Learning and Sophistication in Coordination Games

Kyle Hyndman[1] Antoine Terracol[2] Jonathan Vaksmann[3]

[1]Southern Methodist University, Dallas, TX

[2]EQUIPPE, Universités de Lille and Centre d'Économie de la Sorbonne, Université Paris 1

[3]GAINS-TEPP, Université du Maine and Centre d'Économie de la Sorbonne, Université Paris 1

Workshop AlgeCoFail Dec. 2009

Introduction I

Motivations

- Behavioral approaches used to describe players' behavior regard people as purely adaptive learners who only best respond to what they have experienced in the past without any awareness of the impact of their actions on their opponents' behavior.
- Along these approaches strategic interactions do not play any role in games!
- Thus a few recent studies exhibit sophistication into players' behavior. In these approaches players might realize that their opponents are capable of learning and could use this opportunity to play strategically and manipulate them.
- This is how strategic teaching might arise.

Introduction II

Previous Research

- Camerer, Ho and Chong (2002), devised a model of strategic teaching in a population of players. A fraction of them is purely adaptive as postulated by usual learning models and the remaining fraction of players is fully sophisticated and can teach them.

Other studies focus on teaching in fixed pairs of players.

- Ehrblatt, Hyndman, Ozbay, Schotter (2009): Teaching a rapid learner facilitates convergence to a unique NE.
- Terracol and Vaksmann (2009): More tenacious teachers take the leadership and drive coordination.

Our goal: Highlighting the determinants of strategic behavior.

Experimental Design I

The Experimental Games

Table: Payoff Matrices

TP_H/TC_L	X	Y	TP_H/TC_H	X	Y
X	**40,45**	8,37	X	**40,45**	0,37
Y	39,0	**12,32**	Y	37,0	**12,32**

TP_L/TC_L	X	Y	TP_L/TC_H	X	Y
X	**20,45**	8,37	X	**20,45**	0,37
Y	19,0	**12,32**	Y	17,0	**12,32**

Game structure: two pure strategy Nash equilibria: **(X,X)** and **(Y,Y)** and one MSNE: $\{(0.8, 0.2); (0.8, 0.2)\}$

Experimental Design II

Teaching Incentives

Teaching as an investment: Players are likely to forego short-run payoffs to teach and get more in the long-run.

- *Teaching Cost* (Optimization Premium for Battalio *et al.* Ecta 2002): $E_i^Y(p) - E_i^X(p) = \theta_i\,(0.8 - p)$, p =prob. attached to X. Where

$$\theta_i = \pi_i\,(X, X) - \pi_i\,(X, Y) + \pi_i\,(Y, Y) - \pi_i\,(Y, X)\,.$$

- *Teaching Premium*: $\psi_i = \frac{\pi_i(X,X)-\pi_i(Y,Y)}{\pi_i(Y,Y)}$, i =Row, Column.

Experimental Design III

Teaching Incentives

Table: Row Players' Incentives For Teaching

Game	ψ_r	θ_r
$\mathrm{TP}_H/\mathrm{TC}_L$	2.33	5
$\mathrm{TP}_H/\mathrm{TC}_H$	2.33	15
$\mathrm{TP}_L/\mathrm{TC}_L$	0.67	5
$\mathrm{TP}_L/\mathrm{TC}_H$	0.67	15

Column players' teaching incentives remain unchanged through games: $\psi_C = 0.4$, $\theta_C = 40$.

Experimental Design IV

The Data

- Parisian Experimental Economics Laboratory (LEEP).
- 30-40 subjects in each game.
- 20 repetitions of each stage game, $\simeq$1hour and €13.5 on average.
- In each period, prior to choosing an action, players are asked (and incentivized) to report their beliefs.

Belief Formation Process (BFP) I

- Precondition for teaching: Players' might take strategic interactions into account.
- Usual proxies used to describe players' BFP postulate that strategic considerations do not play any role.
- Test of a Sophistication Bias: The impact of players' previous action on their BFP.

Belief Formation Process (BFP) II

Empirical Strategy

- Usual Proxies

$$B_i^a(t+1) = \frac{\mathbb{1}_{\{a_j(t)=a\}} + \sum_{u=1}^{t-1} \gamma^u \mathbb{1}_{\{a_j(t-u)=a\}}}{1 + \sum_{u=1}^{t-1} \gamma^u}$$

  $0 \leq \gamma \leq 1$.
  $\gamma = 0 \Rightarrow$ Cournot model.
  $\gamma = 1 \Rightarrow$ Fictitious Play model.

- Elicited Beliefs (using a standard quadratic scoring rule), $b_i^a(t)$.
- Belief Differences, $D_i^a(t) = b_i^a(t) - B_i^a(t)$.

Empirical strategy: A positive impact of $\mathbb{1}_{\{a_j(t-1)=a\}}$ on $D_i^a(t)$ indicates the presence of a sophistication bias.

Belief Formation Process (BFP) — Results

Table: Random-Effects Panel Regression: The Sophistication Bias

	TP_H/TC_L	TP_H/TC_H	TP_L/TC_L	TP_L/TC_H
All	0.149^{***} (0.032)	0.210^{***} (0.041)	0.137^{***} (0.042)	0.187^{***} (0.062)
Row players	0.138^{***} (0.046)	0.230^{***} (0.068)	0.167^{**} (0.065)	0.173^{*} (0.091)
Column players	0.163^{***} (0.045)	0.195^{***} (0.049)	0.098^{**} (0.048)	0.199^{**} (0.086)

* 10% level of significance; ** 5% level of significance; *** 1% level of significance.
Robust standard errors in parentheses.

Choice Behavior

A player over responds to a given action when he plays this action despite the fact that it is not a best response to his static beliefs.

Table: Frequency of Choice Behaviour Categorised By Best Response

ROW PLAYERS

TP_h/TC_ℓ

	$BR = X$	$BR = Y$
X	0.25	0.38
Y	0.02	0.36

TP_h/TC_h

	$BR = X$	$BR = Y$
X	0.31	0.26
Y	0.01	0.42

TP_ℓ/TC_ℓ

	$BR = X$	$BR = Y$
X	0.37	0.23
Y	0.04	0.36

TP_ℓ/TC_h

	$BR = X$	$BR = Y$
X	0.29	0.17
Y	0.06	0.48

The numbers in each matrix should sum to 1, modulo rounding.

Choice Behavior

Table: Frequency of Choice Behaviour Categorised By Best Response

COLUMN PLAYERS

TP_h/TC_ℓ

	$BR = X$	$BR = Y$
X	0.27	0.24
Y	0.04	0.45

TP_h/TC_h

	$BR = X$	$BR = Y$
X	0.37	0.19
Y	0.02	0.43

TP_ℓ/TC_ℓ

	$BR = X$	$BR = Y$
X	0.39	0.18
Y	0.03	0.40

TP_ℓ/TC_h

	$BR = X$	$BR = Y$
X	0.29	0.20
Y	0.04	0.47

The numbers in each matrix should sum to 1, modulo rounding.

Table: Two-sample t-tests Across Games: Frequency of Over Response to X.

ROW PLAYERS

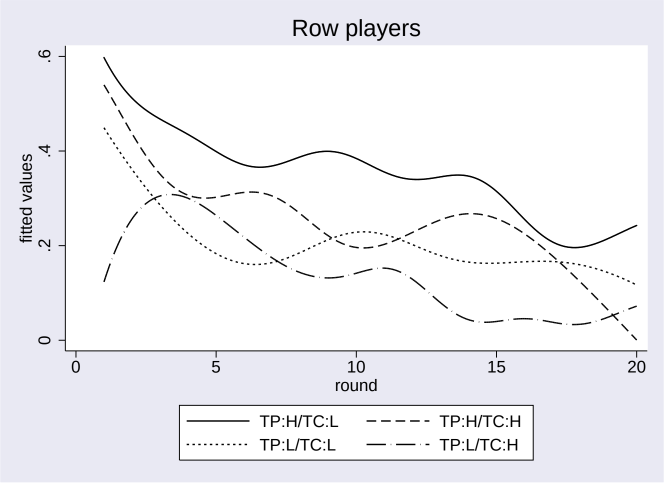
	TP_H/TC_L	TP_H/TC_H	TP_L/TC_L	TP_L/TC_H
TP_H/TC_L	-	1.75*	2.79***	4.19***
TP_H/TC_H	-	-	0.83	2.03**
TP_L/TC_L	-	-	-	1.26
TP_L/TC_H	-	-	-	-

COLUMN PLAYERS

	TP_H/TC_L	TP_H/TC_H	TP_L/TC_L	TP_L/TC_H
TP_H/TC_L	-	0.94	1.52	0.56
TP_H/TC_H	-	-	0.54	0.30
TP_L/TC_L	-	-	-	-0.79
TP_L/TC_H	-	-	-	-

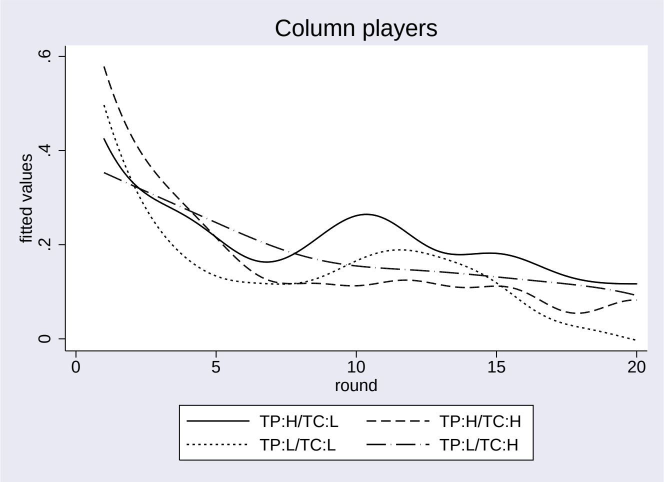
Choice Behavior—Dynamic pattern

Proportion of over responses to X.

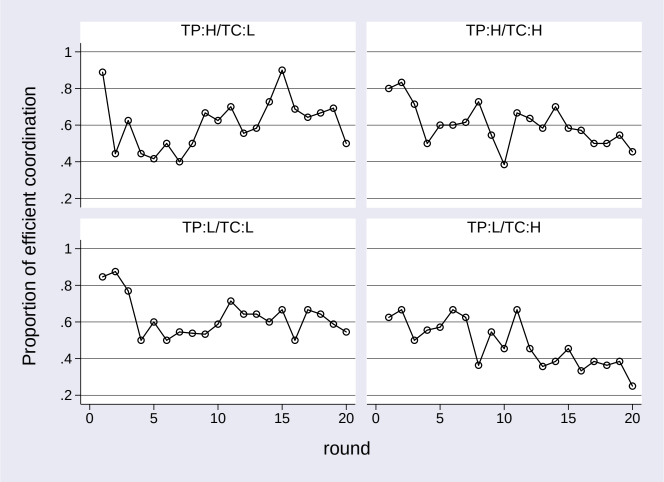


Choice Behavior—Dynamic pattern

Proportion of over responses to X.



Coordination



Tracking players' behavior I

A Model of Sophisticated Learning I

Players see their opponent as a γ-learner:

- Teachers can build their opponent's beliefs and actions and are allowed to re-evaluate their opponent's responsiveness (its parameter γ) at each period on the basis on the information gathered.
- $\Rightarrow$ Continuation strategies: $\sigma_i(t) = (a_i(t), a_i(t+1), ..., a_i(T))$.

Tracking players' behavior II

A Model of Sophisticated Learning II

Players seek to maximize their intertemporal expected payoffs

$$\begin{aligned} E_i(\sigma_i^a(t)) &= b_i^X(t) \cdot \pi_i(a, X) + (1 - b_i^X(t)) \cdot \pi_i(a, Y) \\ &+ \sum_{u=t+1}^{T} \delta^{u-t} \sum_{z=X,Y} b_i^z(u|\sigma^a(t)) \cdot \pi_i(a, z) \end{aligned}$$

When $\delta = 0$, the red part vanishes and the model reduces to the adaptive/myopic model.

Tracking players' behavior III

A Model of Sophisticated Learning III

As usual in this kind of models, we assume that players optimize stochastically.

Players' choice probabilities:

$$P_i^X(t) = \frac{\exp\left[\lambda\left[E_i(\sigma^X(t)) - E_i(\sigma^Y(t))\right]\right]}{1 + \exp\left[\lambda\left[E_i(\sigma^X(t)) - E_i(\sigma^Y(t))\right]\right]}.$$

$$P_i^Y(t) = 1 - P_i^X(t).$$

Where, $\lambda > 0$. When $\lambda \to 0$, players tend to randomize over the set of actions. When $\lambda \to +\infty$, players tend to optimize deterministically.

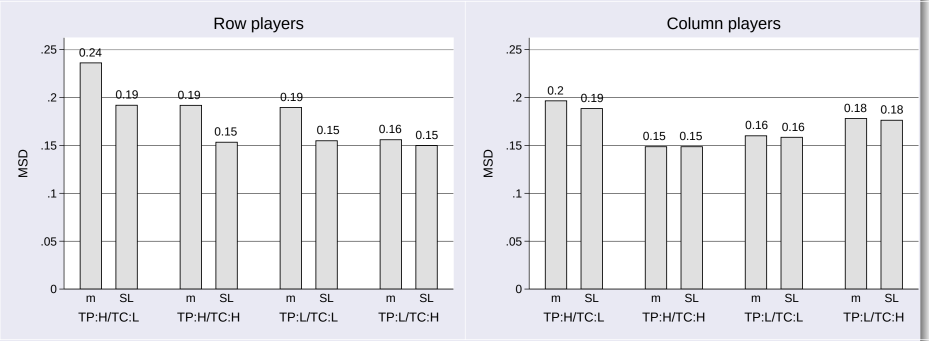
Tracking players' behavior—Estimations

Table: Estimations for each type in each game

		Myopic Model				SL Model			
		TP_H/TC_L	TP_H/TC_H	TP_L/TC_L	TP_L/TC_H	TP_H/TC_L	TP_H/TC_H	TP_L/TC_L	TP_L/TC_H
Row players	λ	0.215** (0.086)	0.192*** (0.036)	0.555*** (0.138)	0.259*** (0.043)	0.394*** (0.089)	0.224*** (0.031)	0.581*** (0.108)	0.231*** (0.051)
	δ	-	-	-	-	0.114*** (0.027)	0.187*** (0.034)	0.228*** (0.046)	0.224 (0.235)
	N	340	320	380	300	340	320	380	300
	LL	-	-	-	-	-	-	-	-
		226.208	181.323	215.601	141.961	190.115	151.519	182.614	138.354
Column players	λ	0.070*** (0.019)	0.112*** (0.020)	0.096*** (0.025)	0.073*** (0.013)	0.051*** (0.016)	0.112*** (0.020)	0.061*** (0.018)	0.047* (0.026)
	δ	-	-	-	-	0.483 (0.333)	0 (0)	0.569** (0.291)	0.561 (0.726)
	N	340	320	380	300	340	320	380	300
	LL	-	-	-	-	-	-	-	-
		199.732	151.371	192.131	160.169	194.102	151.371	190.316	159.929

Tracking players' behavior—MSD

$$MSD = \frac{1}{N}\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[\mathbb{1}_{\{a_i(t)=X\}} - P_i^X(t)\right]^2.$$



Conclusion

- Sophistication Bias: Players think more strategically than postulated by usual theories of learning $\Rightarrow$ This paves the way for strategic sophistication.
- When players are given high incentives to teach, they are particularly likely to over respond, i.e. they forego short-run payoffs to get more in the long-run. Doing so promotes efficiency.
- When players are given high incentives to teach, learning models are particularly limited. Adding a forward-looking component significantly improves the fit and provides a unifying framework to account for different types of behavior.