_{x_i} v_j(y_j, x_i) = -y_j$ as the minimum payoff, and (iii) $v_j^l(y_j)$ as the minimum among the efficient payoffs player j can get, given y_j . Since j 's payoff monotonically decreases in i 's effort, the latter payoff depends on player i 's best response to i 's beliefs about j 's effort choice, which is given by

$$X_i(y_j) = \begin{cases} \sqrt{\frac{V y_j}{c}} - y_j & \text{for } y_j \leq V/c, \\ 0 & \text{else.} \end{cases} \quad (3)$$

We thus define $v_j^e(y_j) \equiv v_j(y_j, X_i(y_j))$. A fair reference point, coined *equitable payoff*, is defined as $v_j^e(y_j) = (v_j^h(y_j) + v_j^l(y_j))/2$. Thus,

$$f_i(x_i, y_j) = \frac{v_j(y_j, x_i) - v_j^e(y_j)}{v_j^h(y_j) - v_j^{min}(y_j)} = p_j(y_j, x_i) - \frac{1}{2} \left(1 + p_j(y_j, X_i(y_j)) \right) \in [-1, 1/2]. \quad (4)$$

If we substitute actions in (4) by first-order beliefs (x_i by y_j) and first-order beliefs in (4) by second-order beliefs (y_j by z_i), we get

$$\tilde{f}_j(y_j, z_i) = \frac{v_i(z_i, y_j) - v_i^e(z_i)}{v_i^h(z_i) - v_i^{min}(z_i)} = p_i(z_i, y_j) - \frac{1}{2} \left(1 + p_i(z_i, Y_j(z_i)) \right) \in [-1, 1/2], \quad (5)$$

where $Y_j(z_i)$ is defined in accordance with (3). Note that in the contest framework the denominator in both kindness functions is positive by definition. Hence, $f_i(x_i, y_j) > 0$ ($\tilde{f}_j(y_j, z_i) > 0$) if and only if i treats j kindly (i believes that j treats her kindly) and *vice versa*.

Player i 's overall expected utility for given x_i, y_j , and z_i is defined as follows (see Rabin (1993)):

$$u_i(x_i, y_j, z_i) = v_i(x_i, y_j) + \tilde{f}_j(y_j, z_i) (1 + f_i(x_i, y_j)). \quad (6)$$

3. Equilibrium analysis

A solution of the game is a fairness equilibrium (FE):

Definition 1 (Fairness Equilibrium, Rabin (1993))

An intention-based FE for a two-player contest is a tuple x_i^F, x_j^F , such that

1. $x_i^F \in \underset{x_i}{\operatorname{argmax}} u_i(x_i, y_j, z_i)$,
2. $x_i = y_i = z_i$,

with $i \in \{1, 2\}$, and $i \neq j$.

Given that $\tilde{f}_j(y_j, z_i)$ is a parameter to i 's optimization problem and $p_j(y_j, x_i) \equiv 1 - p_i(x_i, y_j)$, (6) can be rewritten as follows:

$$u_i(x_i, y_j, z_i) = p_i(x_i, y_j) \left((V - \tilde{f}_j(y_j, z_i)) - c x_i + \tilde{f}_j(y_j, z_i) (1 - p_j(y_j, X_i(y_j))) \right). \quad (6')$$

Note that the third term of the above equation is independent of x_i . The formulation shows that from a behavioral point of view fairness preferences “perturb” the perception of the value of the prize in any interior equilibrium. Intuitively, the effective prize $(V - \tilde{f}_j(y_j, z_i))$ decreases in the kindness of player j .

Proposition 2

The lottery contest with fairness preferences defined by (6) has a unique fairness equilibrium $\mathbf{x}^F = (x_i^F, x_j^F)$ with

1.

$$x_i^F(c, V) = x_j^F(c, V) = \begin{cases} \frac{1+2V}{8c} & \text{for } 2V(4c-1) < 1, \\ \frac{1+32V^2 + \sqrt{1+64V^2}}{128cV} & \text{for } 2V(4c-1) \geq 1, \end{cases} \quad (7)$$

2. $x_i^F(c, V) = y_i^F(c, V) = z_i^F(c, V) = y_j^F(c, V) = z_j^F(c, V)$,

3. $f_i(x_i^F(c, V), y_j^F(c, V)) = \tilde{f}_j(y_j^F(c, V), z_i^F(c, V)) = -\frac{1}{4}$,

for $i \in \{1, 2\}$ and $i \neq j$.

Proof. Assume an interior equilibrium exists. If it exists it can be characterized by the first-order condition of (6'):

$$\frac{\partial p_i(x_i, y_j)}{\partial x_i}(V - \tilde{f}_j(y_j, z_i)) = c. \quad (8)$$

Note that this conjecture implicitly assumes that the effective prize is positive ($V > \tilde{f}_j(y_j, z_i)$). In this case (6) is strictly concave in x_i . Using (1), (3), (5), and the equilibrium conditions $x_i = y_i = z_i$, $x_j = y_j = z_j$, (8) becomes

$$\frac{\sqrt{c}x_i x_j (x_i + x_j) + \sqrt{V}x_j((2V-1)x_i + (2V+1)x_j)}{2\sqrt{V}(x_i + x_j)^3} - c = 0 \Leftrightarrow \frac{V}{c} \geq x_i, \quad (9a)$$

$$\frac{x_j(V(x_i + x_j) + x_j)}{(x_i + x_j)^3} - c = 0 \Leftrightarrow \frac{V}{c} < x_i. \quad (9b)$$

We have to analyze all possible permutations of the above cases.

Symmetric interior equilibria: In a symmetric and interior FE ($x_i^F = x_j^F = x^F > 0$), the equilibrium efforts are implicitly given by

$$\frac{1}{8} \left(\frac{2V}{x^F} + \sqrt{\frac{c}{x^F V}} \right) - c = 0 \Leftrightarrow \frac{V}{c} \geq x^F, \quad (10a)$$

$$\frac{1}{8} \left(\frac{2V}{x^F} + \frac{1}{x^F} \right) - c = 0 \Leftrightarrow \frac{V}{c} < x^F. \quad (10b)$$

The solution yields (7). Note that $2V(4c-1) < 1$ implies $y > V/c$, and thus $X(y) = 0$ (cf. (3)). If $2V(4c-1) \geq 1$, we find that $y \leq V/c$ and thus $X(y) = \sqrt{V y_j/c} > 0$. In both cases $\tilde{f}_j(y_j^F(c, V), z_i^F(c, V)) = -\frac{1}{4}$ so that the optimization problem is well defined (cf. eq. (8)).

Asymmetric interior equilibria: Next, we need to rule out that asymmetric equilibria exist with (i) $x_1 > V \wedge x_2 > V$, (ii) $x_1 < V \wedge x_2 < V$, (iii) $x_1 > V \wedge x_2 < V$, (iv) $x_1 < V \wedge x_2 > V$. The resulting system of FOCs cannot be solved analytically in these cases (cf. the polynomials in (9)). The proof of the non-existence of fairness equilibria in the above cases has been carried out using the software package Mathematica 8. The programming code can be sent upon request.

Corner equilibria: Two candidates for corner equilibria exist: $\mathbf{x} = (0, 0)$ and $\mathbf{x} = (0, x_j)$, with $x_j > 0$. In the former case, $\tilde{f}_j(0, 0) = 0$ so that both players' actions (as well as their first and second order

beliefs) are neutral. Then, however, the contest boils down to a standard two-player contest with selfish preferences (cf. eq. (2)) where $\mathbf{x} = \mathbf{0}$ can never be an equilibrium (see Nitzan (1994)). In case $\mathbf{x} = (0, x_j)$ player i believes that the other player is unkind, since $\tilde{f}_j(0, x_j) < 0$. This leads to a perception of the value of the prize $V - \tilde{f}_j(0, x_j) > V$: the contest resembles a standard contest with (possibly) asymmetric valuations, and it follows from the above mentioned literature that in this case $\mathbf{x} = (0, x_j)$ cannot be an equilibrium either. ■

In the FE both players perceive the other player as unkind and therefore act unkindly themselves. The reason why only unkindness is a mutually consistent belief becomes apparent if one takes a closer look at the equitable payoff $v_j^e(y_j)$. Payoffs larger (smaller) than $v_j^e(y_j)$ are considered as a signal for the kindness (unkindness) of the other player. Interestingly, $v_j^e(y_j) = (V + \sqrt{cy_jV})/2 - cy_j$, which is for $y_j > 0$ strictly larger than the payoff that would result if the players shared the rent equally or had identical winning probabilities ($V/2 - cy_j$). In an extreme case (if $y_j = 0$) player i is considered fair if and only if she leaves more than half of the prize for player j , which, given the CSF in (1), is unfeasible for all $y_i \geq 0$. It follows that even if player i invests zero effort ($y_i = 0$) the ascribed intentions are neutral. In situations where $y_j > 0$, the equity threshold increases beyond one-half, which implies that even with identical effort levels the players ascribe negative intentions to the other player.

Lab experiments on contests consistently produce several anomalies compared to the reference model with selfish preferences (see Dechenaux et al. (2012)). We show that the following anomalies can be explained by intention-based fairness preferences.

Corollary 3 *We find that the following holds in the FE:*

1. **Overspending.** *Players spend more effort compared to the NE, i.e.*

$$X^F(c, V) \equiv \sum_{i=1}^2 x_i^F(c, V) > \sum_{i=1}^2 x_i^N(c, V) \equiv X^N(c, V) \quad \forall c > 0, V > 0. \quad (11a)$$

2. **Overdissipation.** *Aggregate effort exceeds the monetary value of the prize if V is sufficiently small, i.e.*

$$cX^F(c, V) > V \Leftrightarrow V < \frac{1}{2\sqrt{2}}. \quad (11b)$$

3. **Joy of winning.** *As the monetary value of the prize converges to zero, aggregate spending stays strictly positive, i.e.*

$$\lim_{V \rightarrow 0} X^F(c, V) = \frac{1}{8c}. \quad (11c)$$

Proof. Note that the aggregate spending in the selfish NE (cf. eq. (2)) equals $V/(2c)$. Comparing this to (7) leads to (11a). (11b) as well as (11c) follow immediately from (7). ■

Figure 1 shows the difference between the effort invested in a NE (X^N) and the FE (X^F) contingent on V for $c = 1$. The dotted graph represents V , the dashed graph represents aggregate effort with selfish, and the bold graph aggregate effort with fairness preferences. For $V < \frac{1}{6}$ individual effort exceeds the value of the prize such that the regime shift from Proposition 1 occurs. If $V < 1/(2\sqrt{2})$ aggregate spending exceeds the value of the prize such that overdissipation exists up to this value. For increasing (but finite) values of V the difference between X^F and X^N is reduced but does not vanish. Overspending results

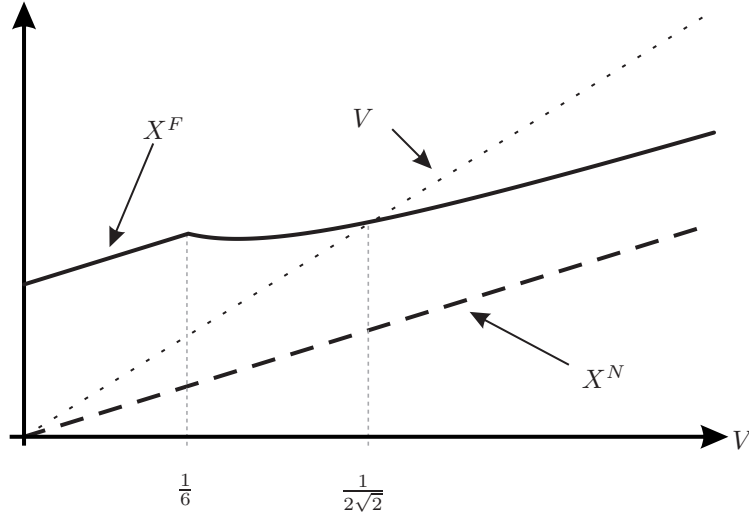


Figure 1
Equilibrium effort with selfish and fairness preferences

because both players regard their opponent as being unkind and act accordingly. However, this finding seems to be at odds with the construction of the psychological (dis-) utility which is normalized to be in $[-1, 1/2]$, and whose relative importance should therefore vanish if V becomes sufficiently large. It is in fact true that (4) and (5) are independent of the valuation V of the relevant player. The reason why this effect does not appear in our results is that (3) is a function of the *opponent's* valuation V and therefore enters the equitable payoff. Given that both players have identical valuations, the comparative-static experiment underlying Figure 1 therefore assumes a joint variation of V and is therefore different from the ones in Rabin (1993).

Overdissipation results because in any FE the normalized psychological utilities represented by the kindness terms become more relevant the smaller the material payoff is (cf. eq. (6)). Thus, if V is sufficiently small, the psychological utilities become relatively more important, such that eventually aggregate effort exceed the value of the prize; a property which does not exist with selfish preferences (see Ellingsen (1991)).

This line of reasoning taken to its extreme leads to (11c): Even if the value of the prize converges to zero, wasteful effort spending persists. This result can be viewed as an attempt to explain the frequent observation that subjects invest positive effort even if the monetary value of the prize is zero. Our model gives the idea of a “joy of winning” motive that has been brought forward to explain this type of behavior a specific meaning: for $V \rightarrow 0$ both players almost exclusively exert effort in order to retaliate effort spending by the competitor. The joy the players experience when winning a prize of (monetary) value zero comes from the fact that they avoid the bad feelings that would result from losing.

4. Discussion

Our analysis has shown that intention-based preferences do a reasonably good job in explaining empirical behavior in contests. The only pattern the model is not able to explain is overspreading (Dechenaux et

al. (2012)), i.e. the fact that the variance of observed efforts is large.

A possible way to extend the model to also explain overspreading (in addition to the obvious observation that overspreading can result from asymmetric information about preferences) is to allow for heterogeneity with respect to the perception of the equitable payoff, which turned out to be crucial for our results. The intuition for Proposition 1 has revealed that it is in fact this reference point that drives the result. We use Rabin's (1990) specification in our model, and its construction has a lot of intuitive appeal. However, it can be scrutinized if it is the only adequate and plausible reference point in a conflict situation like the one analyzed in this paper. In the end players know that they are in a situation of mutual conflict where the only possibility to profit is at the expense of the other player. Assuming negative intentions for any positive effort level therefore seems to be a rather strong feature of the Rabin model because it implies that the players consider only the fully efficient outcome with zero effort as one where intentions are not unkind. An alternative (and admittedly *ad hoc*) candidate payoff that would potentially qualify as equitable would be the selfish Nash outcome. The interpretation would be that players assume kindness of the other player if they receive more than in the selfish Nash case and *vice versa*. In a situation like this there is more room for mutually ascribed kindness, and it is in fact straightforward to check that kindness equilibria can exist in this case where players undercommit effort compared to the selfish Nash case. This line of reasoning points in the direction of situation-specific reference points.

References

- Dechenaux, Emmanuel, Dan Kovenock, and Roman M. Sheremeta**, "A Survey of Experimental Research on Contests, All-Pay Auctions and Tournaments," *Working Paper*, 2012.
- Ellingsen, Tore**, "Strategic Buyers and the Social Cost of Monopoly," *American Economic Review*, June 1991, 81 (3), 648–57.
- Fonseca, M.A.**, "An Experimental Investigation of Asymmetric Contests," *International Journal of Industrial Organization*, 2009, 27, 582–591.
- Gill, David and Rebecca Stone**, "Fairness and desert in tournaments," *Games and Economic Behavior*, 2010, 69 (2), 346–364.
- Grund, Christian and Dirk Sliwka**, "Envy and Compassion in Tournaments," *Journal of Economics and Management Strategy*, 2005, 14 (1), 187–207.
- Konow, James**, "Which Is the Fairest One of All? A Positive Analysis of Justice Theories," *Journal of Economic Literature*, December 2003, 41 (4), 1188–1239.
- Lim, N**, "Social Loss Aversion and Optimal Contest Design," *Journal of Marketing Research*, 2010, 47, 777–787.
- Nitzan, Shmuel**, "Modelling Rent-Seeking Contests," *European Journal of Political Economy*, 1994, 10, 41–60.
- Rabin, Matthew**, "Incorporating Fairness Into Game Theory and Economics," *The American Economic Review*, 1993, 83, 1281–1302.
- Tullock, Gordon**, "Efficient Rent Seeking," in J. Buchanan, R. Tollison, and G. Tullock, eds., *Towards a Theory of the Rent-Seeking Society*, A & M University Press, 1980, pp. 97–112.