Learning in Games

Teck H. Ho
University of California, Berkeley

Joint work with Colin Camerer and Juin-Kuan Chong
Outline

- Research Question

- Criteria of a Good Model

- Experience Weighted Attraction (EWA) Learning Model
  - Choice Reinforcement
  - Weighted Fictitious Play

- Self-running EWA Learning
# Median-Effort Game

## The Median Effort Game

<table>
<thead>
<tr>
<th>Xi</th>
<th>Median {X_i}</th>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.85</td>
<td>1.00</td>
<td>1.05</td>
<td>1.00</td>
<td>0.85</td>
<td>0.60</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.50</td>
<td>0.75</td>
<td>0.90</td>
<td>0.95</td>
<td>0.90</td>
<td>0.75</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.05</td>
<td>0.40</td>
<td>0.65</td>
<td>0.80</td>
<td>0.85</td>
<td>0.80</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>−0.50</td>
<td>−0.05</td>
<td>0.30</td>
<td>0.55</td>
<td>0.70</td>
<td>0.75</td>
<td>0.70</td>
<td></td>
</tr>
</tbody>
</table>
Median-Effort Game

Figure 1a: Actual choice frequencies

Van Huyck, Battalio, and Beil (1990)
The learning setting

- In games where each player is aware of the payoff table
- Player $i$’s strategy space consists of discrete choices indexed by $j$ (e.g., 1, 2, …, 7)
- The game is repeated for several rounds
- At each round, all players observed:
  - Strategy or action history of all other players
  - Own payoff history
Research question

- To develop a good descriptive model of adaptive learning to predict the probability of player $i$ ($i=1,\ldots,n$) choosing strategy $j$ at round $t$ in any game

$$P_{ij}(t)$$
Criteria of a “good” model

- Use (potentially) all available information subjects receive in a sensible way
- Satisfies plausible principles of behavior (i.e., conformity with other sciences such as psychology)
- Fits and predicts choice behavior well
- Ability to generate new insights
- As simple as the data allow
Models of Learning

- **Introspection** (\( P^j \)): Requires too much human cognition
  - Nash equilibrium (Nash, 1950)
  - Quantal response equilibrium (Nash-\( \lambda \)) (McKelvey and Palfrey, 1995)

- **Evolution** (\( P^j(t) \)): Players are pre-programmed
  - Replicator dynamics (Friedman, 1991)
  - Genetic algorithm (Ho, 1996)

- **Learning** (\( P^i_j(t) \)): Uses about the right level of cognition
  - Experience-weighted attraction learning (Camerer and Ho, 1999)
    - Reinforcement (Roth and Erev, 1995)
    - Belief-based learning
      - Cournot best-response dynamics (Cournot, 1838)
      - Simple Fictitious Play (Brown, 1951)
      - Weighted Fictitious Play (Fudenberg and Levine, 1998)
  - Directional learning (Selten, 1991)
Information Usage in Learning


- Belief-based Learning (Cournot, 1838; Brown, 1951; Fudenberg and Levine 1998): *form beliefs based on opponents’ action history and choose according to expected payoffs*

- The information used by reinforcement learning is own payoff history and by belief-based models is opponents’ action history

- EWA uses both kinds of information
“Laws” of Effects in Learning

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>M</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>10</td>
</tr>
</tbody>
</table>

Row player’s payoff table

- **Law of actual effect**: successes increase the probability of chosen strategies
  - Teck is more likely than Colin to “stick” to his previous choice (other things being equal)

- **Law of simulated effect**: strategies with simulated successes will be chosen more often
  - Colin is more likely to switch to T than M

- **Law of diminishing effect**: Incremental effect of reinforcement diminishes over time
  - $1 has more impact in round 2 than in round 7.

Colin chose B and received 4
Teck chose M and received 5
Their opponents chose L
Assumptions of Reinforcement and Belief Learning

- Reinforcement learning ignores simulated effect
- Belief learning predicts actual and simulated effects are equally strong
- EWA learning allows for a positive (and smaller than actual) simulated effect
The EWA Model

- Initial attractions and experience (i.e., $A_{ij}(0), N(0)$)
- Updating rules

$$A_{ij}(t) = \begin{cases} 
\phi \cdot N(t-1) \cdot A_{ij}(t-1) + \pi_i(s_{ij}, s_{-i}(t)) & s_{ij} = s_i(t) \\
\phi \cdot N(t-1) \cdot A_{ij}(t-1) + \delta \cdot \pi_i(s_{ij}, s_{-i}(t)) & s_{ij} \neq s_i(t) 
\end{cases}$$

$$N(t) = \rho \cdot N(t-1) + 1$$

- Choice probabilities

$$P_{ij}(t+1) = \frac{e^{\lambda \cdot A_{ij}(t)}}{\sum_{k=1}^{m_i} e^{\lambda \cdot A_{ik}(t)}}$$

Camerer and Ho (Econometrica, 1999)
EWA Model and Laws of Effects

\[ A_{ij}(t) = \begin{cases} 
\frac{\phi \cdot N(t-1) \cdot A_{ij}(t-1) + \pi_i(s_{ij}, s_{-i}(t))}{N(t)} & \text{if } s_{ij} = s_i(t) \\
\frac{\phi \cdot N(t-1) \cdot A_{ij}(t-1) + \delta \cdot \pi_i(s_{ij}, s_{-i}(t))}{N(t)} & \text{if } s_{ij} \neq s_i(t)
\end{cases} \]

\[ N(t) = \rho \cdot N(t-1) + 1 \]

- Law of **actual** effect: successes increase the probability of **chosen** strategies (positive incremental reinforcement increases attraction and hence probability)
- Law of **simulated** effect: strategies with simulated successes will be chosen more often (\( \delta > 0 \))
- Law of **diminishing** effect: Incremental effect of reinforcement diminishes over time (\( N(t) \geq N(t-1) \))
The EWA model: An Example

Row player’s payoff table

Period 0: \( A^T (0), A^B (0) \)

Period 1:

\[
A^T (1) = \frac{\phi \cdot A^T (0) \cdot N (0) + \delta \cdot 8}{\rho \cdot N (0) + 1}
\]

\[
A^B (1) = \frac{\phi \cdot A^B (0) \cdot N (0) + 4}{\rho \cdot N (0) + 1}
\]

History: Period 1 = (B,L)
Reinforcement Model: An Example

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>10</td>
</tr>
</tbody>
</table>

Row player’s payoff table

- **Period 0**: $R^T(0)$, $R^B(0)$

- **Period 1**: $R^T(1) = \phi \cdot R^T(0)$
  
  $R^B(1) = \phi \cdot R^B(0) + 4$

  $A^T(1) = \frac{\phi \cdot A^T(0) \cdot N(0) + \delta \cdot 8}{\rho \cdot N(0) + 1}$

  $A^B(1) = \frac{\phi \cdot A^B(0) \cdot N(0) + 4}{\rho \cdot N(0) + 1}$

  If $\delta = 0$, $\rho = 0$, $N(0) = 1$,
  
  $EWA \Rightarrow RF$
Belief-based (BB) model: An Example

- **Period 0:**

\[ N(0) = N^L(0) + N^R(0) \]

\[ B^L(0) = \frac{N^L(0)}{N(0)} \quad B^R(0) = \frac{N^R(0)}{N(0)} \]

\[ E^T(0) = 8 \cdot B^L(0) + 9 \cdot B^R(0) = \frac{N^L(0) \cdot 8 + N^R(0) \cdot 9}{N(0)} \]

\[ E^B(0) = 4 \cdot B^L(0) + 10 \cdot B^R(0) = \frac{N^L(0) \cdot 4 + N^R(0) \cdot 10}{N(0)} \]

- **Period 1:**

\[ B^L(1) = \frac{\rho \cdot N^L(0) + 1}{\rho \cdot N(0) + 1} \quad B^R(1) = \frac{\rho \cdot N^R(0) + 0}{\rho \cdot N(0) + 1} \]

\[ E^T(1) = 8 \cdot B^L(1) + 9 \cdot B^R(1) = \frac{\rho \cdot E^T(0) \cdot N(0) + 8}{\rho \cdot N(0) + 1} \]

\[ E^B(1) = 4 \cdot B^L(1) + 10 \cdot B^R(1) = \frac{\rho \cdot E^B(0) \cdot N(0) + 4}{\rho \cdot N(0) + 1} \]

Bayesian Learning with Dirichlet priors
Relationship between Belief-based (BB) and EWA Learning Models

**BB**

\[
B^L (1) = \frac{\rho \cdot N^L (0) + 1}{\rho \cdot N (0) + 1} \quad B^R (1) = \frac{\rho \cdot N^R (0) + 0}{\rho \cdot N (0) + 1}
\]

\[
E^T (1) = 8 \cdot B^L (1) + 9 \cdot B^R (1)
\]

\[
E^B (1) = 4 \cdot B^L (1) + 10 \cdot B^R (1)
\]

**EWA**

\[
A^T (1) = \frac{\phi \cdot A^T (0) \cdot N (0) + \delta \cdot 8}{\rho \cdot N (0) + 1}
\]

\[
A^B (1) = \frac{\phi \cdot A^B (0) \cdot N (0) + 4}{\rho \cdot N (0) + 1}
\]

If \( \delta = 1, \rho = \phi \)

\[
EWA \implies BB
\]
Model Interpretation

- Simulation or attention parameter ($\delta$): measures the degree of sensitivity to foregone payoffs

- Exploitation parameter ($\kappa = \frac{\phi - \rho}{\phi}$): measures how rapidly players lock-in to a strategy (average versus cumulative)

- Stationarity or motion parameter ($\phi$): measures players’ perception of the degree of stationarity of the environment
Model Interpretation

Weighted Fictitious Play

Cournot

Fictitious Play

Cumulative Reinforcement

Average Reinforcement

September, 2006
Learning in Games
Teck H. Ho
New Insight

- Reinforcement and belief learning were thought to be fundamental different for 50 years.

- For instance, “….in rote [reinforcement] learning success and failure directly influence the choice probabilities. … Belief learning is very different. Here experiences strengthen or weaken beliefs. Belief learning has only an indirect influence on behavior.” (Selten, 1991)

- *EWA shows that belief and reinforcement learning are related and special kinds of EWA learning*
Actual versus Belief-Based Model Frequencies

Figure 1a: Actual choice frequencies

Figure 1b: Belief-based Model frequencies
Actual versus Reinforcement Model Frequencies

Figure 1a: Actual choice frequencies

Figure 1c: Choice reinforcement model frequencies
Actual versus EWA Model Frequencies

Figure 1a: Actual choice frequencies

Figure 1d: EWA model frequencies
## Estimation and Results

| Game Model | No. of Parameters | Calibration | Validation | |
|------------|------------------|-------------|-------------|
|            |                  | LL          | AIC         | BIC | $\rho_2$ | LL | MSD |
| Median Action (M=378) | | | | | |
| 1-segment  | | | | | |
| Random Choice | 0 | -677.29 | -677.29 | -677.29 | 0.0000 | -315.24 | 0.1217 |
| Choice Reinforcement | 8 | -341.70 | -349.70 | -365.44 | 0.4837 | -80.27 | 0.0301 |
| Belief-based | 9 | -438.74 | -447.74 | -465.45 | 0.3389 | -113.90 | 0.0519 |
| EWA | 11 | -309.30 | -320.30 | -341.94* | 0.5271 | -41.05 | 0.0185 |
| 2-Segment  | | | | | |
| Random | 0 | -677.29 | -677.29 | -677.29 | 0.0000 | -315.24 | 0.1217 |
| Choice Reinforcement | 17 | -331.25 | -348.25 | -381.70 | 0.4858 | -66.32 | 0.0245 |
| Belief-based | 19 | -379.24 | -398.24 | -435.62 | 0.4120 | -70.31 | 0.0250 |
| EWA | 23 | -290.25 | -313.25* | -358.51 | 0.5374* | -34.79* | 0.0139* |
Table 1a: A summary of EWA parameter estimates and forecast accuracy (games estimated by us)

<table>
<thead>
<tr>
<th>Citation</th>
<th>Game</th>
<th>$\delta$</th>
<th>$\phi$</th>
<th>$\rho(1-\kappa)$</th>
<th>EWA</th>
<th>Choice reinforcement - EWA</th>
<th>Belief - EWA</th>
<th>In / Out of sample</th>
<th>Fit technique</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camerer, Ho and Hiasa (2000)</td>
<td>Sealed bid mechanism*</td>
<td>n.a.</td>
<td>1.00</td>
<td>0.91</td>
<td>1102.0</td>
<td>30.8</td>
<td>65.5</td>
<td>IN</td>
<td>-LL</td>
<td>$\psi$, $\phi$ &amp; $\kappa$ replace $\delta$ &amp; $\rho$</td>
</tr>
<tr>
<td>Camerer, Ho and Wang (1999)</td>
<td>&quot;Continental divide&quot; coordination</td>
<td>0.75</td>
<td>0.61</td>
<td>0.00</td>
<td>346.9</td>
<td>86.1</td>
<td>235.8</td>
<td>OUT</td>
<td>-LL</td>
<td></td>
</tr>
<tr>
<td>Camerer and Ho (1998)</td>
<td>Weak-link coordination</td>
<td>0.65</td>
<td>0.58</td>
<td>0.20</td>
<td>358.1</td>
<td>29.1</td>
<td>438.6</td>
<td>IN</td>
<td>-LL</td>
<td></td>
</tr>
<tr>
<td>Anderson and Camerer (in press)</td>
<td>Signaling games (game 5) 95% Confidence Interval</td>
<td>0.69 (0.47,1.00)</td>
<td>1.020 (0.99,1.04)</td>
<td>1.00 (0.98,1.00)</td>
<td>72.2</td>
<td>6.5</td>
<td>10.1</td>
<td>OUT</td>
<td>-LL</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Signaling games (game 5) 95% Confidence Interval</td>
<td>0.54 (0.45,0.63)</td>
<td>0.65 (0.59,0.71)</td>
<td>0.46 (0.39,0.54)</td>
<td>139.5</td>
<td>14.1</td>
<td>23.7</td>
<td>OUT</td>
<td>-LL</td>
<td></td>
</tr>
<tr>
<td>Camerer and Ho (1999b)</td>
<td>Median-action coordination</td>
<td>0.85 (0.01)</td>
<td>0.80 (0.02)</td>
<td>0.00 (0.00)</td>
<td>41.1</td>
<td>39.2</td>
<td>72.8</td>
<td>OUT</td>
<td>-LL</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4x4 Mixed-strategy games</td>
<td>0.00 (0.04)</td>
<td>1.04 (0.01)</td>
<td>0.96 (0.01)</td>
<td>326.4</td>
<td>9.1</td>
<td>-40.8</td>
<td>OUT</td>
<td>-LL</td>
<td>Payoff = 5 rupees</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.73 (0.10)</td>
<td>1.01 (0.01)</td>
<td>0.95 (0.01)</td>
<td>341.7</td>
<td>18.0</td>
<td>8.4</td>
<td>OUT</td>
<td>-LL</td>
<td>Payoff = 10 rupees</td>
</tr>
<tr>
<td></td>
<td>6x6 Mixed-strategy games</td>
<td>0.41 (0.08)</td>
<td>0.99 (0.01)</td>
<td>0.94 (0.01)</td>
<td>301.7</td>
<td>6.8</td>
<td>-5.4</td>
<td>OUT</td>
<td>-LL</td>
<td>Payoff = 5 rupees</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.55 (0.05)</td>
<td>0.99 (0.01)</td>
<td>0.93 (0.02)</td>
<td>362.3</td>
<td>13.7</td>
<td>8.9</td>
<td>OUT</td>
<td>-LL</td>
<td>Payoff = 10 rupees</td>
</tr>
<tr>
<td></td>
<td>p-beauty contests*</td>
<td>0.95 (0.01)</td>
<td>0.11 (0.00)</td>
<td>0.00 (0.00)</td>
<td>1917.0</td>
<td>647.0</td>
<td>35.0</td>
<td>OUT</td>
<td>-LL</td>
<td>Experienced and inexperienced combined</td>
</tr>
<tr>
<td>Camerer, Ho and Wang (1999)</td>
<td>Normal form centipede (odd player)</td>
<td>0.32 (0.32)</td>
<td>0.91 (0.14)</td>
<td>0.00 (0.00)</td>
<td>1016.8</td>
<td>57.6</td>
<td>536.3</td>
<td>OUT</td>
<td>-LL</td>
<td>Clairvoyance full update, $\kappa$</td>
</tr>
<tr>
<td></td>
<td>Normal form centipede (even player)</td>
<td>0.24 (0.32)</td>
<td>0.90 (0.14)</td>
<td>0.95 (0.03)</td>
<td>951.3</td>
<td>46.4</td>
<td>604.7</td>
<td>OUT</td>
<td>-LL</td>
<td>Clairvoyance full update, $\kappa$</td>
</tr>
</tbody>
</table>

* In Figure 1, we did not include this study.
* Unlike the previous estimates, these new estimates assume that subjects do not know the winning numbers.
Table 1b: A summary of EWA parameter estimates and forecast accuracy (games estimated by others)

<table>
<thead>
<tr>
<th>CITATION</th>
<th>GAME</th>
<th>EWA estimates (standard error)</th>
<th>Model accuracy</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chen and Khoroshilov (1999)</td>
<td>Cost allocation+</td>
<td>80~1.0</td>
<td>73.88</td>
<td>IN</td>
</tr>
<tr>
<td>Morgan and Selton (1999)</td>
<td>&quot;Unprofitable&quot; games (baseline games)</td>
<td>0.08 (0.07)</td>
<td>1729.5</td>
<td>IN</td>
</tr>
<tr>
<td></td>
<td>&quot;Unprofitable&quot; games (upside games)</td>
<td>0.14 (0.06)</td>
<td>1906.5</td>
<td>IN</td>
</tr>
<tr>
<td>Stahl (1999)</td>
<td>5x5 matrix games</td>
<td>0.66 (0.02)</td>
<td>4803.7</td>
<td>OUT</td>
</tr>
<tr>
<td>Hisa (1999)</td>
<td>Call markets</td>
<td>0.47 (0.32)</td>
<td>1915.0</td>
<td>IN</td>
</tr>
<tr>
<td>Amaldoss (1998)</td>
<td>Same function alliance – equal profit sharing</td>
<td>0.00 (0.00)</td>
<td>886.3</td>
<td>High reward</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.00 (0.00)</td>
<td>767.5</td>
<td>Med. Reward</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.07 (0.01)</td>
<td>1399.7</td>
<td>Low reward</td>
</tr>
<tr>
<td></td>
<td>Same function alliance – proportional sharing</td>
<td>0.00 (0.00)</td>
<td>910.8</td>
<td>High reward</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.00 (0.00)</td>
<td>1055.0</td>
<td>Med. Reward</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.00 (0.00)</td>
<td>1013.7</td>
<td>Low reward</td>
</tr>
<tr>
<td></td>
<td>Parallel development of product – equal sharing</td>
<td>0.00 (0.00)</td>
<td>1194.2</td>
<td>High reward</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.17 (0.01)</td>
<td>1321.5</td>
<td>Med. Reward</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.21 (0.01)</td>
<td>1297.7</td>
<td>Low reward</td>
</tr>
<tr>
<td>Rapoport and Amaldoss (2000)</td>
<td>Patent race game – symmetric players</td>
<td>0.00 (0.00)</td>
<td>3551.7</td>
<td>Low reward</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.00 (0.00)</td>
<td>2908.1</td>
<td>High reward</td>
</tr>
<tr>
<td></td>
<td>Patent race game – asymmetric players</td>
<td>0.48 (0.08)</td>
<td>3031.5</td>
<td>Strong player</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.14 (0.06)</td>
<td>2835.5</td>
<td>Weak player</td>
</tr>
</tbody>
</table>

+ In Figure 1, we did not include this study.
Extensions

- Heterogeneity (JMP, Camerer and Ho, 1999)

- Payoff learning (EJ, Camerer, Ho, and Wang, 2006)

- Sophistication and strategic teaching
  - Sophisticated learning (JET, Camerer, Ho, and Chong, 2002)
  - Reputation building (GEB, Chong, Camerer, and Ho, 2006)

- EWA Lite (Self-tuning EWA learning) (JET, Ho, Camerer, and Chong, forthcoming)

- Applications:
  - Auction markets (Book Chapter, Camerer, Ho, and Hsia, 2000)
  - Product Choice at Supermarkets (JMR, Ho and Chong, 2004)
Learning in Games

- Research Question

- Criteria of a Good Model

- Adaptive Experience Weighted Attraction (EWA) Learning Model
  - Choice Reinforcement
  - Weighted Fictitious Play

- Self-running EWA Learning
Two Open Questions

- **A theory to explain why parameters vary across games**

- **Metrics for judging model performance**
  - Statistical Measures
  - Economic value
EWA Cube

Weighted Fictitious Play

Cournot

Fictitious Play

Cumulative Reinforcement

Average Reinforcement
The EWA Model

- Initial attractions and experience (i.e., $A_{ij}(0), N(0)$)

- Updating rules

$$A_{ij}(t) = \begin{cases} 
\frac{\phi \cdot N(t-1) \cdot A_{ij}(t-1) + \pi_i(s_{ij}, s_{-i}(t))}{N(t)} & s_{ij} = s_i(t) \\
\frac{\phi \cdot N(t-1) \cdot A_{ij}(t-1) + \delta \cdot \pi_i(s_{ij}, s_{-i}(t))}{N(t)} & s_{ij} \neq s_i(t)
\end{cases}$$

$$N(t) = \rho \cdot N(t-1) + 1$$

- Choice probabilities

$$P_{ij}(t+1) = \frac{e^{\lambda \cdot A_{ij}(t)}}{\sum_{k=1}^{m_i} e^{\lambda \cdot A_{ik}(t)}}$$
Self-Tuning EWA Model: Initialization

- Initial attractions determined by CH Model (with $\tau = 1.5$; Camerer, Ho, Chong, 2004, QJE)

- $N(0)=1$ (it fades away quickly with experience anyway)

- Set $\kappa = 0$ (i.e., $\phi = \rho$)
Self-tuning EWA Model:
Change-detector function ($\phi_i(t)$, $t \geq 1$)

- The core of the change-detector function is the surprise index, $S_i(t)$

$$S_i(t) = \sum_{k=1}^{m-i} [h^k_i (t) - r^k_i (t)]^2$$

Distance between history and current round

$$h^k_i (t) = \sum_{\tau=1}^{t} \frac{I(s^k_{-i}, s_{-i} (\tau))}{t}$$

history

$$r^k_i (t) = I(s^k_{-i}, s_{-i} (t))$$

current round

- Change-detector function is:

$$\phi_i (t) = 1 - \frac{1}{2} \cdot S_i(t)$$
Change-detector function \( (\phi_i(t), t>1) \):

**Examples**

- If the other player chooses the strategy she has always chosen before, then \( S_i(t) = 0 \), and \( \phi_i(t) = 1 \).

- If the other player chooses a new strategy which was never chosen before in a very long run of history, \( S_i(t) = 2 \) and \( \phi_i(t) = 0 \).

- If a player chose the same strategy for each of nine periods and a new strategy in period 10, then \( S_i(t) = (0.9-0.0)^2 + (0.1-1.0)^2 = 1.62 \) and \( \phi_i(t) = 1 - 0.5 \times 1.62 = 0.19 \).

- If unique strategies have been played in every period up to \( t-1 \), and another unique strategy is played in period \( t \), then \( \phi_i(t) = 0.5 + 1/(2 \times t) \).
Self-tuning EWA Model:
Attention function \((\delta_{ij}(t), t>1)\)

- Attention function is:

\[
\delta_{ij}(t) = \begin{cases} 
1 & \text{if } \pi_i(s^k_i, s_{-i}(t)) > \pi_i(t) \\
0 & \text{otherwise}
\end{cases}
\]

- Pay attention to only strategies that give strictly better ex-post payoffs because of limited attention
Attention function ($\delta_{ij}(t)$, $t>1$): Examples

- If subjects are strictly best-responding (ex post), then no other strategies have a higher ex-post payoff so $\delta_{ij}(t) = 0$, which reduces the model to choice reinforcement.

- If subjects always choose the worst strategy, then $\delta_{ij}(1) = 1$, which corresponds to weighted fictitious play.

- A natural way to formalize “learning direction” theory.

- Create “exploration-exploitation” shift over time: start with poor choices and then lock in to the best choice.
Table 1: A Description of the Seven Games Used in the Estimation of Various Learning Models

<table>
<thead>
<tr>
<th>Game</th>
<th>Number of Players</th>
<th>Number of Strategies</th>
<th>Number of Pure Strategy Equilibria</th>
<th>Number of Subjects</th>
<th>Number of Rounds</th>
<th>Matching Protocol</th>
<th>Experimental Treatment</th>
<th>Description of Games</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixed Strategies</td>
<td>2</td>
<td>4, 6</td>
<td>0</td>
<td>80</td>
<td>40</td>
<td>Fixed</td>
<td>Stake Size</td>
<td>A constant-sum game with unique mixed strategy equilibrium.</td>
</tr>
<tr>
<td>Mookerjee and Sopher (1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patent Race</td>
<td>2</td>
<td>5, 6</td>
<td>0</td>
<td>36</td>
<td>80</td>
<td>Random</td>
<td>Strong vs Weak</td>
<td>Strong (weak) player invests between 0 and 5 (0 and 4) and the higher investment wins a fixed prize.</td>
</tr>
<tr>
<td>Rapoport and Amaldoss (2000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continental Divide</td>
<td>7</td>
<td>14</td>
<td>2</td>
<td>70</td>
<td>15</td>
<td>Fixed</td>
<td>None</td>
<td>A coordination game with two pure strategy equilibria</td>
</tr>
<tr>
<td>Van Huyck et al. (1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median Action</td>
<td>9</td>
<td>7</td>
<td>7</td>
<td>54</td>
<td>10</td>
<td>Fixed</td>
<td>None</td>
<td>A order-statistic game with individual payoff decreases in the distance between individual choice and the median</td>
</tr>
<tr>
<td>Van Huyck et al. (1990)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pot Games</td>
<td>3, 6, 9, 18</td>
<td>2</td>
<td>1</td>
<td>84</td>
<td>25 (manual)</td>
<td>Fixed</td>
<td>Number of Players</td>
<td>An entry game where players must decide which of the two ponds of sizes 2n and n they wish to enter. Payoff is the ratio of the pond size and number of entries.</td>
</tr>
<tr>
<td>Amaldoss and Ho (2001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>29 (computer)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price Matching</td>
<td>2</td>
<td>121</td>
<td>1</td>
<td>52</td>
<td>10</td>
<td>Random</td>
<td>Penalty Size</td>
<td>Players choose claims between 80 and 200. Both players get lower claim but the high-claim player pays a penalty to the low-claim player.</td>
</tr>
<tr>
<td>(Traveller's Dilemma)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capra et al. (1999)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-Beauty Contest</td>
<td>7</td>
<td>101</td>
<td>1</td>
<td>196</td>
<td>10</td>
<td>Fixed</td>
<td>Experienced vs. Inexperienced</td>
<td>Players simultaneously choose a number from 0 to 100 and the winner whose number is closest to p (&lt;1) times the group average</td>
</tr>
<tr>
<td>Ho et al. (1998)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note 1: Continuous strategies of 80 to 200 are discretized to 121 integer strategies
Example 2: $P$-Beauty Contest

- $n$ players
- Every player simultaneously chooses a number from 0 to 100
- Compute the group average
- Define Target Number to be 0.7 times the group average
- The winner is the player whose number is the closest to the Target Number
- The prize to the winner is US$20
$P$-beauty contests: Actual Choices of experienced Subjects

Ho, Camerer, and Weigelt (AER, 1998)
$P$-beauty contests:
Prediction of EWA

Figure 3c: Adaptive EWA (Experienced Subject)
$P$-beauty contests:
Prediction of Self-tuning EWA
Two Open Questions

- A theory to explain why parameters vary across games

- Metrics for judging model performance
  - Statistical Measures
  - Economic value
Economic Value

- Evaluate models based on their value-added rather than statistical fit (Camerer and Ho, 2000)

- Treat models like consultants and seek advice on what opponents will do

- If players were to hire either Mr. Nash or Ms. EWA as consultant and listen to his or her advice at the beginning of each round, would they have made a higher payoff?
Economic Value: Definition and Motivation

- Economic value of a model = how much more players would earn if they use the model to forecast what others will do, and best-respond given that forecast, compared to how much they actually earn.

- A measure of degree of disequilibrium, in dollar terms.
  - If players are in equilibrium, then an equilibrium theory will advise them to make the same choices they would make anyway, and hence will have zero economic value.

  - If players are not in equilibrium, then players are mis-forecasting what others will do. A theory with more accurate beliefs will have positive economic value (and an equilibrium theory can have negative economic value if it misleads players).
## Economic Value: Example

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>8,8</td>
<td>8,6</td>
</tr>
<tr>
<td>B</td>
<td>6,8</td>
<td>12,12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Column's Choice</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td></td>
<td>L</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Row's Choice</th>
<th>T</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Actual Payoff</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 1's Prediction of Prob(R)</th>
<th>0.5</th>
<th>0.75</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Model 1's Recommendation</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>- Improvement in Payoff</td>
<td>4</td>
<td>-2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 2's Prediction of Prob(R)</th>
<th>0.5</th>
<th>0.25</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Model 2's Recommendation</td>
<td>B</td>
<td>T</td>
</tr>
<tr>
<td>- Improvement in Payoff</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>
# Economic Value of Models

## Table 4: Economic Value

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Mixed Strategies</th>
<th>Patent Race</th>
<th>Continental Divide²</th>
<th>Median Action²</th>
<th>Pot Games</th>
<th>p-Beauty Contest²</th>
<th>Price Matching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Payoff</td>
<td>1560</td>
<td>2657</td>
<td>872</td>
<td>432</td>
<td>304</td>
<td>519</td>
<td>192</td>
</tr>
<tr>
<td>Economic Value Achieved as a Percentage of Actual Payoff ¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>%V(Rational Expectation)</td>
<td>17.7%</td>
<td>47.9%</td>
<td>1.9%</td>
<td>0.9%</td>
<td>35.9%</td>
<td>585.4%</td>
<td>19.3%</td>
</tr>
<tr>
<td>%V(Self-tuning EWA)</td>
<td>11.8%</td>
<td>3.8%</td>
<td>1.4%</td>
<td>0.7%</td>
<td>6.9%</td>
<td>53.6%</td>
<td>15.5%</td>
</tr>
<tr>
<td>%V(EWA)</td>
<td>3.9%</td>
<td>5.0%</td>
<td>0.8%</td>
<td>0.5%</td>
<td>4.7%</td>
<td>49.8%</td>
<td>15.2%</td>
</tr>
<tr>
<td>%V(Belief Based)</td>
<td>5.5%</td>
<td>1.4%</td>
<td>-0.7%</td>
<td>-0.5%</td>
<td>3.4%</td>
<td>48.9%</td>
<td>13.9%</td>
</tr>
<tr>
<td>%V(Reinforcement)</td>
<td>11.9%</td>
<td>5.0%</td>
<td>-5.2%</td>
<td>-0.1%</td>
<td>7.3%</td>
<td>-27.6%</td>
<td>9.7%</td>
</tr>
<tr>
<td>%V(QRE)</td>
<td>2.5%</td>
<td>1.3%</td>
<td>-8.6%</td>
<td>-0.5%</td>
<td>8.1%</td>
<td>-60.2%</td>
<td>12.0%</td>
</tr>
</tbody>
</table>

Note 1: We assume that each bionic subject use the respective model to predict other's behavior and best responds with the strategy that yields the highest expected payoff.

Note 2: The expected value of each strategy in these games is computed with 1000 simulated instances for a given round due to high computational burden for actual derivation.
Takeaways

- EWA cube provides a simple but useful framework for studying learning in games.

- EWA model fits and predicts better than reinforcement and belief learning in many classes of games because it allows for a positive (and smaller than actual) simulated effect.

- Self-tuning EWA can approximate EWA reasonably well.

- Economic value is an alternative measure for judging models.
Thank You