Chapter 16

THE PARADOX OF POWER

YVONNE DURHAM

University of Arkansas

JACK HIRSHLEIFER

University of California at Los Angeles

VERNON L. SMITH

Economic Science Laboratory
University of Arizona

Individuals and organizations – if rational and self-interested – will equalize the marginal returns of two ways of generating income: (1) production and exchange, versus (2) 'appropriative' efforts for capturing resources previously controlled by other parties (or to defend against appropriation by others). For example, management and labor jointly generate the output of the firm, but also are in contention over the distribution of the firm's net revenues. This balancing problem has been examined in a number of theoretical studies, among them Haavelmo (1954), Skogh and Stuart (1982), Hirshleifer (1989, 1991), Skaperdas (1996), and Grossman and Kim (1995).

Following the model proposed by Hirshleifer (1991) in our experimental investigation, decision-makers simultaneously interact through *joint production* (requiring cooperation) and a *distributive struggle* (conflict).

It might be expected that in appropriative struggles between stronger and weaker contenders, the strong would grow ever stronger and the weak always weaker. The "paradox of power" (POP) is the observation that in actual contests, poorer or smaller combatants often end up improving their position relative to richer opponents. A notable instance is the political struggle over income redistribution. Although citizens in the upper half of the income spectrum surely have more political strength than those in the lower half, modern governments have mainly been attempting to transfer income from the former to the latter group (Browning, and Browning, 1994, pp. 259–261).

The theoretical explanation is that the marginal payoff of appropriative effort relative to productive effort is typically greater at low levels of income. Thus, while the rich always have the capability of exploiting the poor, it often does not pay them to do so.

Nevertheless, in some contexts, initially richer contestants do exploit weaker rivals. So the question is, when does POP occur? The crucial factor is a parameter (m) reflecting the *decisiveness* of conflictual effort. When decisiveness is low, the rich are content to concentrate upon producing a larger social pie of income even though the poor will be gaining an improved share thereof. But when conflictual preponderance makes a suf-

Handbook of Experimental Economics Results, Volume 1

DOI: 10.1016/S1574-0722(07)00016-9

ficiently weighty difference for achieved income, the rich cannot afford to let the poor win the contest over distributive shares.

1. The Model

Each of two contenders ($i = 1, 2$) must divide his/her exogenously given resource endowment R_i between productive effort E_i and appropriative ('fighting') effort F_i:

$$E_1 + F_1 = R_1, \qquad E_2 + F_2 = R_2. \tag{1}$$

The efforts (E_i) are inputs to a joint production function. For simplicity here we have assumed the simple additive form:

$$I = E_1 + E_2. \tag{2}$$

Thus, the parties can cooperate by combining their productive efforts so as to generate a common pool of income, I, available to the two of them jointly. But the respective shares p_1 and p_2 (where $p_1 + p_2 = 1$) are determined in a conflictual process: the contest success function (CSF) takes the fighting efforts F_i as inputs, yielding the distributive shares:

$$p_1 = F_1^m / \left(F_1^m + F_2^m\right), \qquad p_2 = F_2^m / \left(F_1^m + F_2^m\right). \tag{3}$$

Here, m is a 'decisiveness parameter' controlling the mapping of the input ratio F_1/F_2 into the success ratio p_1/p_2. For $m \leqslant 1$ the CSF is characterized by diminishing marginal returns as F_1 increases with given F_2, or vice versa. However, for $m > 1$ there will be an initial range of increasing returns before diminishing marginal returns set in.

The resulting incomes are:

$$I_1 = p_1 I, \qquad I_2 = p_2 I. \tag{4}$$

For each level of fighting effort by contender 2, there is a corresponding optimal effort for contender 1 (and vice versa). Thus, 1's optimization problem is to choose $F_1 \geqslant 0$ so as to solve:

$$\text{Max } I_1 = p_1(F_1|F_2) \times I(E_1|E_2) \quad \text{subject to} \quad E_1 + F_1 = R_1.$$

The solution is similar for side 2.

For an interior equilibrium, the Nash–Cournot reaction functions are defined by

$$\frac{F_1}{F_2^m} = \frac{m(R_1 + R_2 - F_1 - F_2)}{F_1^m + F_2^m}, \qquad \frac{F_2}{F_1^m} = \frac{m(R_1 + R_2 - F_1 - F_2)}{F_1^m + F_2^m}. \tag{5}$$

The right-hand sides being identical, $F_1 = F_2$ is always a solution of these equations. That is, the reaction curves intersect along the 45° line between the F_1 and F_2 axes. In fact, this is the sole intersection in the positive quadrant.

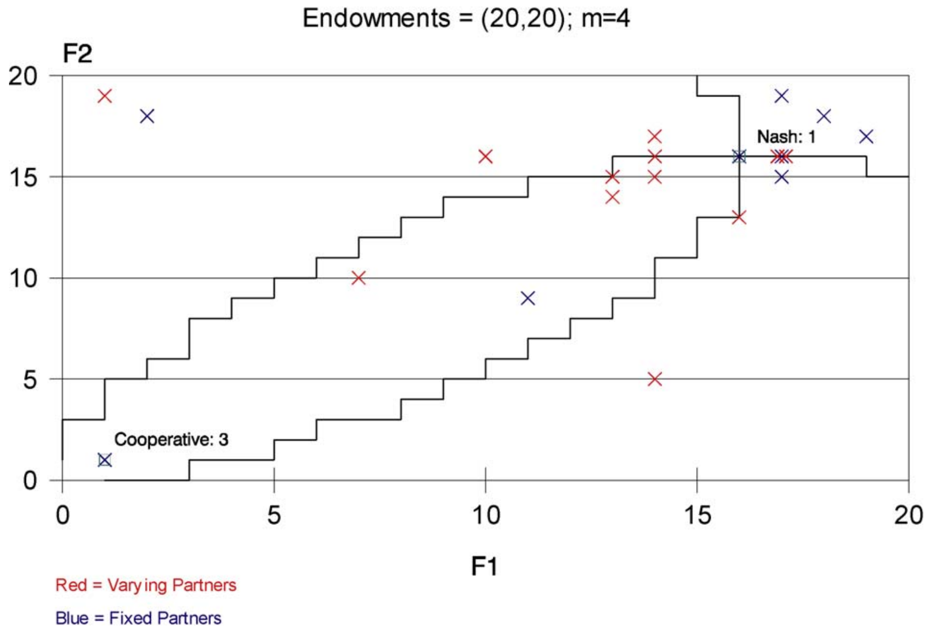

Figure 1. The step function in Figures 1–3 are the optimal reaction functions for each of the two protagonists for various resource endowments and value for the decisiveness parameter, m. Thus in this figure, the lower step function defines the optimal fighting effect, F_1, for player 1 for each fighting effort F_2 chosen by player 2 and vice versa for the upper step function. The Nash equilibrium is at the intersection (16, 16), while the joint maximizing cooperative outcome is at (1, 1). The $\times$s in this figure plot the observations on the 16th round of play in experiments 13 and 14 for varying partners, shown in red, and experiments 15 and 16 for fixed partners, shown in blue. Note the tendency for the observations to cluster nearer to the Nash equilibrium than to the cooperative; this tendency is stronger for varying, than for fixed partners; also, the error relative to Nash tends to the cooperative side.

If however the boundary constraint is binding for the poorer side (which we always take to be contender 2), the second equation would be replaced by

$$F_2 = R_2. \tag{5a}$$

In that case, at equilibrium, F_1 and F_2 are in general unequal, but the intersection of the reaction functions still determines the Nash–Cournot equilibrium values of the fighting efforts. Figures 1–3 plot the integer valued reaction functions in (5) and (5a) for the different values of (R_1, R_2) and m used in the experiments we report here. In each figure the lower reaction function is for subject 1, the upper one for subject 2. Thus, for each level of fighting effort by 2, we read off the conditionally optimal fighting effort for 1 using the lower step function. Each figure corresponds to a different parameterization of the model.

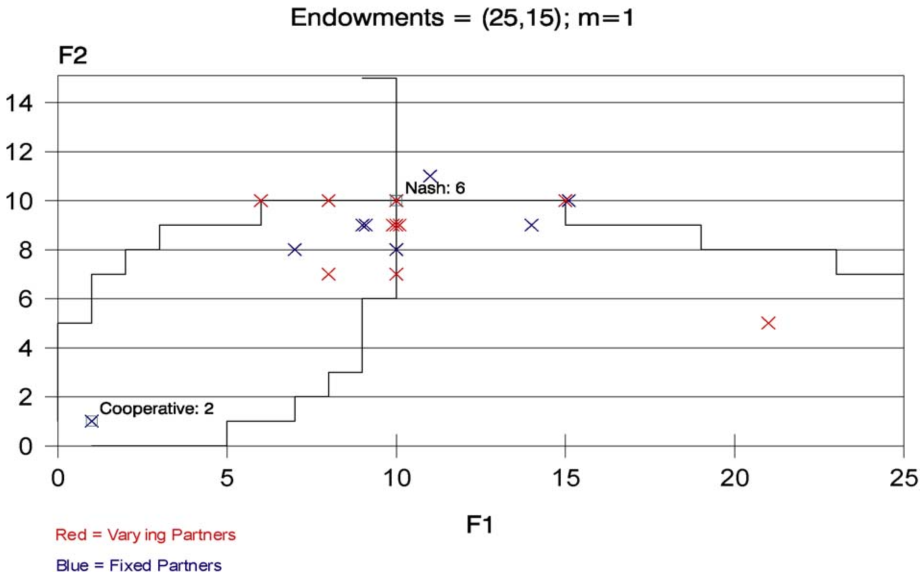

Figure 2. The $\times$s plot the final observations in experiments 5 and 6 for varying partners, shown in red, and experiments 7 and 8 for fixed partners, shown in blue. The tendency toward Nash is strong, stronger for varying than fixed partners, with many error deviations from Nash tending to occur on the cooperative side.

The experiments were intended in part to challenge a number of specific predictions derived from the model.

(i) *Fighting intensities*: If the decisiveness parameter m is increased, it pays both sides to 'fight harder.' As the F_i increase the ultimate achieved incomes (I_i) must fall. (Compare Figure 1 with Figures 2 and 3.)

(ii) *Conflict as an equalizing process* (*paradox of power*)*, strong vs weak form*. For sufficiently low values of the decisiveness parameter m, disparities in achieved income will be smaller than the initial disparities in resource endowments. Letting contender 1 be the initially better endowed side:

$$R_1/R_2 > I_1/I_2 \geqslant 1. \tag{6}$$

When the *equality* on the right holds we have the 'strong form' of the POP. It can be shown that there will be interior solutions up to some critical value ρ of the resource ratio:

$$\rho = (2+m)/m. \tag{7}$$

When $m = 1$ for the decisiveness parameter, the prediction is that the strong form of the POP will hold for low resource ratios, specifically for $R_1/R_2 \leqslant 3$. For resource ratios larger than $\rho = 3$, only the weak form, i.e., the strict *inequality* on the right of Equation (6), is predicted.

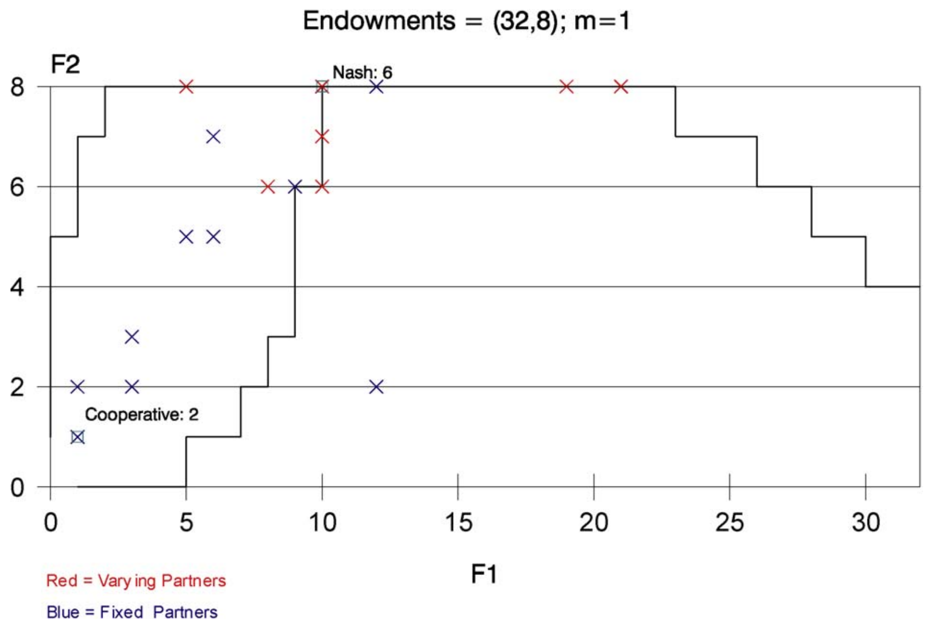

Figure 3. The data are from the last round of experiments 9–12. In comparison with Figures 1 and 2, the observations are not as strongly supportive of Nash, with many errors tending to the cooperative side of Nash. This is explained by the boundary Nash equilibrium for subjects 2. If they err relative to Nash it must be toward cooperation. Also, subjects 2 are limited by this boundary in responding with punishing choices. These elements encourage cooperation attempts although it is rarely achieved.

(iii) *Conflict as an inequality-aggravating process*. The model also indicates that for sufficiently high values of the decisiveness coefficient m and the resource ratio R_1/R_2, the POP will *not* apply. The rich would get richer and the poor poorer. Specifically, for our experiments using the high decisiveness coefficient $m = 4$, the critical value τ of the resource ratio for this condition is approximately 2.18. (The value of τ was obtained by finding the resource ratio where the condition $I_1/I_2 = R_1/R_2$ was met for $m = 4$.) Also, from (7), when $m = 4$ the critical ρ separating the weak from the strong forms of the POP equals 1.5. Thus in our experiments using the low resource ratio $25/15 = 1.67$ we expect the weak form of the POP to hold, since 1.67 lies between ρ and τ. However, for the experiments with $R_1/R_2 = 32/8 = 4 > \tau = 2.18$, the prediction is that the initially better endowed party will improve its relative position compared to the less well endowed side:

$$I_1/I_2 = (F_1/F_2)^m > R_1/R_2. \tag{8}$$

2. Implementing the Model

Certain game-theoretic and implementational concerns are also addressed in our experimental test of the above model. In the strict game-theoretic sense, the non-cooperative equilibrium is about strangers who meet once, interact strategically in their self-interest, and will never meet again. Such conditions control for repeated game effects, since the protagonists have no history or future. Yet in many contexts individuals interact in repeated games, where they can signal, punish, and build reputations. In the particularly simple version where the one-shot game is iterated with the same payoffs each round, we have a supergame. The study of such games has been motivated by the intuition or "folk theorem" that repetition makes cooperation possible (Mertens, 1984).

Consequently, in addition to testing the substantive predictions associated with the paradox of power, we address some of these issues that have arisen in the experimental and game-theoretic traditions. Specifically, we will be comparing the results of experiments in which the partners are randomly varied in each round with experiments in which the partners are fixed throughout the supergame. As suggested by previous experimental studies and the "folk theorem," we anticipate that the condition of fixed partners will favor somewhat more cooperative behavior.

3. Experimental Procedures and Design

We conducted 24 experiments using a total of 278 subjects. No subject participated in more than one experiment. There were 6 bargaining pairs in each experiment, except for a few cases with only 4 or 5 pairs. Each experiment involved repeated play, the payoffs being constant in each round. Within each round, each subject pair chose simultaneously a (row, column) in a payoff matrix. Subjects were not informed how many rounds would take place; in fact, in each experiment there were 16 or 17 rounds before termination. Subjects were recruited for two-hour sessions but the experiments took much less time, making credible the condition of an unknown horizon.

In every round each subject allocated his/her initial endowment of tokens between an "Investment Account" (IA) and a "Rationing Account" (RA). (We deliberately avoided using any terminology suggestive of "fighting".) Tokens contributed to the IA corresponded to productive effort, E_i, in the theoretical model: the paired IA contributions generated an aggregate pool of income (in the form of 'experimental pesos') in accordance with Equation (2) above. Funds put into the RA corresponded to fighting effort, F_i, and determined the respective distributive shares p_1 and p_2 in accordance with Equations (3). For simplicity, only integer choices were permitted. (More precisely, each subject could allocate, within his/her resource constraint, amounts in integral hundreds of tokens to invest in the IA, the remainder, of course, going into the RA.) The totals of pesos ultimately achieved were converted into actual dollars at the end of the experiment, so subjects had a substantial motivation to make self-interested choices. (The payoffs ranged from \$.25 to \$75.25, not including the \$5 show-up fee. The average payoff was \$17.66.)

Table 1

There are two classes of experimental treatment variables: (1) model parameters which include resource endowments (R_1, R_2), and decisiveness, m; (2) pairing protocol, either with "variable partners" (subjects randomly rematched on each round) or "fixed partners" (subjects randomly matched once who then play repeatedly). Two experiments were run under each parameterization with variable, and two with fixed, partners. Most sessions were run in groups of 6 pairs (12 subjects), but in a few experiments not all subjects appeared on time, and were run with 4 or 5 pairs. Each experiment was run for either 16 or 17 repeat rounds

Model parameters		Number of experiments (subjects)		
Endowments (R_1, R_2)	Decisiveness m	Variable pairing	Fixed pairing	Totals
20, 20	1	2(24)	2(24)	4(48)
25, 15	1	2(24)	2(24)	4(48)
32, 8	1	2(22)	2(24)	4(46)
20, 20	4	2(22)	2(22)	4(44)
25, 15	4	2(20)	2(24)	4(44)
32, 8	4	2(24)	2(24)	4(48)
	Totals	12(136)	12(142)	24(278)

To challenge the implications of the model, we manipulated the resource endowments R_1 and R_2 and also the decisiveness coefficient m. Four experiments were run with each of the three endowment vectors $(R_1, R_2) = (20, 20)$, $(25, 15)$, and $(32, 8)$ – the first series using a low value $m = 1$ of the decisiveness parameter, and the next using a high value $m = 4$. Thus there were 24 experiments in all.

Also, in view of the well-established results (e.g., McCabe, Rassenti, and Smith, 1996) that cooperation is promoted by repeated play with the same partners, each group of four experiments was further subdivided into alternative matching protocols. In the first ('varying partners') protocol, partners were randomly changed each round. Under the second ('fixed partners') protocol, subjects were randomly paired at the beginning of the experiment but played repeatedly with the same partner throughout.

Overall there were eight experiments under each of the three endowment conditions. Four of the eight involved varying partners, and four fixed partners. There was an analogous subdivision between experiments conducted using $m = 1$ and using $m = 4$. The design is summarized in Table 1. The pairing treatment and model parameters for each of the 24 experiments are listed in columns 1–7 of Table 2.

4. Results

4.1. Nash vs Cooperative Comparisons

Column 8 in Table 2 lists the results of each experiment as measured by the average fighting efforts (F_1, F_2) of the 4 to 6 pairs participating in each experiment. Columns 9–

Table 2

Each experiment listed in column 1, consists of 4 to 6 pairs of subjects (Table 1). Columns 2–5 list the corresponding model parameters, and columns 6 and 7 the Nash solution. In columns 8–10, are listed the average of the 4 to 6 observations on (F_1, F_2) occurring in round 16, and the ratios F_1/F_2, and I_1/I_2. The last two columns list the values of S_1 and S_2, the index of average slippage toward cooperation for the type 1 and 2 players respectively, where $S_i = (N_i - F_i)/(N_i - C)$. The S_i are mostly positive indicating that errors in hitting Nash tend to be biased in the direction of the cooperative solution at (1, 1). When the parameter m increases from 1 to 4 (column 2), the POP model predicts an increase in (F_1, F_2); in the 48 comparisons between the bottom and top halves of the table, 45 observations satisfy this prediction. When $m = 1$, POP predicts $I_1/I_2 < R_1/R_2$ for the 8 experiments with $R_1 > R_2$. As is seen in rows 5–12, all 8 observations conform to this prediction. See text for additional predictions only partially conforming to the data

Exp. #	Treatment parameters				Nash solution		Avg. results (16th obs.)			Avg. slippage	
	m	Pairing	R_1, R_2	R_1/R_2	N_1, N_2	N_1/N_2	F_1, F_2	F_1/F_2	I_1/I_2	S_1	S_2
1	1	V	20, 20	1	10, 10	1	7.83, 6.83	1.15	1.15	0.24	0.35
2	1	V	20, 20	1	10, 10	1	8,9	0.89	0.89	0.22	0.11
3	1	F	20, 20	1	10, 10	1	8.67, 6.67	1.30	1.30	0.15	0.37
4	1	F	20, 20	1	10, 10	1	4, 5	0.8	0.8	0.67	0.56
5	1	V	25, 15	1.67	10, 10	1	10.83, 8.5	1.27	1.27	−0.09	0.17
6	1	V	25, 15	1.67	10, 10	1	9, 9.17	0.98	0.98	0.11	0.09
7	1	F	25, 15	1.67	10, 10	1	10.17, 9	1.13	1.13	−0.02	0.11
8	1	F	25, 15	1.67	10, 10	1	7.67, 6.83	1.12	1.12	0.26	0.41
9	1	V	32, 8	4	10, 8	1.25	11.83, 7.67	1.54	1.54	−0.20	0.05
10	1	V	32, 8	4	10, 8	1.25	10.33, 7.5	1.38	1.38	−0.04	0.07
11	1	F	32, 8	4	10, 8	1.25	5.17, 3.17	1.63	1.63	0.54	0.69
12	1	F	32, 8	4	10, 8	1.25	5.4, 4.6	1.17	1.17	0.51	0.46
13	4	V	20, 20	1	16, 16	1	10.33, 12.83	0.81	0.42	0.38	0.21
14	4	V	20, 20	1	16, 16	1	14.67, 15.33	0.96	0.84	0.09	0.04
15	4	F	20, 20	1	16, 16	1	11.67, 13.83	0.84	0.51	0.29	0.15
16	4	F	20, 20	1	16, 16	1	10, 9.2	1.09	1.40	0.40	0.45
17	4	V	25, 15	1.67	16, 15	1.07	15.5, 11.83	1.31	2.95	0.33	0.23
18	4	V	25, 15	1.67	16, 15	1.07	16.5, 12.5	1.32	3.04	−0.03	0.18
19	4	F	25, 15	1.67	16, 15	1.07	16.17, 13.83	1.17	1.87	−0.01	0.08
20	4	F	25, 15	1.67	16, 15	1.07	13.5, 12.33	1.10	1.44	0.17	0.19
21	4	V	32, 8	4	12, 8	1.5	11.67, 7.33	1.59	6.42	0.03	0.10
22	4	V	32, 8	4	12, 8	1.5	11.67, 7	1.67	7.72	0.03	0.40
23	4	F	32, 8	4	12, 8	1.5	10.5, 7.5	1.4	3.84	0.14	0.07
24	4	F	32, 8	4	12, 8	1.5	11, 4.67	2.36	30.76	0.09	0.48

12 are calculations derived from (F_1, F_2). Thus, F_1/F_2 and I_1/I_2 are ratios of the average fighting efforts and the attained incomes of the protagonists in each experiment. Since the cooperative solution is $(C_1, C_2) = (1, 1)$, column 8 allows a comparison between the observed average outcome (F_1, F_2) and the Nash and cooperative solutions. It is immediately evident that most of the observations are closer to the Nash than the cooperative solution. However, the error relative to Nash tends toward the cooperative

side. This "slippage" (S_i), indicated in the last two columns, is defined as:

$$S_i = (N_i - F_i)/(N_i - C).$$

A positive S_i indicates that error with respect to Nash is tending toward the cooperative solution while a negative S_i indicates that the error is tending away from the cooperative solution.

4.2. Predictions of the Model

Various predictions of the model can be evaluated from Table 2.

PREDICTION 1. Higher values of the decisiveness parameter m will lead to larger fighting efforts on both sides.

The upper half of Table 2 shows the results for $m = 1$, and the lower half for $m = 4$. There are 48 comparisons, of which a remarkable 45 are in the direction predicted.

PREDICTION 2A. At the low value $m = 1$ of the decisiveness parameter, the initially poorer side will always end up improving its position.

At $m = 1$ the attained income ratio I_1/I_2 should exceed the resource ratio R_1/R_2. This is the "paradox of power." The requirement of unequal initial endowments limits the relevant data to rows 5 through 12 of Table 2. Here all 8 of the 8 comparisons show the predicted relative improvement – that is, $I_1/I_2 < R_1/R_2$.

PREDICTION 2B. For $m = 1$ the poorer side should attain approximate equality of income (strong form of the POP) for initial resource ratios $R_1/R_2 < 3$, but only some relative improvement, $1 < I_1/I_2 < R_1/R_2$, for larger resource ratios (weak form of the POP).

Looking only at the unequal endowments cases, rows 5 through 12 of Table 2, the average of the tabulated results is $I_1/I_2 = 1.125$, on the high side of the predicted $I_1/I_2 = 1$. By way of comparison, for rows 9 through 12 $I_1/I_2 > 1$ is predicted, the average outcome is $I_1/I_2 = 1.43$. In a relative sense, the predicted comparison of the strong form versus weak form predictions is supported.

PREDICTION 2C. At the specific high value $m = 4$ of the decisiveness coefficient, the paradox of power should continue to hold (in its weak form) for $\rho < R_1/R_2 < \tau$, where $\rho = 1.5$ and $\tau = 2.18$. But for higher resource ratios the richer side should end up actually improving on its relative position. That is, in this range $I_1/I_2 = (F_1/F_2)^4$ should exceed R_1/R_2.

For the unequal-endowments rows 17 though 20 of Table 2, the resource ratio is $R_1/R_2 = 25/15 = 1.67$, lying between ρ and τ. So the paradox of power is predicted in these cases. However, for rows 21 through 24 the resource ratio is $R_1/R_2 = 32/8 = 4 > 2.18 = \tau$, so we expect the rich to become richer still.

Taking up the latter group first, 3 of the 4 cases support the prediction $I_1/I_2 = (F_1/F_2)^4 > 4$. In fact, the average of the observed results was a much higher $I_1/I_2 = 12.19$. That is, the model predicted that here the rich would get richer – and they did. Turning to the first group, however, all 4 cases violate the prediction! Quantitatively, the predicted Nash outcome $(N_1, N_2) = (16, 15)$ implies $I_1/I_2 = (16/15)^4 = 1.29 < 1.67$ while the average of the observed results was $I_1/I_2 = 2.32 > 1.67$. Here the poor should have improved their position, but did not. On the whole the rich behaved about as expected, whereas the poor fell short of the fighting effort $F_2 = 15$ predicted for them. But notice that the Nash prediction for the poorer side required them to devote 100% of their resources to fighting. Thus, any error whatsoever could only take the form of a deficiency of fighting effort. Such outcomes commonly occur in experiments with boundary equilibria (Smith and Walker, 1993).

4.3. *Charting the Observations*

The final period (16) observations are plotted for four experiments in each of three parameter classes: Figure 1, experiments 13–16; Figure 2, experiments 5–8; Figure 3, experiments 9–12. Each "×" symbol plots the values (F_1, F_2) chosen by the 12 bargaining pairs in round 16, along with the reaction step functions defined for integer outcomes. The computed Nash equilibrium at the intersection, and cooperative solution at (1, 1) are as shown. The red symbols are the final outcomes for varying partners, while the blue symbols are for fixed partners.

To summarize: (1) as predicted, the outcomes tend to support (are closer to) Nash. (The statistical tests, both classical and Bayesian, strongly support Nash against cooperation for both fixed and varying partners; see Durham, Hirshleifer, and Smith, 1998); (2) but the error deviation or "slippage" from Nash tend to be biased toward cooperation; (3) as predicted by the "folk theorem" the slippage toward cooperation is stronger in the case of fixed than varying partners, as some pairs attempt to coordinate on the cooperative outcome; (4) the predictions of the POP model are broadly supported.

References

Browning, Edgar F., Browning, Jacquelene M. (1994). "Public Finance and the Price System", fourth ed. Prentice Hall, Englewood Cliffs, NJ.

Durham, Yvonne, Hirshleifer, Jack, Smith, Vernon L. (1998). "Do the rich get richer and the poor poorer? Experimental tests of a model of power". American Economic Review 88, 970–983.

Grossman, Herschel, Kim, Minseong (1995). "Swords or plowshares: A theory of the security of claims to property". Journal of Political Economy 103, 1275–1288.

Haavelmo, Trygve (1954). "A Study in the Theory of Economic Evolution". North-Holland Publishing Co., Amsterdam.

Hirshleifer, Jack (1989). "Conflict and rent-seeking success functions: Ratio vs. difference models of relative success". Public Choice 63, 101–112.

Hirshleifer, Jack (1991). "The paradox of power". Economics and Politics 3, 177–200.

McCabe, Kevin, Rassenti, Stephen, Smith, Vernon (1996). "Game theory and reciprocity in some extensive form experimental games". Proceedings National Academy of Science 93, 13421–13428.

Mertens, J. (1984). "Repeated games". In: Eatwell, J., Milgate, M., Newman, P. (Eds.), The New Palgrave. MacMillan Press, London, pp. 151–152.

Skaperdas, Stergios (1996). "Contest success functions". Economic Theory 7, 283–290.

Skogh, Goran, Stuart, Charles (1982). "A contractarian theory of property rights and crime". Scandinavian Journal of Economics 84, 27–40.

Smith, Vernon L., Walker, James M. (1993). "Monetary rewards and decision cost in experimental economics". Economic Inquiry 31, 245–261.