Backward Induction in Finitely Repeated Prisoner's Dilemma: Experimental Evidence

Matthew Embrey *Maastricht*
Guillaume Fréchette *NYU*
Sevgi Yuksel *NYU*

Zurich 2013

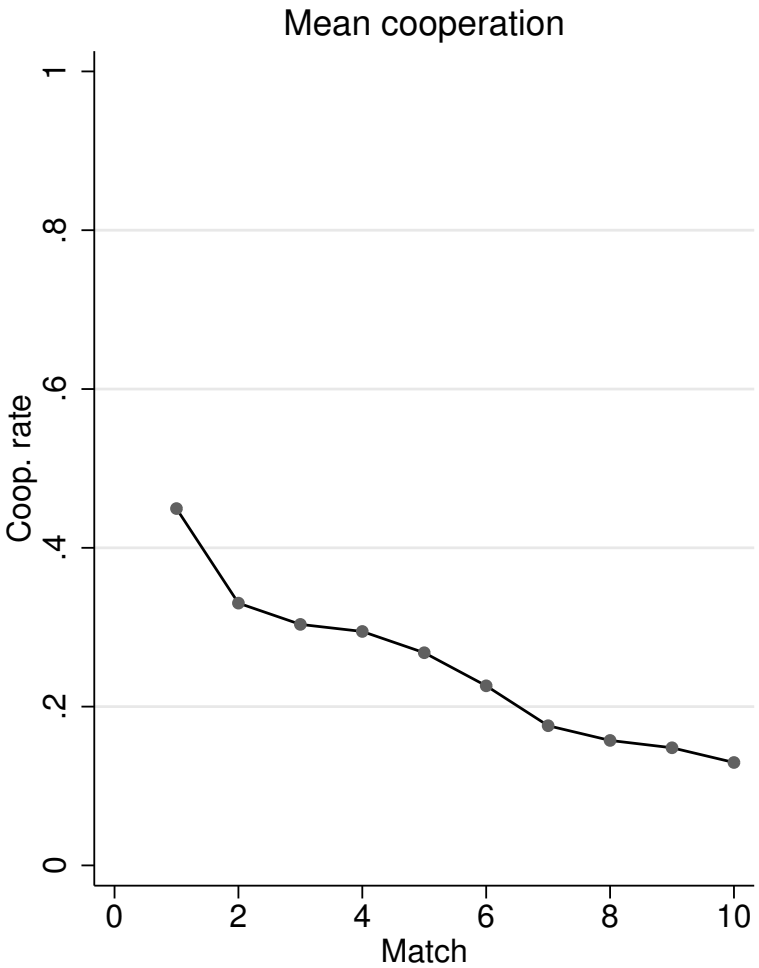
MOTIVATION

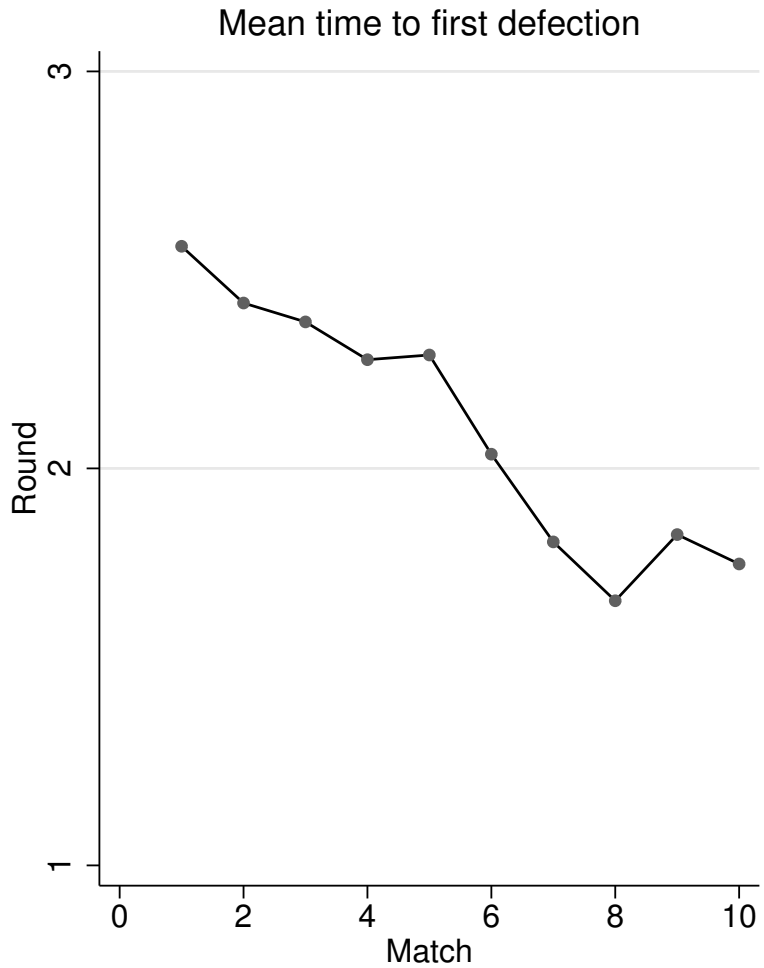
- The prisoner's dilemma is a canonical game in social sciences.
- We know little about how people behave in finitely repeated PDs.
- Results are inconclusive on whether or not subjects learn backward induction:
  - some studies find cooperation to decline with experience,
  - others find that first defection occurs later on in the interaction.
- Folk wisdom: As the horizon of the repeated game increases cooperation rates increase.
  - Backward induction is more difficult as there are more steps of reasoning required.

COOPERATION DECREASES WITH EXPERIENCE

Dal Bo (2005)

PD1, Horizon = 4

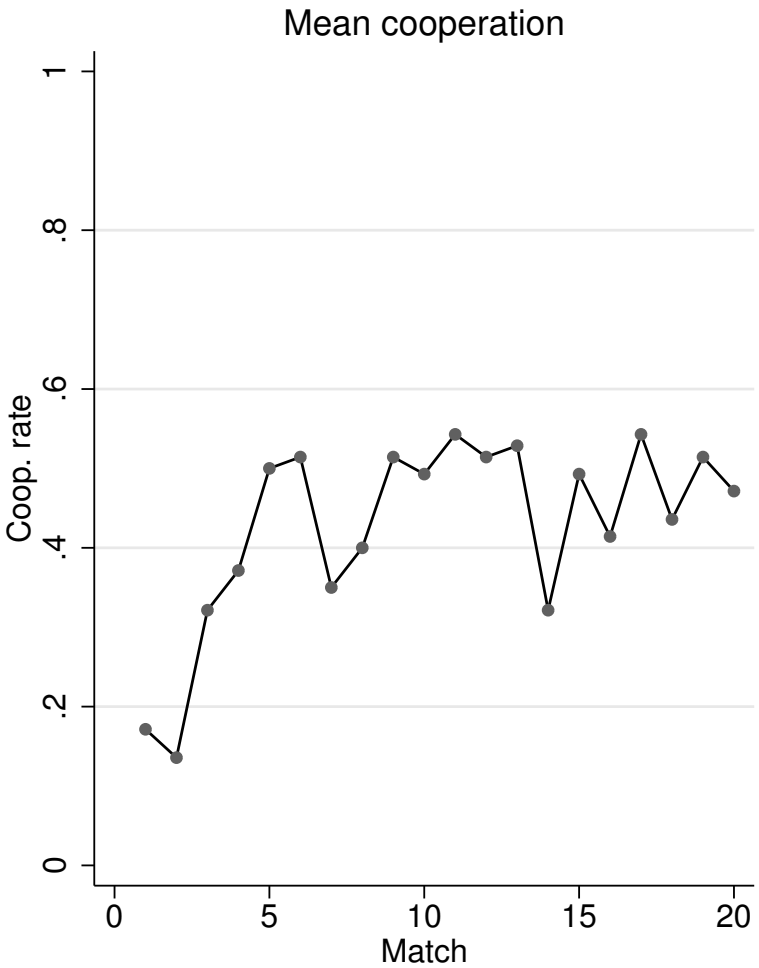


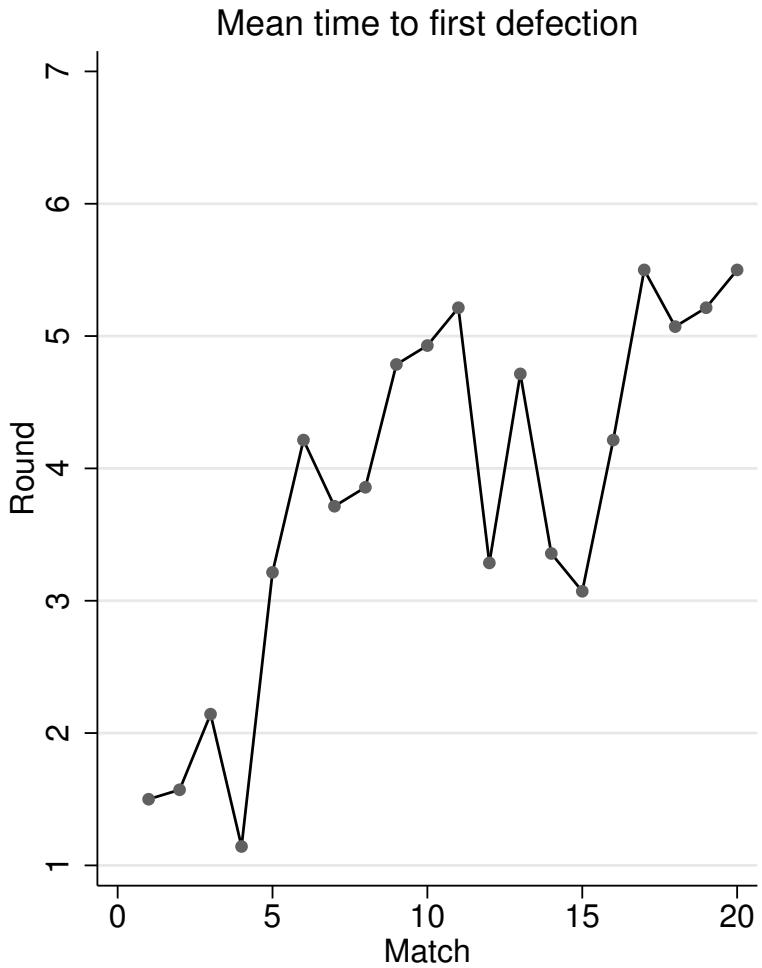


COOPERATION INCREASES WITH EXPERIENCE

Andreoni & Miller (1993)

Horizon = 10





PRIOR WORK

- Studies differ in the parameters of the game: number of matches, the horizon, stage game payoffs; but rarely within studies.
- Data is analyzed differently
  - Round 1 behavior
  - Last round behavior
  - Mean cooperation rate
  - Mean time to first defection

WHAT WE DO

- Meta-analysis of the prior experimental research.
- Conduct a new experiment to understand how the ability to perform backward induction varies with the environment.
  - Document how this varies with the parameters.

OUR CONTRIBUTIONS

Statistical:

- Measures used in the literature (mean cooperation rate, mean time to first defection) can be misleading.

OUR CONTRIBUTIONS

Statistical:

- Measures used in the literature (mean cooperation rate, mean time to first defection) can be misleading.

Design:

- Establish that having many matches is important.
- Conduct the first study varying payoffs and horizon in a between subject design.

OUR CONTRIBUTIONS

Statistical:

- Measures used in the literature (mean cooperation rate, mean time to first defection) can be misleading.

Design:

- Establish that having many matches is important.
- Conduct the first study varying payoffs and horizon in a between subject design.

Conceptual:

- Propose that the impact of the horizon confounds other factors.
- Horizon affects play:
  - not because it requires more steps of backward induction,
  - but because it increases the value of cooperation which can be captured by the basin of attraction.

META: PAPERS AND PARAMETERS

1. Selten and Stoecker (1986)
   - $H = 10$, 25 matches
2. Andreoni and Miller (1993)
   - $H = 10$, 20 matches, 1 session
3. Cooper, DeJong, Forsythe, and Ross (1996)
   - $H = 10$, 2 matches, 3 sessions
4. Hauk and Nagel (2001)
   - $H = 10$, 10 ($\times 6$) matches
5. Dal Bó (2005)
   - $H = \{2, 4\}$, {5,8,9,10} matches, 4 sessions, within design
6. Bereby-Meyer and Roth (2006)
   - $H = 10$, 20 matches, 4 sessions
7. Friedman and Oprea (2012)
   - $H = 8$, 8 matches, 3 sessions, within design

META: STAGE GAME PARAMETERS

Original

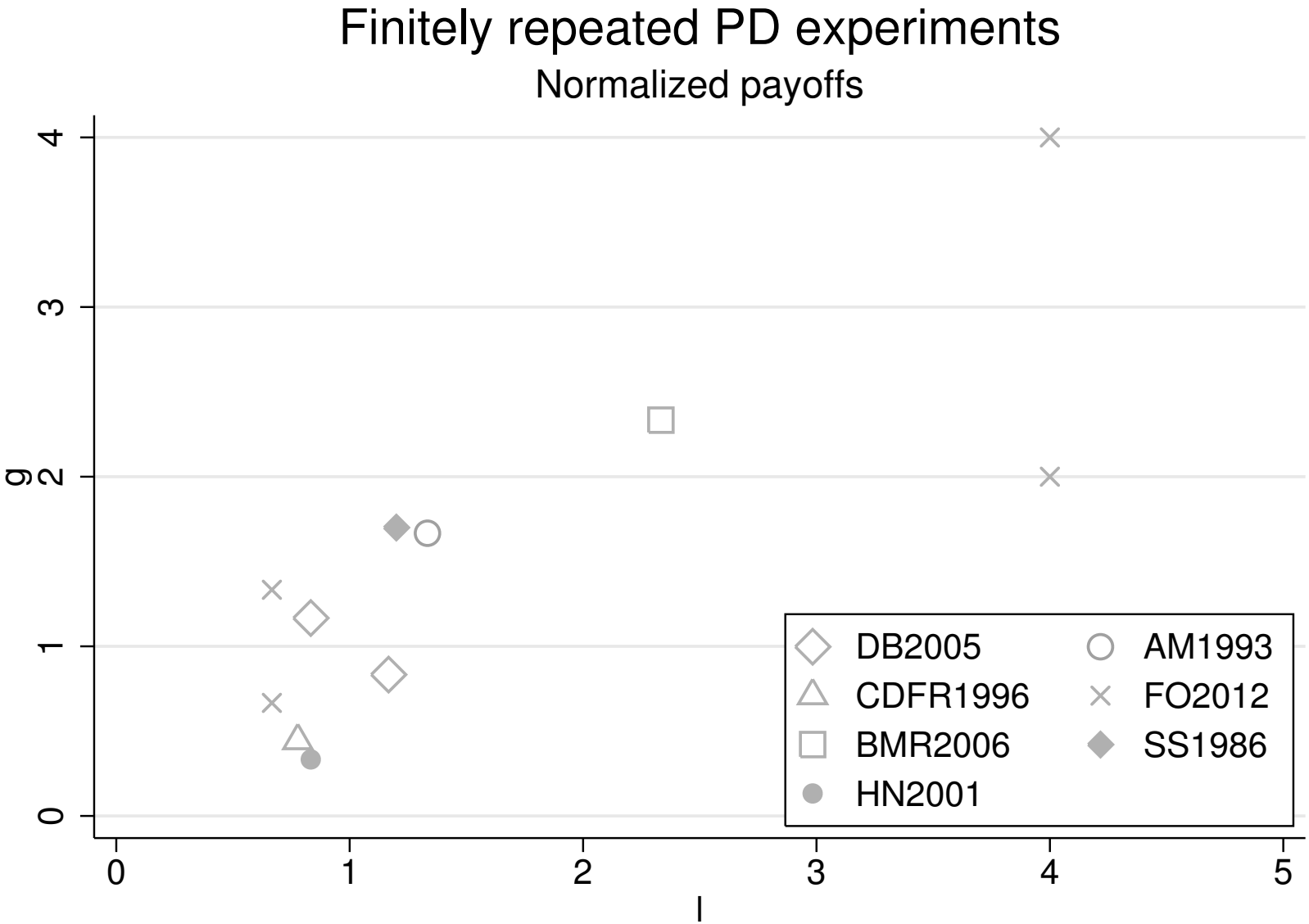
	C	D
C	R	S
D	T	P

Normalized

	C	D
C	$\frac{R-P}{R-P} = 1$	$\frac{S-P}{R-P} = -\ell$
D	$\frac{T-P}{R-P} = 1 + g$	$\frac{P-P}{R-P} = 0$

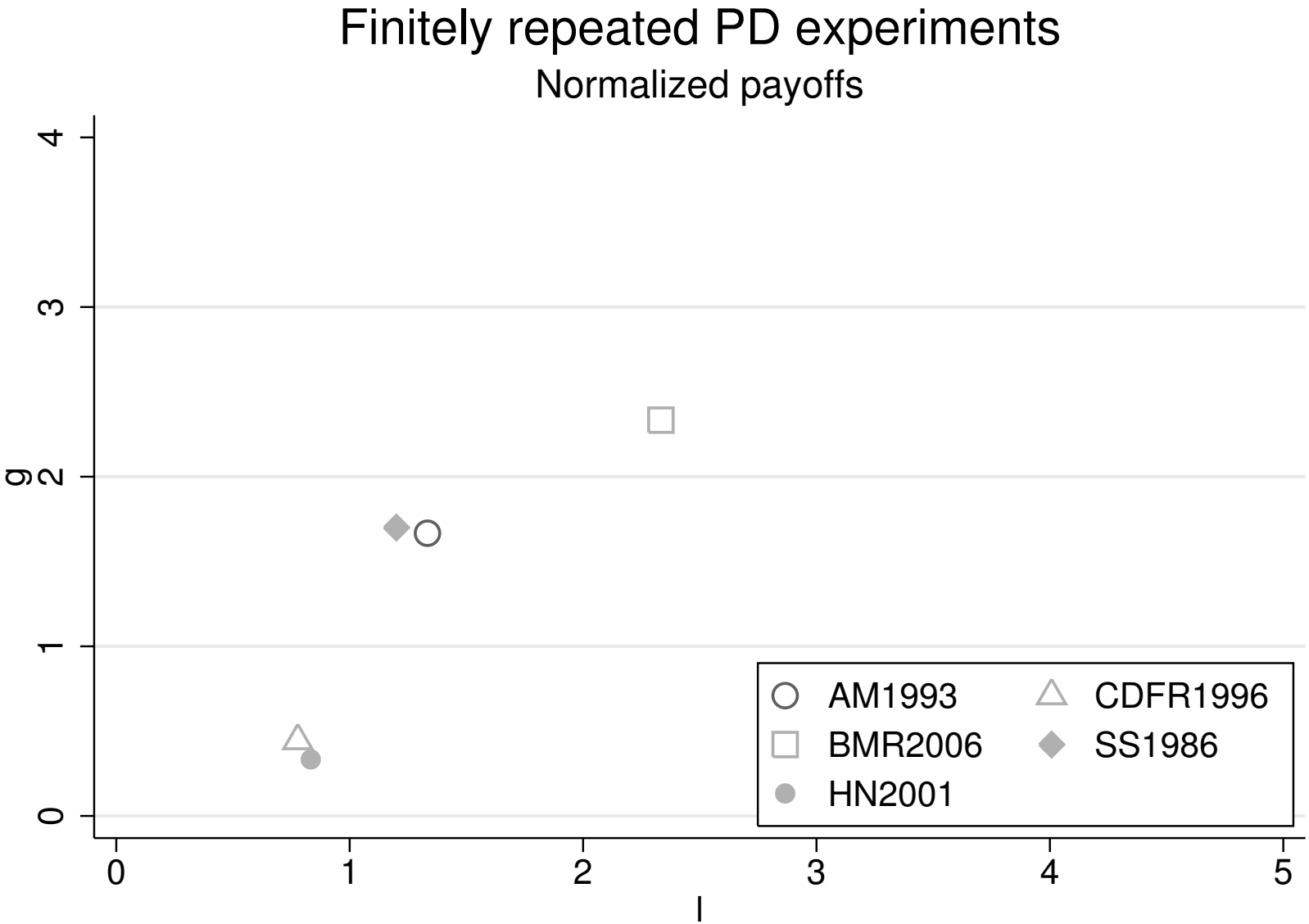
$T > R > P > S$

META: STAGE GAME PARAMETERS



META: STAGE GAME PARAMETERS

BETWEEN SUBJECTS DESIGNS ONLY



META: STANDARD ANALYSIS

- Payoffs matter: high g and high ℓ decrease cooperation.
- With experience: last round cooperation rate close to zero.
- Average and round 1 cooperation, round to first defection:
  - when H is small, they decrease;
  - when H is large, they increase.

META: STANDARD ANALYSIS

				Cooperation rate (%)						Mean round to first def.	
				Average		Round 1		Last Round			
Experiment	H	g	ℓ	1	L	1	L	1	L	1	L
DB2005	2	1.17	0.83	14	13	18	14	10	11	1.21	1.20
within sub.	2	0.83	1.17	25	9	32	13	17	5	1.42	1.14
	4	1.17	0.83	33	20	44	32	25	8	1.99	1.58
	4	0.83	1.17	31	22	37	34	20	12	1.76	1.61
FO212	8	4.00	4.00	33	33	43	67	23	3	2.27	3.53
within sub.	8	2.00	4.00	38	34	43	63	30	3	2.77	3.67
	8	1.33	0.67	40	48	43	73	37	3	2.83	4.43
	8	0.67	0.67	44	69	50	87	30	23	3.1	6.07
BMR2006	10	2.33	2.33	38	66	61	93	22	7	3.19	7.39
AM1993	10	1.67	1.33	17	47	36	86	14	0	1.50	5.50
CDFR1996	10	0.44	0.78	52	57	60	67	20	27	4.63	5.53

First defection is set to Horizon + 1 if there are no defection.

Table : Cooperation rates and mean round to first defection

META: STANDARD ANALYSIS

				Cooperation rate (%)						Mean round to first def.	
				Average		Round 1		Last Round			
Experiment	H	g	ℓ	1	L	1	L	1	L	1	L
DB2005	2	1.17	0.83	14	13	18	14	10	11	1.21	1.20
within sub.	2	0.83	1.17	25	9	32	13	17	5	1.42	1.14
	4	1.17	0.83	33	20	44	32	25	8	1.99	1.58
	4	0.83	1.17	31	22	37	34	20	12	1.76	1.61
FO212	8	4.00	4.00	33	33	43	67	23	3	2.27	3.53
within sub.	8	2.00	4.00	38	34	43	63	30	3	2.77	3.67
	8	1.33	0.67	40	48	43	73	37	3	2.83	4.43
	8	0.67	0.67	44	69	50	87	30	23	3.1	6.07
BMR2006	10	2.33	2.33	38	66	61	93	22	7	3.19	7.39
AM1993	10	1.67	1.33	17	47	36	86	14	0	1.50	5.50
CDFR1996	10	0.44	0.78	52	57	60	67	20	27	4.63	5.53

First defection is set to Horizon + 1 if there are no defection.

Table : Cooperation rates and mean round to first defection

META: STANDARD ANALYSIS

				Cooperation rate (%)						Mean round to first def.	
				Average		Round 1		Last Round			
Experiment	H	g	ℓ	1	L	1	L	1	L	1	L
DB2005	2	1.17	0.83	14	13	18	14	10	11	1.21	1.20
within sub.	2	0.83	1.17	25	9	32	13	17	5	1.42	1.14
	4	1.17	0.83	33	20	44	32	25	8	1.99	1.58
	4	0.83	1.17	31	22	37	34	20	12	1.76	1.61
FO212	8	4.00	4.00	33	33	43	67	23	3	2.27	3.53
within sub.	8	2.00	4.00	38	34	43	63	30	3	2.77	3.67
	8	1.33	0.67	40	48	43	73	37	3	2.83	4.43
	8	0.67	0.67	44	69	50	87	30	23	3.1	6.07
BMR2006	10	2.33	2.33	38	66	61	93	22	7	3.19	7.39
AM1993	10	1.67	1.33	17	47	36	86	14	0	1.50	5.50
CDFR1996	10	0.44	0.78	52	57	60	67	20	27	4.63	5.53

First defection is set to Horizon + 1 if there are no defection.

Table : Cooperation rates and mean round to first defection

META: STANDARD ANALYSIS

				Cooperation rate (%)						Mean round to first def.	
				Average		Round 1		Last Round			
Experiment	H	g	ℓ	1	L	1	L	1	L	1	L
DB2005	2	1.17	0.83	14	13	18	14	10	11	1.21	1.20
within sub.	2	0.83	1.17	25	9	32	13	17	5	1.42	1.14
	4	1.17	0.83	33	20	44	32	25	8	1.99	1.58
	4	0.83	1.17	31	22	37	34	20	12	1.76	1.61
FO212	8	4.00	4.00	33	33	43	67	23	3	2.27	3.53
within sub.	8	2.00	4.00	38	34	43	63	30	3	2.77	3.67
	8	1.33	0.67	40	48	43	73	37	3	2.83	4.43
	8	0.67	0.67	44	69	50	87	30	23	3.1	6.07
BMR2006	10	2.33	2.33	38	66	61	93	22	7	3.19	7.39
AM1993	10	1.67	1.33	17	47	36	86	14	0	1.50	5.50
CDFR1996	10	0.44	0.78	52	57	60	67	20	27	4.63	5.53

First defection is set to Horizon + 1 if there are no defection.

Table : Cooperation rates and mean round to first defection

VALUE OF COOPERATION VS. DEFECTION

BASINS OF ATTRACTION

Normalized PD

	C	D
C	1, 1	$-\ell$, $1 + g$
D	$1 + g$, $-\ell$	0, 0

VALUE OF COOPERATION VS. DEFECTION

BASINS OF ATTRACTION

Normalized PD

	C	D
C	1, 1	$-\ell$, $1+g$
D	$1+g$, $-\ell$	0, 0

Reduced Game

	Grim	AD
Grim	H, H	$-\ell$, $1+g$
AD	$1+g$, $-\ell$	0, 0

VALUE OF COOPERATION VS. DEFECTION

BASINS OF ATTRACTION

Normalized PD

	C	D
C	1, 1	$-\ell$, $1+g$
D	$1+g$, $-\ell$	0, 0

Reduced Game

	Grim	AD
Grim	H, H	$-\ell$, $1+g$
AD	$1+g$, $-\ell$	0, 0

$$sizeBAD = \frac{\ell}{(H-1)+\ell-g}$$

META: DETERMINANTS OF COOPERATION

Table : Probit regression (Marginal Effects): round 1.

	(1)		(2)	
basin of att.			-0.359***	(0.093)
g	-0.063***	(0.020)	-0.034*	(0.018)
l	-0.006	(0.023)	-0.020	(0.025)
horizon	0.034***	(0.009)	0.003	(0.014)
observations	5398		5398	

Clustered (session level) standard errors in parentheses.
***1%, **5%, *10% significance.
Note: Regressors not reported here include cooperation in round 1 of first match, opponent cooperation in round 1 of previous match, and number of matches played.

DESIGN

2×2 factorial design (between subjects)

Summary of treatments		
PD/Horizon	4	8
Easy	E4	E8
Difficult	D4	D8

- Horizon: 4 or 8 rounds
- Payoffs of the PD: Easy vs. Difficult
- In each session subjects played 20-30 finitely repeated PDs

Easy PD

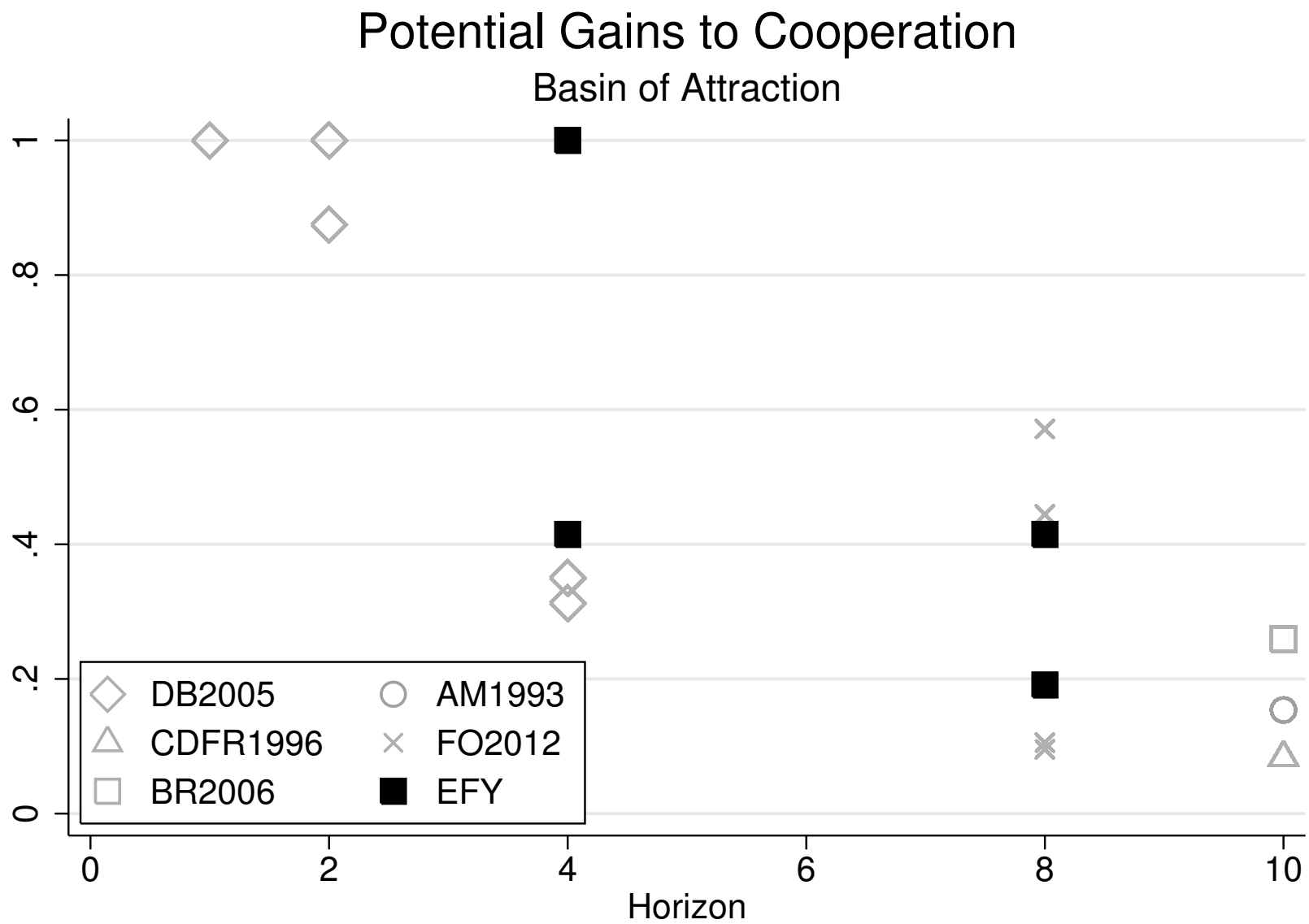
	C	D
C	51, 51	22, 63
D	63, 22	39, 39

Difficult PD

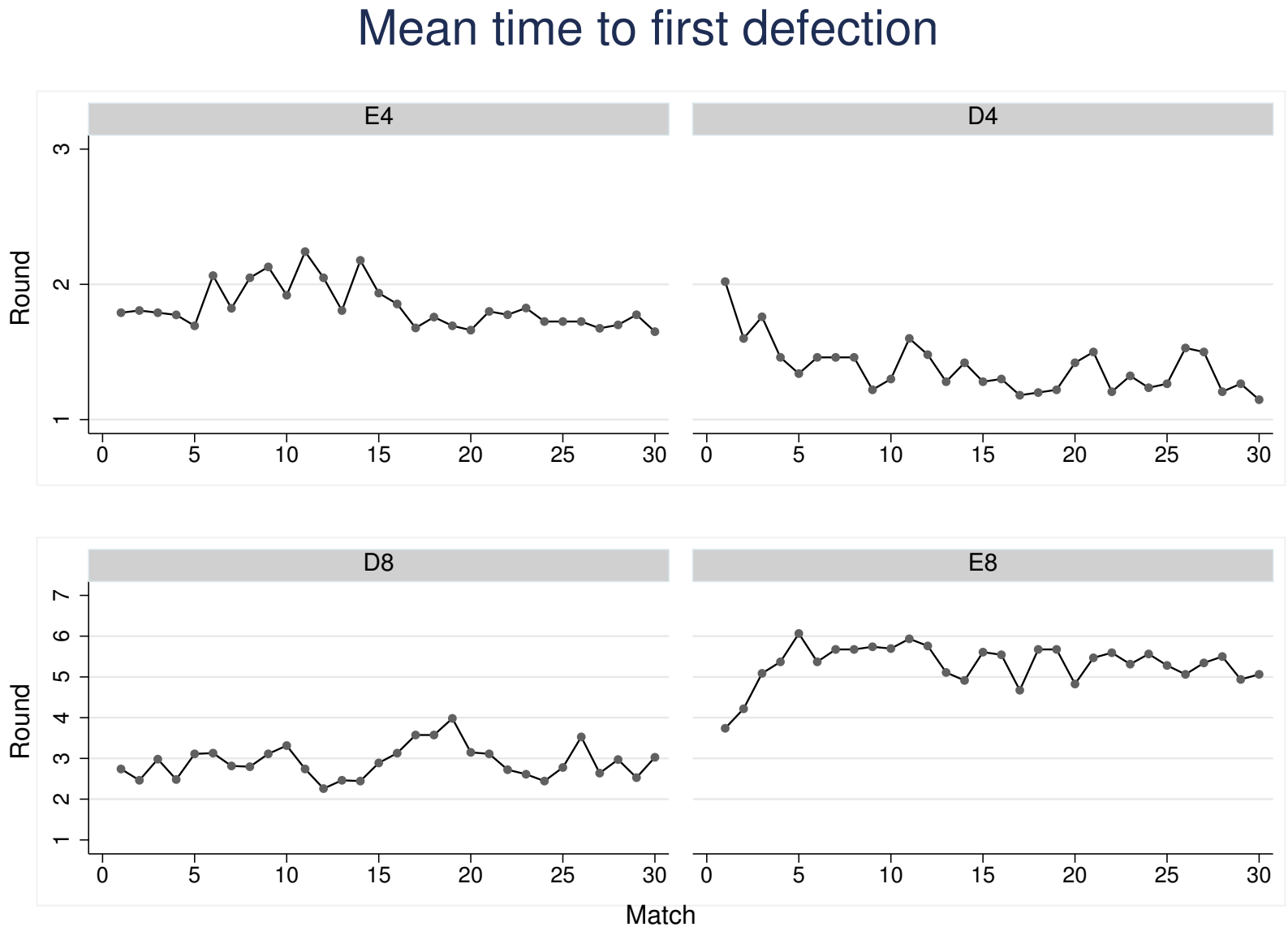
	C	D
C	51, 51	5, 81
D	81, 5	39, 39

DESIGN

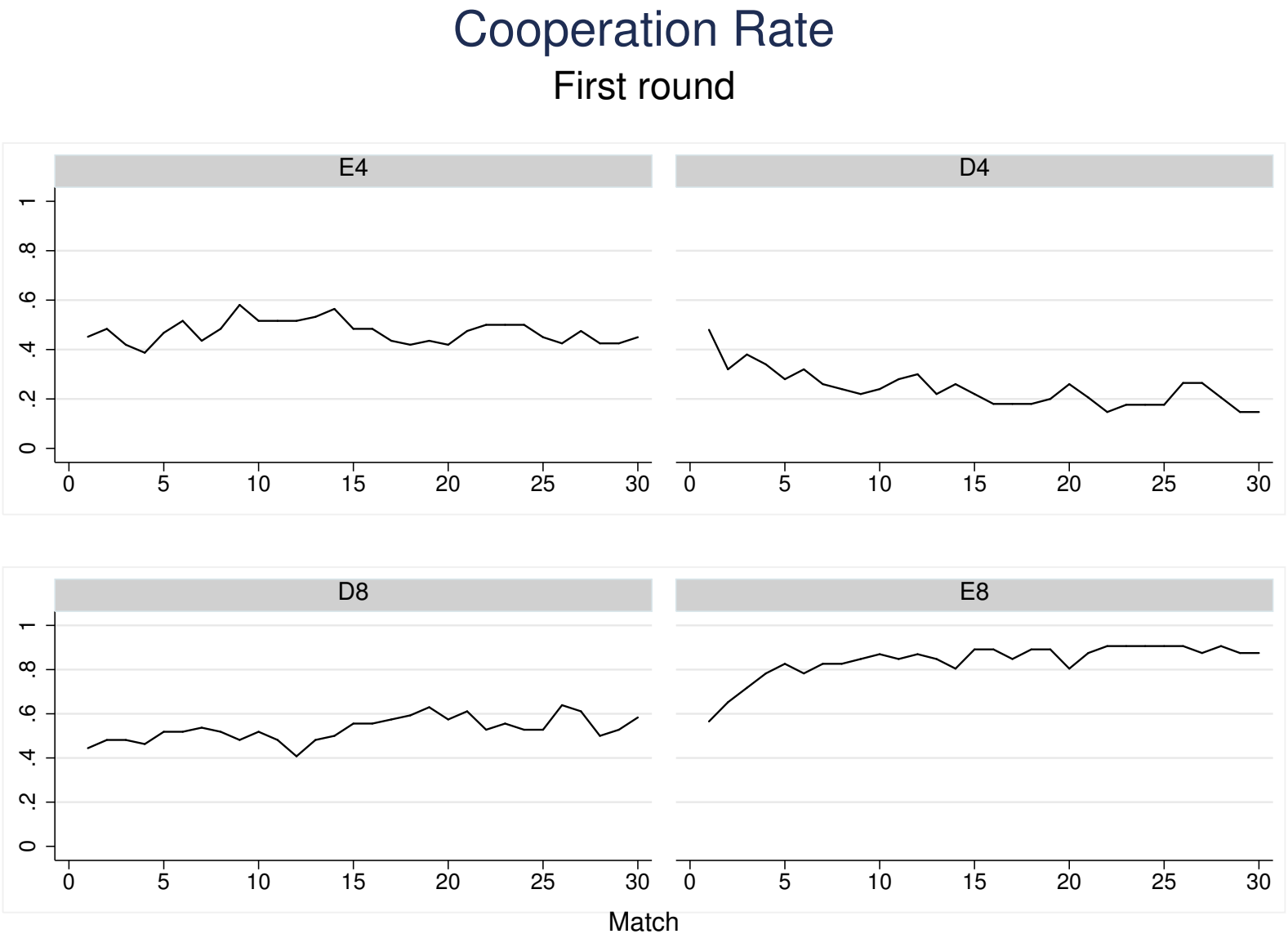
BASIN OF ATTRACTION OF GRIM



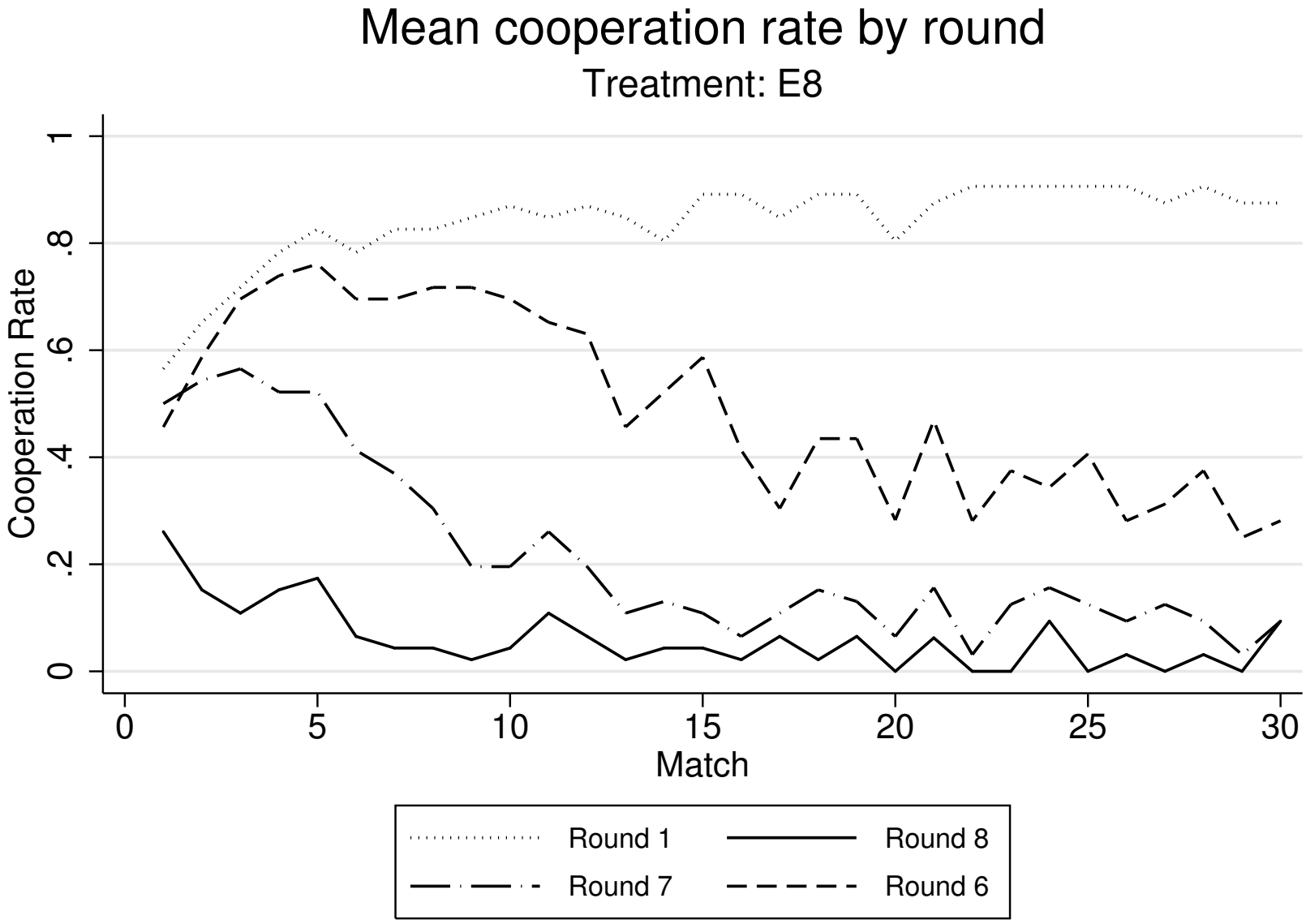
FIRST DEFECTION



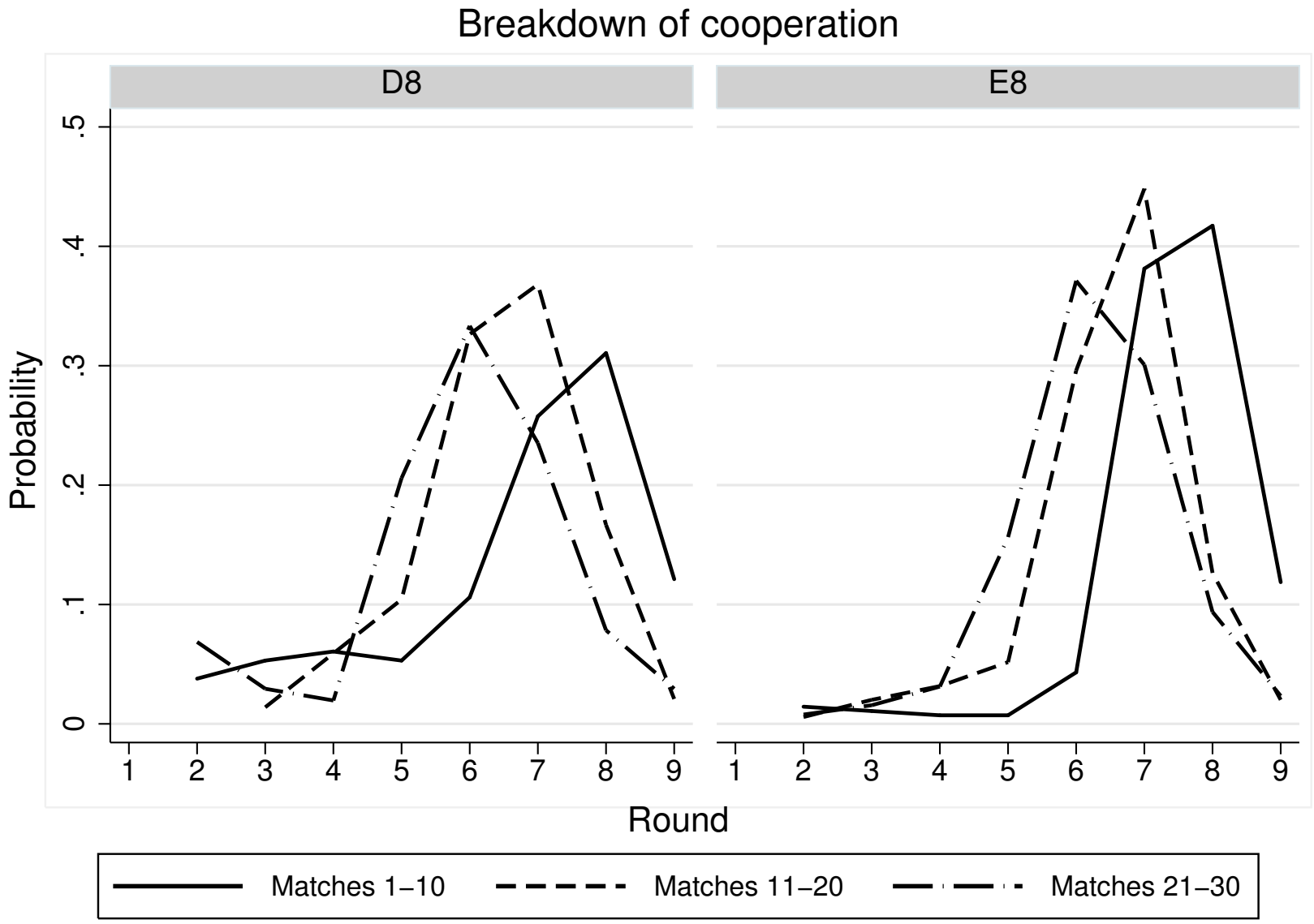
COOPERATION BY ROUND



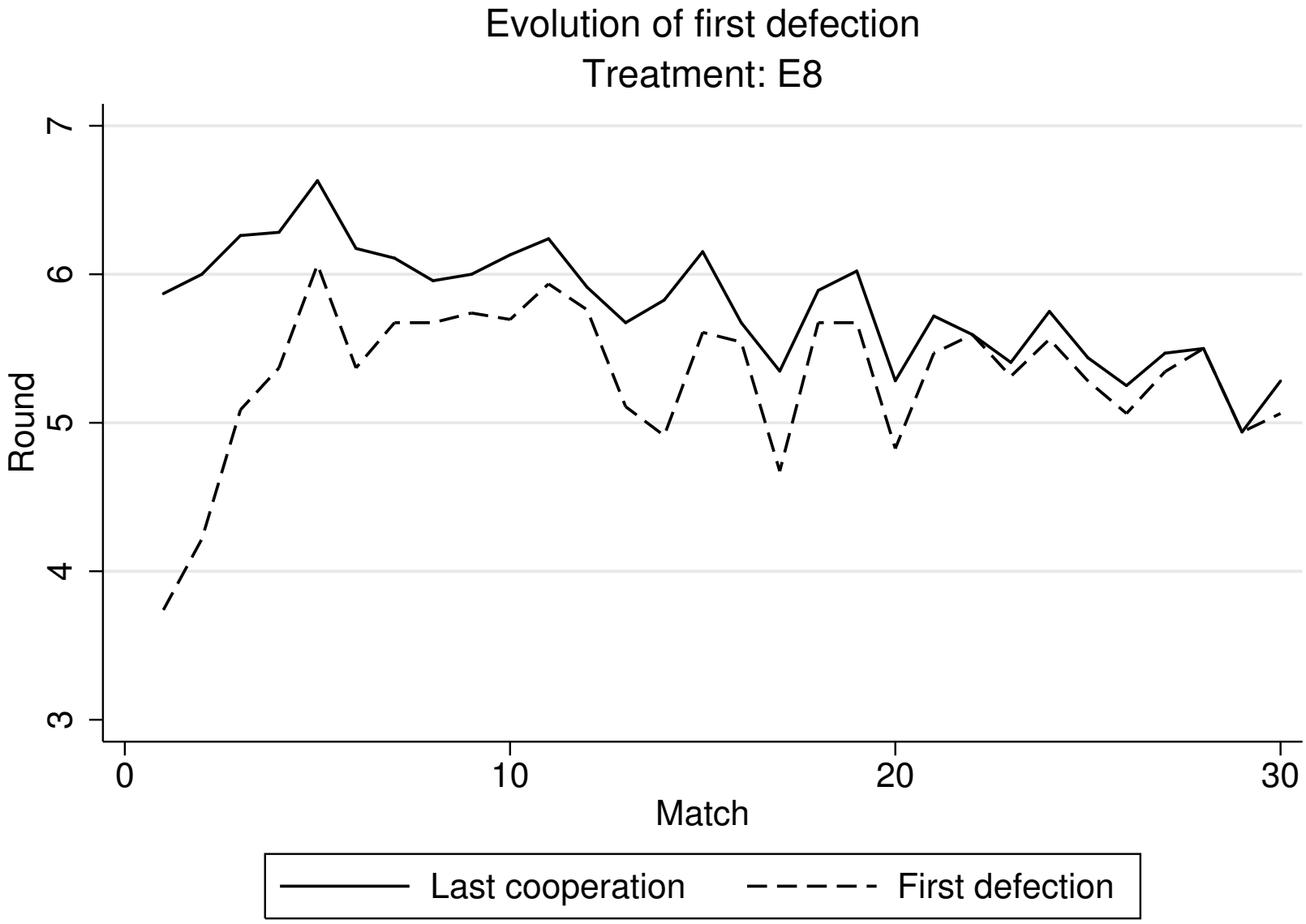
UNRAVELLING OF COOPERATION



BREAKDOWN OF COOPERATION



FIRST DEFECTION AND LAST COOPERATION



CONCLUSION

- Parameters of the game have a significant impact on round 1 cooperation rates.
  - The majority of this effect can be captured by the basin of attraction.
- Despite differences in round 1 behavior, subjects behavior is consistent with the logic of backward induction.