Motivation

- Current game theory is a theory about equilibria because alternatives are hard to work with rather than because it is essentially true.
- Consider “blind testing” task (like in wine tasting) with learning models.
- The paper have shown that models cannot be distinguished because parameters are not very well identified. When we re-do the original exercise, it is what we find as well.
- Our contribution is that we found the way out of this problem: compare predictive quality of models, because then they are clearly distinguishable.
Learning models

Action learning models (a player’s strategy set is updated at the end of each period as traditional action learning models require)

Three leading action learning models:

▶ Belief-based Models
  ▶ Fictitious Play - Brown (1951)
  ▶ Cournot Best Response - Cournot (1960)
  ▶ $\gamma$-Weighted Beliefs - (Cheung and Friedman (1997))

▶ Reinforcement-based Models

▶ Hybridized Models
  ▶ Experience Weighted Attraction (EWA) - Camerer & Ho (1999)
  ▶ STEWA (Ho, Camerer, and Chong (2007))
  ▶ Inertia, Sampling And Weighting (I-SAW) - Erev, Ert, and Roth (2010))
COMPARISON OF LEARNING MODELS WITH HUMAN INTERACTION

EWA Model I

Main idea of EWA is to provide a flexible general model that can account for both reinforcements (reaction to own actions) and best responses (reaction to opponent’s actions)

- $\rho$ and $\phi$ - discounting factors
- $\delta$ - relative weight of hypothetical and actual payoffs
- $\lambda$ - sensitivity to attraction
- $N(0)$ - strength of initial experiences
- $A^j_i(0)$ - form of initial experiences
EWA Model II

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

\[ A^j_i(t) = \frac{\phi N(t - 1)A^j_i(t - 1) + [\delta + (1 - \delta)\Pi(s^j_i, s_i(t))]\pi_i(s^j_i, s_{-i}(t))}{\rho N(t - 1) + 1} \]

\[ P^j_i(t + 1) = \frac{\exp(\lambda A^j_i(t))}{\sum_{k=1}^{m_i}\exp(\lambda A^j_i(t))} \]
Actual CB parameters: $\rho = 0, \delta = 1, \phi = 0, N(0) = 1, \lambda = 1$
Actual FP parameters: $\rho = 1, \delta = 1, \phi = 1, N(0) = 0, \lambda = 1$
(Salmon, 2001) redux

Actual RL parameters: $\rho = 0, \delta = 0, \phi = 0.75, N(0) = 1, \lambda = 1$
Actual Mdl parameters: $\rho = 0.5, \delta = 0.5, \phi = 0.5, N(0) = 0.5, \lambda = 1$
Brier score

- $N$ be number of predictions to be made,
- $R$ - number of possible actions that could be taken,
- $P_{ti}$ - predicted probability to take action $i \in R$ in round $t \in N$,
- $F_{ti}$ - indicator of the action that was actually taken.

Then

$$Score = \frac{1}{N} \sum_{t=1}^{N} \sum_{i=1}^{R} (P_{ti} - F_{ti})^2$$

It increases in number of mistakes, thus lower the the score – better the model.
Main result

Average Brier score vs. alternative model for different models.

- EWA_RL: 0.74***
- EWA.mdl: 0.28***
- EWA_FP: 0.44***
- EWA_CB: 0.05 true model

- EWA_RL: 0.59***
- EWA.mdl: 0.55***
- EWA_FP: 0.4*** true model
- EWA_CB: 0.82***

- EWA_RL: 0.51***
- EWA.mdl: 0.29 true model
- EWA_FP: 0.44***
- EWA_CB: 0.49***

- EWA_RL: 0.1 true model
- EWA.mdl: 0.52***
- EWA_FP: 0.31***
- EWA_CB: 0.84***
Main result under a log-loss predictive metric

Kruskal-Wallis test for two group with closest mean brier $\alpha p<0.001$ log_score worse than random better than random

Average log loss

<table>
<thead>
<tr>
<th>Alternative Model</th>
<th>EWA_RL</th>
<th>EWA_mdl</th>
<th>EWA_FP</th>
<th>EWA_CB</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Model</td>
<td>0.74***</td>
<td>0.28**</td>
<td>0.44**</td>
<td>0.05</td>
</tr>
<tr>
<td>EWA_CB</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EWA_FP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EWA_RL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Brier score</td>
<td>0.0 0.5 1.0 1.5 2.0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2x2

Average Brier score

0.0 0.5 1.0 1.5 2.0
Why this result is observed?

Why Brier score works where classic statistics fails?

- Learning is nonstationary: starts with randomization, ends in a stable (usually equilibrium) state
- Learning is reactive - regime changes in response to changes in environment are very quick, predictive quality reward models that are more sensitive to such changes
Conclusion

- Models of learning are useful if and only if they are comparable on a clear scale
- Parameter estimates are not very reliable
- Predictive quality is a useful substitute both because it works better and because it makes more sense from theoretical standpoint