

# Learning, signaling and social preferences in a public good game

Marco Janssen, TK Ahn  
2005

# Représenter l'apprentissage

- Cadre : expérience de choix social de contribution au bien public
- Face aux dilemmes sociaux, on sait que les individus ont des comportements *non optimisateurs* et *hétérogènes*
- **Mais on n'a toujours pas de modèle robuste, en particulier face à des problèmes complexes, comme avec N joueurs**
- Ici : utilisent plusieurs modèles comme social preferences (Charness and Rabin, 2002) et EWA (Camerer and Ho, 99) pour les évaluer...
- Données venues des expériences de Isaac and Walker (1988) et Walker et al. (1994)

# Contexte du travail

- Choix du modèle tel que l'optimum est la non-contribution au bien public mais les participants aux expé contribuent
- Beaucoup de tests divers: de 4 à 40 joueurs, de 10 à 60 pas de temps – risque d'être difficile de pouvoir tomber sur les bons paramètres – complexité des tests plus grande que d'habitude
- On peut choisir le niveau micro ou le macro comme données pertinentes à reproduire (dans un travail antérieur avaient comparé divers apprentissages et trouvaient certains algorithmes plus efficaces pour atteindre des moyennes et d'autres mieux pour les parcours individuels).
- Approche différente des tests d'apprentissage faits par les économistes expérimentaux, parce que l'hétérogénéité est au centre de la recherche

# Le cadre du modèle

- **Jeu de bien commun linéaire**
  - N joueurs, retour marginal par tête (pcmr)  $r$ , le nombre de répétition  $T$ , chaque joueur a  $\omega$
  - Problème de free-riding si  $r < 1$  et  $N.r > 1$
  - $\Pi_i = \alpha.(\omega - x_i + r. \sum x_j)$
- Ici, si on suit l'équilibre, personne ne donne rien au bien commun
- Si on observe les comportements, il y a contribution, et des tendances en fonction des divers paramètres

# Quelques faits stylisés

- La contribution dépend de **la taille du groupe, le pcmr, la durée de l'expérience**
  - Plus la taille augmente, plus il y a contribution
  - Mprc haut implique plus de contribution
  - Durée de l'expérience augmente la contribution
- A contribution donnée il y a une **grande variabilité individuelle**
  - En particulier il y a beaucoup d'extrêmes : 70% soit 0 soit tout.
- Les joueurs changent de contribution à chaque tour, chacun de façon différente. Cette hétérogénéité dépend du temps qu'il reste et des paramètres

# Le modèle d'apprentissage

- En trois parties
  - Choix probabiliste pour la décision
  - Apprentissage pour le changement de comportement des joueurs
  - Utilité sociale utilisée par les agents pour leurs choix

## Choix probabiliste de la contribution x

- $P_i^x = \exp(\varphi_i \cdot A_i^x) / \sum_{\omega} \exp(\varphi_i \cdot A_i^x)$
- Où  $\varphi_i$  est la sensibilité de réponse et  $A_i^x$  est l'attraction de x pour agent i
- si  $\varphi_i$  augmente, la discrimination entre valeurs augmente
- Si  $A_i^x$  augmente, la probabilité que x soit choisit augmente

# Apprentissage

- C'est en faisant évoluer les  $A_i^x$  qu'on apprend
- EWA : experience weighted attraction  $A_i^x(t)$
- $H(t)$  est l'influence du passé
  - $H(t) = H(t-1) \cdot \lambda_i \cdot (1 - \kappa_i) + 1$
- $\lambda_i$  est l'oubli,  $\kappa_i$  est le taux d'augmentation de l'attraction -> impact de l'expérience

$$A_i^x(t) = (\lambda_i \cdot H(t-1) \cdot A_i^x(t-1) + [(\delta_i + (1 - \delta_i) \cdot I(x_i, x_i(t)))] \cdot u_i(x_i, x_{-i}(t))) / H(t)$$

Où  $I$  vaut 0 pour inégalité et 1 pour égalité (si c'est la contribution de  $i$  au temps  $t$  ou non)

# Préférence sociale

- C'est un facteur de rationalité portant sur une comparaison des gains aux autres aux siens
- $U_i = \rho \cdot \text{moy}(\pi_{-i}) + (1 - \rho) \cdot \pi_i$  si  $\pi_i \geq \text{moy}(\pi_{-i})$
- $U_i = \chi \cdot \text{moy}(\pi_{-i}) + (1 - \chi) \cdot \pi_i$  si  $\pi_i < \text{moy}(\pi_{-i})$
- Et  $\chi < \rho < 1$  – cela signifie qu'on est plus sensible au gain des autres quand il est supérieur au sien
  - Si  $\chi < \rho < 0$  alors l'agent est très compétitif et aime avoir plus que les autres
  - Si  $\chi < 0 < \rho < 1$  – aversion à l'inégalité
  - Si  $0 < \chi < \rho < 1$  – social welfare : s'intéresse à ce que les autres soient aussi bien que lui
  - Si  $0 = \chi = \rho$  alors ne s'intéresse qu'à son gain et se désintéresse des autres

# Signaling

- Hypothèse portant sur le fait que les joueurs jouent haut pour pousser les autres à les suivre.
- Dans ces cas-là ça deviendrait inutile vers la fin de jeu, et on aurait diminution des contributions (ce qui se vérifie dans les expé)
- On ajoute à l'équation déjà obtenue pour  $A_i^x$
- $\text{Signal} = x_i \cdot r \cdot \theta_i \cdot (T-t/T) \eta_i$
- Où  $\theta_i$  est l'espoir d'avoir un impact,  $(T-t/T)$  la proportion de temps restant,  $\eta_i$  est le comportement final de  $i$

# Rationalité et paramètres

$$A_i^x(t) = \frac{\lambda_i H(t-1) A_i^x(t-1) + [\delta_i + (1-\delta_i) I(x_i, x_{-i}(t))] (u_i(x_i, x_{-i}(t)) + x_i \cdot r \cdot \theta \cdot \left(\frac{T-t}{T}\right)^\eta)}{H(t)}$$

$$P_i^x = \frac{e^{\varphi_i A_i^x}}{\sum_{j=0}^{\omega} e^{\varphi_j A_j^x}}$$

$$U_i = \rho \cdot \bar{\pi}_{-i} + (1-\rho) \cdot \pi_i, \quad \text{if } \pi_i \geq \bar{\pi}_{-i} \text{ and}$$

$$U_i = \chi \cdot \bar{\pi}_{-i} + (1-\chi) \cdot \pi_i, \quad \text{if } \pi_i < \bar{\pi}_{-i} \text{ and}$$

$$H(t) = H(t-1) \lambda_i (1 - \kappa_i) + 1$$

Parameter	Interpretation
$\rho$	Weight to others' payoff when $\pi_i > \bar{\pi}_{-i}$
$\chi$	Weight to others' payoff when $\pi_i < \bar{\pi}_{-i}$
$\varphi$	Response sensitivity
$\lambda$	Forgetting rate
$\delta$	Weight to forgone payoffs
$\kappa$	Rate of attraction growth
$\theta$	Signal weight
$\eta$	Weight of future rounds in signaling

# Tests du modèle

- Font trois tests
  - Agent représentatif
  - Découpage par catégories
  - Apprentissage individuel
- Face à la grande complexité : prennent en compte le « fit »  $L$  et le nombre de paramètres, voire même le nombre d'agents
  - AIC et BIC doivent être les plus bas possible
  - $AIC = -\ln L + 2k$
  - $BIC = -\ln L + k \ln N$
  - $k$  : En tout il y a 8 paramètres.

# Agent représentatif

- Plusieurs sous-modèles testés
  - SP : seulement social preference
  - SP + L
  - SP +L + S
  - L : Seulement apprentissage maximisateur
- Résultats
  - C'est SP+L+S le meilleur fit
  - $\lambda = 0.85$  et  $\delta = 0.55$  à  $0.72$  : sont des beliefs learner plutôt qu'optimisateurs
  - Dans les expériences de 10, l'effet du signaling est très fort au début mais dure peu // Dans les expériences de 40-60 l'effet est plus faible mais dure longtemps

# Catégories d'agents

- Ici la rationalité choisie est  $SP + L + S$
- Augmente le nombre jusqu'à ce que ça rende moins bon l'estimation :  $i$ 
  - 8 pour les 10 tours
  - 2 pour les 40-60 tours

# Individualisation

- Tirent une distribution uniforme de chaque valeur
- Trouvent le nombre d'agents dans chacun des 16 types d'apprentissage, qui sont définis par les préférences sociales et par le type d'apprentissage.
- La plupart sont des beliefs learners et s'intéressent donc aux gains qu'ils auraient fait en jouant autre chose.
- La majorité sont rétifs à l'inégalité
- Moins de 10% sont des optimisateurs

# Problématique de la correspondance aux données

- Marco Janssen et Ahn (2003 – 2005)
  - trois formes classiques comparées / combinées - EWA + social preferences + signalling -
- Duffy est un des premiers à immerger des agents artificiels au milieu d'humains pour pousser les apprentissages dans le sens de la rationalité ou évaluer ses propres algorithmes.
- Grossklags: La présence d'agents artificiels transforme les comportements des joueurs humains de façon différentes selon la conscience de leur présence.