Quiz

1. How many "families" of games are used in the experiments in the paper?
2. True or False: The authors find strong correlations between measures of cognitive ability of an individual and the individual's strategic sophistication.
3. True or False: The authors find that behavior is strongly responsive to information about the opponent's cognitive ability.

GHW (2015)

- GHW (2015) starts with a set Γ of games, with generic element $\gamma = (\{i, j\}, S, u)$
- γ is exogenous and determined by the experimenter in each period

GHW (2015)

- GHW (2015) starts with a set Γ of games, with generic element $\gamma = (\{i, j\}, S, u)$
- γ is exogenous and determined by the experimenter in each period
- Player receives some signal $\tau_i \in T$ about his partner's type

GHW (2015)

- GHW (2015) starts with a set Γ of games, with generic element $\gamma = (\{i, j\}, S, u)$
- γ is exogenous and determined by the experimenter in each period
- Player receives some signal $\tau_i \in T$ about his partner's type
- Player i's own type is $\theta_i = (c_i, k_i)$, where $k_i : \Gamma \times T \rightarrow \{0, 1, 2, ...\}$ is the player's *cognitive level* and $c_i : \Gamma \rightarrow \{0, 1, 2, ...\}$ is the player's *capacity*
- $k(\tau_i, \gamma) \leq c(\tau_i)$ for all τ_i and γ
- **Could do this without capacities...** Capacity just upper bound on type in each game

Beliefs

- Beliefs for each k_i are defined by some function $v : \mathbb{N}_0 \to \Delta(\mathbb{N}_0)$ satisfying $v(k)(\{0, 1, 2, ..., k-1\}) = 1$ for all $k > 0$

Beliefs

- Beliefs for each k_i are defined by some function $v : \mathbb{N}_0 \to \Delta(\mathbb{N}_0)$ satisfying $v(k)(\{0, 1, 2, ..., k-1\}) = 1$ for all $k > 0$

- For example, one common assumption is that every player things that everyone else is one level below them. In this case, $v(k)(k-1) = 1$ and $v(k)(l) = 0$ for all $l \neq k-1$ (This is actually the assumption used in the paper...)

Strategies

- Level-0 chooses some fixed strategy $\sigma_i^0 \in \Delta(S_i)$
- For example, chooses a random action

Strategies

- Level-0 chooses some fixed strategy $\sigma_i^0 \in \Delta(S_i)$
- For example, chooses a random action
- Level-k best-responds to beliefs $v(k)$, assuming each type $\kappa < k$ does the same
- Formally,

$$\sigma_i^k = \arg\max_{\sigma_i \in \Sigma_i} \left\{ \sum_{\kappa} u_i(\sigma_i, \sigma_{-i}^{\kappa}) v(k)(\kappa) \right\}$$

Examples

- $(c_i, k_i) = (0, 0) \Rightarrow$ Player i plays σ_i^0

Examples

- $(c_i, k_i) = (0, 0) \Rightarrow$ Player i plays σ_i^0
- $(c_i, k_i) = (1, 0) \Rightarrow$ Player i plays σ_i^0

Examples

- $(c_i, k_i) = (0, 0) \Rightarrow$ Player i plays σ_i^0
- $(c_i, k_i) = (1, 0) \Rightarrow$ Player i plays σ_i^0
- $(c_i, k_i) = (1, 1) \Rightarrow$ Player i plays σ_i^1

Examples

- $(c_i, k_i) = (0, 0) \Rightarrow$ Player i plays σ_i^0
- $(c_i, k_i) = (1, 0) \Rightarrow$ Player i plays σ_i^0
- $(c_i, k_i) = (1, 1) \Rightarrow$ Player i plays σ_i^1
- $(c_i, k_i) = (2, 1) \Rightarrow$ Player i plays σ_i^1
- Notice the chosen strategy depends only on the level, not on the capacity

Predictions on cognitive levels

(1) **Constant:** $k_i(\gamma, \tau_i) = k_i(\gamma', \tau_i')$ for all i, γ, γ', τ_i, and τ_i'.
(2) **Constant Across Games:** $k_i(\gamma, \tau_i) = k_i(\gamma', \tau_i)$ for all i, γ, γ', and τ_i.
(3) **Constant Ordering:** If $k_i(\gamma, \tau) \geq k_j(\gamma, \tau)$ for some γ and τ then $k_i(\gamma', \tau') \geq k_j(\gamma', \tau')$ for all γ' and τ'.
(4) **Responsiveness to Signals:** For every γ and i there is some τ and τ' such that $k_i(\gamma, \tau) > k_i(\gamma, \tau')$.
(5) **Consistent Ordering of Games:** For any τ, if $k_i(\gamma, \tau) \geq k_i(\gamma', \tau)$ for some i, γ and γ', then $k_j(\gamma, \tau) \geq k_j(\gamma', \tau)$ for all j.

Undercutting games

The experiment uses two families of games. The games in the first family ("undercutting") look like this:

	1	2	3	4	5	6	7
1	1 1	10 -10	0 0	0 0	0 0	0 0	-11 0
2	-10 10	0 0	10 -10	0 0	0 0	0 0	0 0
3	0 0	-10 10	0 0	10 -10	0 0	0 0	0 0
4	0 0	0 0	-10 10	0 0	10 -10	10 -10	10 -10
5	0 0	0 0	0 0	-10 10	0 0	0 0	0 0
6	0 0	0 0	0 0	-10 10	0 0	0 0	0 0
7	0 -11	0 0	0 0	-10 10	0 0	0 0	-11 -11

Fig. 1. Undercutting game 1 (UG1).

K-level logic: ...

Undercutting games

- Although the paper says these strategies are dominated, 5 and 6 are not weakly or strictly dominated
- Not sure what the point of 5-7 is? Not motivated
- New game. Why not used established game? Bigger action space not convincing argument

Guessing game

- Player a chooses a number in $[a_i, b_i]$, motivated to guess some fraction p_i of the other player's number
- Helpful to look at instructions: `http://piotr-evdokimov.com/persistence_instructions.pdf`
- Lets think about an example with $A_1 = [300, 500]$, $p_1 = 0.7$, $A^2 = [100, 900]$, $p_2 = 0.5$

IESDA, Step 1

Recall that $A_1 = [300, 500]$, $p_1 = 0.7$, $A^2 = [100, 900]$, $p_2 = 0.5$

$$
\begin{aligned}
300 \leq a_1 \leq 500 &\Rightarrow p_2 300 \leq p_2 a_1 \leq p_2 500 \\
&\Rightarrow 150 \leq p_2 a_1 \leq 250 \\
&\Rightarrow 150 \leq a_2 \leq 250
\end{aligned}
$$

$\Rightarrow$ A rational player 2 will not choose actions below 150 or above 250 (can eliminate those)

IESDA, Step 2

Now, $A_1 = [300, 500]$, $p_1 = 0.7$, $A_2^2 = [150, 250]$, $p_2 = 0.5$

$$\begin{aligned} 150 \leq a_2 \leq 250 &\Rightarrow p_1 150 \leq p_1 a_2 \leq p_1 250 \\ &\Rightarrow 105 \leq p_1 a_2 \leq 175 \end{aligned}$$

$\Rightarrow$ Player's 1's target falls outside the range of available actions. The best he can do is choose 300.

IESDA, Step 2

Now, $A_1 = [300, 500]$, $p_1 = 0.7$, $A_2^2 = [150, 250]$, $p_2 = 0.5$

$$\begin{aligned} 150 \leq a_2 \leq 250 \Rightarrow p_1 150 \leq p_1 a_2 \leq p_1 250 \\ \Rightarrow 105 \leq p_1 a_2 \leq 175 \end{aligned}$$

$\Rightarrow$ Player's 1's target falls outside the range of available actions. The best he can do is choose 300.

- **Step 3:** Player 2 best responds to 300 by choosing $p_2 \times 300 = 150$.

Level-k

Recall that $A_1 = [300, 500]$, $p_1 = 0.7$, $A^2 = [100, 900]$, $p_2 = 0.5$

Level	**Player 1's choice**	**Player 2's choice**
0	400	500
1	350	200
2	300	175
3	300	150

Level-k

Recall that $A_1 = [300, 500]$, $p_1 = 0.7$, $A^2 = [100, 900]$, $p_2 = 0.5$

Level	**Player 1's choice**	**Player 2's choice**
0	400	500
1	350	200
2	300	175
3	300	150

- Compare this to undominated actions:

Step	**Player 1's undominated actions**	**Player 2's undominated actions**
0	[300,500]	[100,900]
1	[300,500]	[150,250]
2	300	[150,250]
3	300	150

The difference in the two sets of predictions is explored in Crawford and Costa-Gomes (2006)

Experimental Design

Subjects played 10 games:

Table 1
The ten games used in the experiment.

Game ID	Game type	Player's limits & target	Opponent's limits & target
UG1	Undercutting game	See Fig. 1	
UG2	Undercutting game	See Fig. 2	
UG3	Undercutting game	See Fig. 3	
UG4	Undercutting game	See Fig. 4	
GG5	Guessing game	([215, 815], 1.4)	([0, 650], 0.9)
GG6	Guessing game	([100, 500], 0.7)	([300, 900], 1.3)
GG7	Guessing game	([100, 500], 0.5)	([100, 900], 1.3)
GG8	Guessing game	([0, 650], 0.9)	([215, 815], 1.4)
GG9	Guessing game	([300, 900], 1.3)	([100, 500], 0.7)
GG10	Guessing game	([100, 900], 1.3)	([100, 500], 0.5)

- One against a random subject
- One against a high type
- One against a low type

Types determined by quizzes

Eye gaze quiz

☐ jealous ☐ panicked ☐ arrogant ☐ hateful

☐ aghast ☐ fantasizing ☐ impatient ☐ alarmed

Fig. 5. Sample questions from the Eye Gaze test.

CRT

A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost?

CRT

If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?

CRT

In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

MLE and Discrete Choice

- Assume that an individual made T choices $c_1, c_2, ..., c_T$
  - I.e., are $\{1, 2, ..., T\}$ observations, with some choice c_t chosen in each observation

MLE and Discrete Choice

- Assume that an individual made T choices $c_1, c_2, ..., c_T$
  - I.e., are $\{1, 2, ..., T\}$ observations, with some choice c_t chosen in each observation
- K options in the choice set $c_1, c_2, ..., c_K$

MLE and Discrete Choice

- Assume that an individual made T choices $c_1, c_2, ..., c_T$
  - I.e., are $\{1, 2, ..., T\}$ observations, with some choice c_t chosen in each observation
- K options in the choice set $c_1, c_2, ..., c_K$
- Known values $\{u_1, u_2, ..., u_K\}$ of each choice in the choice set (this is part of our model)

MLE and Discrete Choice

- Assume that an individual made T choices $c_1, c_2, ..., c_T$
  - I.e., are $\{1, 2, ..., T\}$ observations, with some choice c_t chosen in each observation
- K options in the choice set $c_1, c_2, ..., c_K$
- Known values $\{u_1, u_2, ..., u_K\}$ of each choice in the choice set (this is part of our model)
- Utility determining each choice t is random, with $U_{kt} = \lambda u_k + \epsilon_{kt}$, where ϵ_{kt} is an i.i.d. random error from an extreme value distribution

MLE and Discrete Choice

- Assume that an individual made T choices $c_1, c_2, ..., c_T$
  - I.e., are $\{1, 2, ..., T\}$ observations, with some choice c_t chosen in each observation
- K options in the choice set $c_1, c_2, ..., c_K$
- Known values $\{u_1, u_2, ..., u_K\}$ of each choice in the choice set (this is part of our model)
- Utility determining each choice t is random, with $U_{kt} = \lambda u_k + \epsilon_{kt}$, where ϵ_{kt} is an i.i.d. random error from an extreme value distribution
- For each t, decision maker compares U_{kt}'s and chooses the k with the highest U_{kt}

MLE and Discrete Choice

- If ϵ_{kt} has an extreme value distribution, then

$$p(c_k;\lambda) = \frac{exp(\lambda u_k)}{\sum_j exp(\lambda u_j)}$$

MLE and Discrete Choice

- If ϵ_{kt} has an extreme value distribution, then

$$p(c_k; \lambda) = \frac{exp(\lambda u_k)}{\sum_j exp(\lambda u_j)}$$

- What is the probability of observing our data set? This is called the *likelihood function*:

$$L(\lambda) = \prod_t \left(p(c_1; \lambda)\right)^{I(c_t = c_1)} \left(p(c_2; \lambda)\right)^{I(c_t = c_2)} \ldots \left(p(c_K; \lambda)\right)^{I(c_t = c_K)}$$

- Take the derivative, set it equal to zero. This is the MLE estimate $\hat{\lambda}$ of λ

MLE and Discrete Choice

- If ϵ_{kt} has an extreme value distribution, then

$$p(c_k;\lambda) = \frac{exp(\lambda u_k)}{\sum_j exp(\lambda u_j)}$$

- What is the probability of observing our data set? This is called the *likelihood function*:

$$L(\lambda) = \prod_t \left(p(c_1;\lambda)\right)^{I(c_t=c_1)} \left(p(c_2;\lambda)\right)^{I(c_t=c_2)} \dots \left(p(c_K;\lambda)\right)^{I(c_t=c_K)}$$

- Take the derivative, set it equal to zero. This is the MLE estimate $\hat{\lambda}$ of λ
- As $\lambda \to 0$, choice probabilities converge to a uniform distribution
- As $\lambda \to \infty$, choice converges to putting probability one on the option with the highest u_k

Identification of k-levels in the data

- For each type $k \in \{0, 1, 2, 3, N\}$ and each subject i, an estimated likelihood function $L(\lambda_{ik}, \epsilon_{ik}; k)$ of observing the data is estimated
- λ_{ik} and ϵ_{ik} are subject- and $k-$specific parameters
  - ϵ_{ik}: probability of making the choice a $k-$type would make if $v(k)(k-1) = 1$
  - $1 - \epsilon_{ik}$: probability of choice coming from a random choice model
  - λ_{ik}: a noise parameter (as before)
- The k associated with the highest likelihood is defined as the subject's k

Results

- Stability of the k-estimates
- Quiz scores and k-levels
- Effect of information about other people

Results: Are types stable across games and families?

Table 2
Frequency of levels in each game, and when pooling each family of games.

Game	L0	L1	L2	L3	Nash
UG1	7.76%	32.76%	19.83%	10.34%	29.31%
UG2	7.76%	32.76%	22.41%	7.76%	29.31%
UG3	5.17%	27.59%	18.10%	5.17%	43.97%
UG4	6.03%	31.03%	29.31%	5.17%	28.45%
UGs pooled	4.31%	28.45%	26.72%	5.17%	35.34%
GG5	6.03%	70.69%	9.48%	12.07%	1.72%
GG6	0.86%	65.52%	17.24%	11.21%	5.17%
GG7	43.10%	37.07%	13.79%	1.72%	4.31%
GG8	6.90%	39.66%	24.14%	21.55%	7.76%
GG9	5.17%	42.24%	23.28%	4.31%	25.00%
GG10	9.48%	38.79%	24.14%	19.83%	7.76%
GGs pooled	1.72%	50.00%	10.34%	10.34%	27.59%

- Levels jump around in the guessing games
- 14.22% of observations in the guessing games correspond exactly to some level-k choice
- **Most common: players best-respond to midpoint of opponent's interval**

Results: Transitions

Table 3
Markov transitions from the pooled undercutting games to the pooled guessing games.

From ↓ to →	L0	L1	L2	L3	Nash
L0	0.0%	**60.0%**	0.0%	20.0%	20.0%
L1	6.1%	**42.4%**	6.1%	9.1%	36.4%
L2	0.0%	**51.6%**	16.1%	9.7%	22.6%
L3	0.0%	*33.3%*	*33.3%*	0.0%	*33.3%*
Nash	0.0%	**56.1%**	7.3%	12.2%	24.4%
Overall	1.7%	50.0%	10.3%	10.3%	27.6%

- Here, $k_i(\Gamma)$ is estimated for each subject using all games from family Γ for both families separately
- How would we expect this table to look if $k_i(\Gamma) = k_i(\Gamma')$?
- Most transitions from undercutting to guessing are into L1 types and Nash types

Results: Transitions (Undercutting)

Table 4
Markov transition between single-game levels within the four undercutting games.

From ↓ to →	L0	L1	L2	L3	Nash
L0	**43.0%**	22.6%	7.5%	9.7%	17.2%
L1	4.9%	**59.7%**	14.6%	4.4%	16.4%
L2	2.2%	20.2%	**57.1%**	9.0%	11.5%
L3	9.1%	19.2%	**28.3%**	18.2%	25.3%
Nash	3.5%	15.6%	7.9%	5.5%	**67.5%**
Overall	6.7%	31.0%	22.4%	7.1%	32.8%

- More stability... why?

Results: Transitions (Guessing)

Table 5
Markov transition between single-game levels within the six guessing games.

From ↓ to →	L0	L1	L2	L3	Nash
L0	8.7%	**48.2%**	18.1%	12.3%	12.8%
L1	11.7%	**53.1%**	16.8%	11.2%	7.1%
L2	11.5%	**44.2%**	27.4%	10.0%	6.9%
L3	12.4%	**46.6%**	15.9%	13.2%	12.0%
Nash	17.7%	**40.3%**	15.0%	16.3%	10.7%
Overall	11.9%	**49.0%**	18.7%	11.8%	8.6%

- Transition probabilities don't depend very much on type in the guessing game

Results: Quiz scores

- **Undercutting games:** The only predictive quiz is Eye Gaze ($P = 0.015$)
- **Guessing games:** The only predictive quiz is CRT ($P = 0.01$)
- The authors claim the latter correlation is spurious... Why not the former?
- Why not look at the effect of total quiz score (since this serves as τ) on levels?

Effect of τ (undercutting)

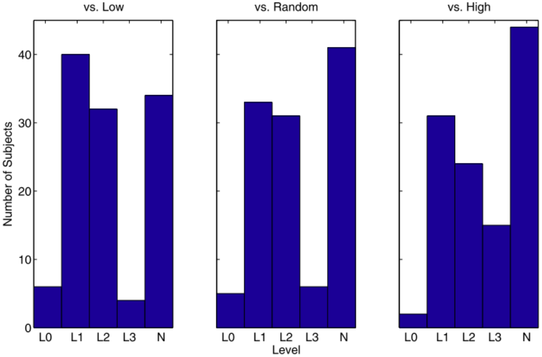
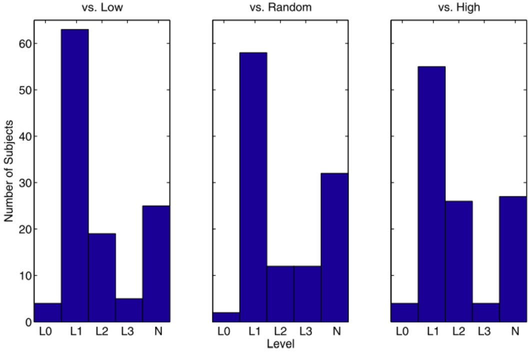


Fig. 8. Level distributions by opponent in the pooled undercutting games.

- Sig. difference for low vs. high ($P < 0.05$)
- No other pairwise differences

Effect of τ (guessing)



- No sig. difference between low and random or low and high
- Sig. difference for random vs. high ($P < 0.05$)
- ...But mean for random is higher than mean for high

Results (Summary)

- **K-levels fluctuate across games and families**
- **K-levels don't correlate with individual characteristics**
- **K-levels are not affected by signals about others**
  - People do not think strategically?
  - Or wrong signals?