Nov. 29, 2015
実験社会科学コンファレンス
Univ. Tokyo

繰り返しゲームの理論と実験：モニタリングが不完全なケース

松島斉
（東京大学大学院経済学研究科教授）

Monitoring Accuracy and Retaliation in Infinitely Repeated Games with Imperfect Private Monitoring: Theory and Experiments

Yutaka Kayaba（Hitotsubashi University）： データ解析担当
Hitoshi Matsushima（University of Tokyo）： 理論、実験デザイン担当
Tomohisa Toyama（Kogakuin University）： プログラミング、実験デザインその他担当

Older Version: Matsushima and Toyama (11)

＊我々は「助け合う」「協力しあう」関係である。ただし、必ずしもお互い寛容というわけではない。

まずは繰り返しゲームの簡単な入門と経済学者の中心的な関心事について

「助け合う」「協力しあう」：どのように動機付けられるか？

利他心、公共心、互恵性：　**Dictator Game, Trust Game**
Ultimatum Game, etc.

コーディネーション、契約、社会システム設計：　**Mechanism Design**

＊長期的関係（**Stick and Carrot**）：　繰り返しゲーム

長期的関係による「暗黙の協調」

協力行動へ調整：　模倣（**Imitation**）
相手の裏をかく
裏の裏をかく
……

＊安定した協力関係：　調整の余地のない状態、調整の行き着く果て
均衡：ナッシュ均衡

模倣（１）：付き合いの作法

Tit-for-Tat (TFT) 戦略

今日相手が私を助ける（協力する　**Cooperation, C**）　⇒ 明日私は相手を助ける（**C**）
今日相手が私を助けない（協力しない　**Defection, D**）　⇒ 明日私は相手を助けない（**D**）

⇒ 相手とほぼ同じ付き合いの程度（**C** の頻度）を達成できる
誰とでもそれなりにやっていける
Axelrod (84)

模倣（２）：郷に入れば郷に従え

C-Society vs D-Society

C-Society： 「助け合う（**C**）」が社会規範（ノルム）
D-Society： 「自己チュー（**D**）」が社会規範

C-Society の方が **D-society** よりパフォーマンスよい ⇒ **C-Society** へ移行（郷に従え）

バブル期の日本経済論： 「**J-Society**（終身雇用、年功序列）」**vs** 「**A-Society**」

相手の裏をかく

C-Society に入って **D-Play** する

裏の裏をかく

D-play に対抗する措置、**C-**ノルムの崩壊（**TFT？**）

裏の裏の裏をかく、裏の裏の裏の裏をかく**……**

出口のない調整の連鎖へ
複雑な調整過程

∴　経済学では、調整過程についての標準モデルは確立していない

企業戦略転換、新規参入退出、クラッシュからクライシス：　スピーディーな側面？
不必要悪な慣行、新しい社会システムの導入、クライシスへの収束：　スローな側面？

経済学者の中心的な関心は

＊安定した協力関係
調整の余地のない状態：ナッシュ均衡（ＮＥ）

相手の行動（パターン）を予想して、最適行動（パターン）を選択
ナッシュ均衡：　「相手に対する予想　　＝ 相手の最適行動（パターン）」

例：価格カルテル

ライバル企業１，２：　「価格競争をやめましょう」
「独占（高い）価格をつけあいましょう」
この約束は守られるか？
「ちょっと値引きすれば、お客さんをいっぱいとれる。」
「ならばカルテル破りを黙ってしよう。」

⇒　　「**Stick and Carrot**（飴とムチ）」でカルテルを守らせよう
繰り返しゲームの均衡分析！

Example：囚人のジレンマ（Symmetry + Additive Separability）を繰り返す

	c (cooperative)	d (defective)
c (cooperative)	X　　X	X-Z　　X+Y
d (defective)	X+Y　　X-Z	X+Y-Z　　X+Y-Z

Y:　「助ける（協力）」にかかる費用
Z:　「助けられる」メリット
$Z \gg Y$：メリット大きい

(c,c)が二人にとって望ましい
しかし、一回限りのゲームでは、(d,d)が唯一のナッシュ均衡（利己的動機にもとづく）

繰り返しゲームではどうか？

「Stick and Carrot」型戦略

c をとり続ける　⇒ 相手も c をとり続ける（協力）
今日 d をとる　　⇒ 明日、あるいは明日以降、相手は d をとって報復

⇒ 報復をおそれ、d をとらない
⇒ 協力関係が安定的に維持
　強制せずとも自発的に協力：暗黙の協調

ここから徐々に本題へ…

「暗黙の協調」の三つの必須項目

１： プレーヤーは「無限に」繰り返されているというパースペクティブをもつこと
cf. Backward Induction（有限回のパースペクティブ）
最後は協力しない。ならば、最後から二番目でも協力しない…　⇒　結局最初から協力しない

２： 割引ファクターが十分な値（１に近い）であること
明日相手に **d** とられることの損失δZ（あるいはそれ以降の報復）が
今日自身が **d** とることの便益Yより大きい

３： 相手の行動が事後的に観察できること
現実：完全には観察できない（ノイズのかかったシグナル）
完全モニタリング **vs** 不完全モニタリング
Secret Price Cutting：　だまって値下げする
⇒ 相手の売り上げが下がる：カルテル破り？不景気？
High discount factor implies Frequent Access to Information Channel, but still imperfectly monitors.

「Stick and Carrot」型戦略：類型

Grim-Trigger, Grim-2, Grim-3
TFT
g-TFT
その他： **2TFT, TF2T, g-2TFT, g-TF2T, etc.**

Stick and Carrot 以外： **All-D, All-C**

List of Strategies for our Experiments

TFT	Tit for Tat
g-TFT-25f/50f/75f	Generous-Tit for Tat, retaliation with 75%/50%/25% (forgiving a bad signal with 25%/50%/75%, no retaliation for good signal)
TF2T	Tit for Two Tat (retaliate if signal B in all of the last two rounds)
g-TF2T-25f/50f/75f	Generous-Tit for Two Tat, retaliation with 75%/50%/25%
TF3T	Tit for Three Tat (retaliate if signal B in all of the last three rounds)
2TFT	Two Tit for Tat (retaliate twice convectively if signal B occur)
g-2TFT-25f/50f/75f	Generous-Two Tit for Tat, retaliation with 75%/50%/25% in the second punishing phase
2TF2T	Two Tit for Two Tat (retaliate twice convectively if signal B in all of the last two rounds)
All-C	Always choose c
All-D	Always choose d (the only non-cooperative strategy in the set)
Grim	Choose c until either player chose d (signal B), keep choosing d after that
Grim-2	Choose c until the case happens twice in a row, in which either player chose d (signal B), keep choosing d after that
Grim-3	Choose c until the case happens three times in a row, in which either player chose d (signal B), keep choosing d after that

例：トリガー戦略

完全モニタリング（**Perfect Monitoring**）：

最初に **c** をとる
$(\boldsymbol{c},\boldsymbol{c})$をとりつづける限り **c** をとり続ける
相手あるいは自身が$\boldsymbol{d}$をとると、以降ずっと$\boldsymbol{d}$をとり続ける

$\Rightarrow(\boldsymbol{c},\boldsymbol{c})$の繰り返しから$(\boldsymbol{d},\boldsymbol{d})$の繰り返しへ転落

今日$\boldsymbol{d}$とる便益：　$\boldsymbol{Y}$

明日以降の損失：　$\delta(Z-Y)+\delta^2(Z-Y)+\delta^3(Z-Y)\cdots=\dfrac{\delta}{1-\delta}(Z-Y)$

$Y<\dfrac{\delta}{1-\delta}(Z-Y)$、つまり、$\delta>{}^{Y}\!/_{Z}$（十分高い割引ファクター）なら
トリガー戦略ナッシュ均衡が成立！

不完全モニタリング（Imperfect Monitoring）

モニタリング精度（Accuracy） $p\left(>\frac{1}{2}\right)$

	Signal G (Good)	Signal B (Bad)
Action (c)	p	$1-p$
Action (d)	$1-p$	p

相手に自分の **action** 伝わらない。しかしシグナル（**G or B**）は伝わる

公的モニタリング（Public Monitoring）： 自分についてのシグナルと
相手についてのシグナルを共通に観察

私的モニタリング（Private Monitoring）： 自分についてのシグナルは見えない
相手は自分をどうみてるのかわからない

(G,B)＝（プレーヤー１についてのシグナル、プレーヤー２についてのシグナル）

Public Monitoring（不完全、公的）

トリガー戦略

最初に c をとる。(G,G) が観察される限り c をとり続ける

(G,B)、(B,G)、(B,B) のどれかが観察されると、(c,c) から (d,d) の繰り返しへ転落

割引ファクター δ およびモニタリング精度 p がともに 1 に近ければ、

トリガー戦略ナッシュ均衡成立

ただし、いずれは (d,d) の繰り返しへ転落

C をとっても **B** が観察される

⇒　**Welfare Loss** 発生

有限トリガー戦略

弱い報復によって **Welfare Loss** を **Save**

(G,G) 以外が観察されると、以降 T 期間 (d,d) を繰り返す

$T+1$ 期後は **(c,c)** から再スタート

⇒　　低い δ、p でもナッシュ均衡

今日の中心テーマ

Private Monitoring（不完全、私的）

「今は協力の時？報復の時？」わからん

相手の観察結果（自分の選択についてのシグナル、自分がどう見られているか）
が観察できない

∴ トリガー戦略はもはや使えない！

TFT あるいはその変形、**g-TFT**、を
「**Stick and Carrot**」の基本に据えよう

（再び）TFT

「模倣の作法」としてではなく
「目には目を歯には歯を」（**Stick and Carrot**）の作法
相手から協力的態度を引き出すための作法と考えよ！

今日 $\boldsymbol{d}$ をとることの便益：　$\boldsymbol{Y}$
明日相手から $\boldsymbol{d}$ の報復を受けることの期待損失：　$\boldsymbol{\delta(p-(1-p))Z}$

昨日相手が **Bad** の場合は $\boldsymbol{d}$ をとらねばならない：　報復するインセンティブは？
昨日相手が **Good** の場合は $\boldsymbol{c}$ をとらねばならない：　協力するインセンティブは？

今日は報復の時？協力の時？わからない：　無差別条件が不可欠

$$\boldsymbol{Y=\delta(p-(1-p))Z}$$

∴　特殊な $\boldsymbol{(\delta, p)}$ のケース以外には **TFT** ナッシュ均衡（利己的動機にもとづく）不成立

⇒ より一般的な「**g-TFT**」へ！

Generous-Tit-for-Tat（g-TFT）

BAD を観察 ⇒ 確率 $r(B;p)$ で **c** を選択する
GOOD を観察 ⇒ 確率 $r(G;p)\,(>r(B;p))$ で **c** を選択する

ex. TFT: $r(B;p)=0,\ r(G;p)=1$

Retaliation Intensity (RI, 報復の強度): $r(G;p)-r(B;p)$

Stick and Carrot を特徴付ける最重要コンセプト

TFT: $RI=1$

RI が高い（低い） ⇒ 相手が c とる（協力）インセンティブ高める（下げる）
d をとる（報復）インセンティブ下げる（高める）
Welfare Loss アップ（ダウン）

g-TFT Nash Equilibrium $(r(G;p), r(B;p))$

（利己的動機にもとづく）

RI が

$$w(p) \equiv r(G;p) - r(B;p) = \frac{Y}{\delta(2p-1)Z}$$

であれば、その時にのみ

g-TFT NE 成立！

$w(p)$ を RI の理論値としよう

確率的選択によって報復の強さ（**RI**）が微調整される

Belief-Free Approach: 協力報復ともにインセンティブ（無差別）

Ely and Valimaki (02)

Obara (01)

進化生物学： **Nowak & Sigmund (92)**

ランダムな選択を実際におこなう？： **No empirical supports ever.**

重要な性質
RI 理論値$\boldsymbol{w(p)}$は精度$\boldsymbol{p}$の減少関数でなければならない

精度 $\boldsymbol{p}$ が高いと

相手が$\boldsymbol{d}$をとったかどうかをより正確に見定めることができる

⇒　弱い報復強度でも十分効果的

RI（$\boldsymbol{w(p)}$）が小さくても大丈夫

（むしろ小さくないと **C** と **D** 両方のインセンティブ付けられない）

Welfare Loss 減らせる！

合理的な経済主体なら不完全モニタリングの制約を
さらに上手に解決できるはず

１： レビュー戦略（あるいは **Lenience Strategy**）

情報蓄積をしてから、報復するか否かを決める
Radner (85), Matsushima (04), Sugaya (14)

２： **Long (Finite) Retaliation**

１期間のみではなく一定期間報復
RI をより低い値に保つことができる

g-TFT NE（**No Lenience, Just One-Shot Retaliation**）では **RI** が高くなってしまう
理論値 $\boldsymbol{w(p)}$ は全均衡のなかでもとくに高い **RI**

合理的プレーヤーは協調達成の工夫を実行するインセンティブをもつか
「均衡」として成立するのか

例：　レビューの途中で、**BAD** がたくさん続いてしまった
レビューの途中で $\boldsymbol{d}$ に切り替えるかも
相手はそれを読み込むから $\boldsymbol{c}$ をとるインセンティブうしなう？

例：　有限期間の報復途中で、あえて $\boldsymbol{c}$ をとって報復期間をスキップ？

Folk Theorems：日本（東大）学派

Perfect Monitoring: **Fudenberg and Maskin (86)**
Public Monitoring: **Matsushima (88), Fudenberg, Levine, and Maskin (94)**
Private Monitoring: **Matsushima (91a, 91b), Sekiguchi (97), Ely and Välimäki (02), Obara (01)**
Kandori and Matsushima (98), Matsushima (04)
Sugaya (14)
その他日本人経済学者：**Aoyagi, Satoru Takahashi, Yuichi Yamamoto, Awaya, etc**

Folk Theorem とは？

割引ファクターδが十分に１に近いならば、不完全モニタリングであろうとも、私的モニタリングであろうとも、ほぼどんなゲームにおいても、効率的配分を含む広範囲の利得配分が近似的にナッシュ均衡によって達成できる。

非常に複雑な均衡分析が必要

Sugaya (14)：　今のところ「究極のフォーク定理」

Belief-Free Approach
Review (Lenient) Strategies（with finite punishment phases）
even (much) more complicated

実験経済学

被験者は、実際に協力をうながす戦略的工夫をとるか？
どの程度？どのように？
理論値からどの程度乖離している？
協力行動をとっている？どんな行動パターン？

さらには…

実験結果に即した均衡理論をいかにして再構築するか？
利己的動機以外の要因（**Naivete, Reciprocity**）を織り込んだ均衡理論は？

繰り返しゲームの実験は古くからある
しかし経済学的に意味のある研究は２１世紀に入ってから！

ターニングポイントは **Dal Bo (05)**
「無限回のパースペクティブ（**Shadow of Future**）」

Infinitely repeated Games with Perfect Monitoring:
Dal Bo (05), Dal Bo and Frechette (11)
Public Monitoring:
Aoyagi and Frechette (09), Fudenberg, Rand, and Dreber (12)
Private Monitoring:
The previous draft (Matsushima and Toyama (11)) is the first paper.
Subsequently, Aoyagi, Bhaskar, and Frechette (14, in preparation)

Dal Bo and Frechette (15) is an excellent general survey.

実験デザイン

High Discount Factor: $\delta = 96.7\%$

Constant Random Termination

Continuation probability 29/30

High accuracy treatment: $p = 90\%$

Low accuracy treatment: $p = 60\%$

Length of each experiment

	Number of Subjects	Order of Treatments (games in each interaction)
October 5, 2006 (10:30 ~ 12:30)	28	0.6 (24, 40, 25), 0.9 (28, 33, 14)
October 5, 2006 (14:30 ~ 16:30)	24	0.6 (20, 23, 37), 0.9 (34, 34, 19)
October 6, 2006 (10:30 ~ 12:30)	28	0.9 (38, 21, 25), 0.6 (25, 28, 29)
October 6, 2006 (14:30 ~ 16:30)	28	0.9 (25, 35, 23), 0.6 (36, 30, 21)

Total 108 subjects for both treatment

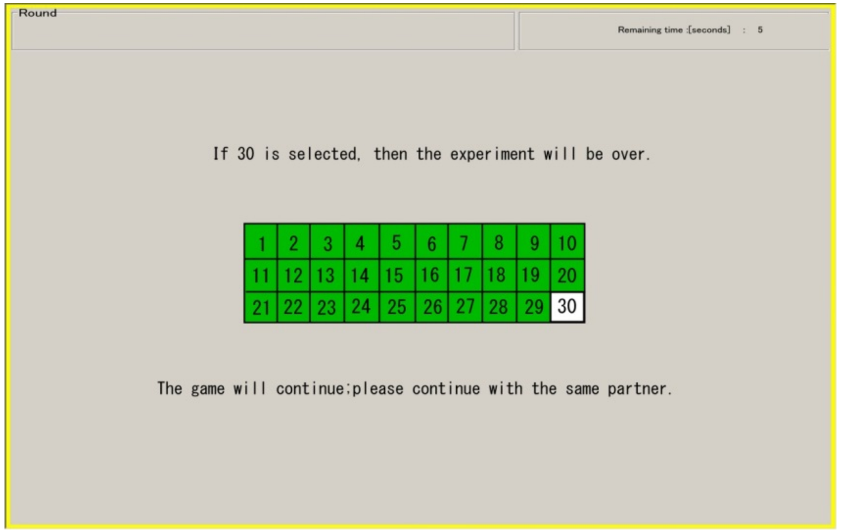
Three repeated games in each treatment

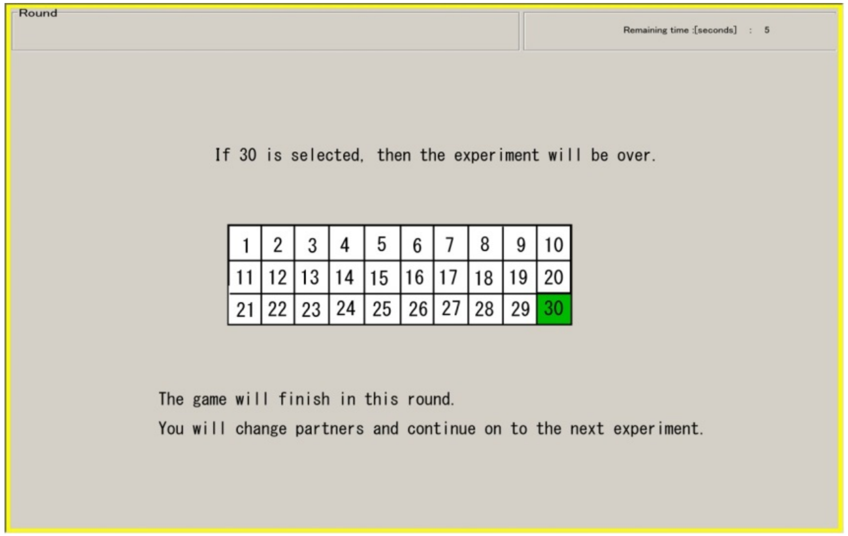
Treatment is counter-balanced

利己的動機付け: **0.6 yen per point + 1,500 yen participation fee**

無限繰り返しのパースペクティブ

各期末に（フェイク）ルーレットを見せて **Constant Random Termination** をアピール





その結果

繰り返しゲームの前半と後半で行動パターンに大きな変化なし

繰り返しゲームを「繰り返し」ても大きな変化なし

囚人のジレンマ：数値の特定

$$\boldsymbol{(X, Y, Z) = (60, 10, 55)}$$

	c	d
c	60　60	5　70
d	70　5	15　15

$$\boldsymbol{Z \gg Y}$$

メリットが費用を大きく上回る
Low Accuracy でも暗黙の協調ＯＫ

実験で調べたいこと

Aggregate Behaviors

精度が高い（$p=0.9$）方がＣをとる頻度が高い？
RI（頻度差としての報復度）は $p=0.9$ の方が低い？
RI は理論値 $w(p)$ と同じかそれ以下である？
Lenience する傾向ある？

Individual Strategies

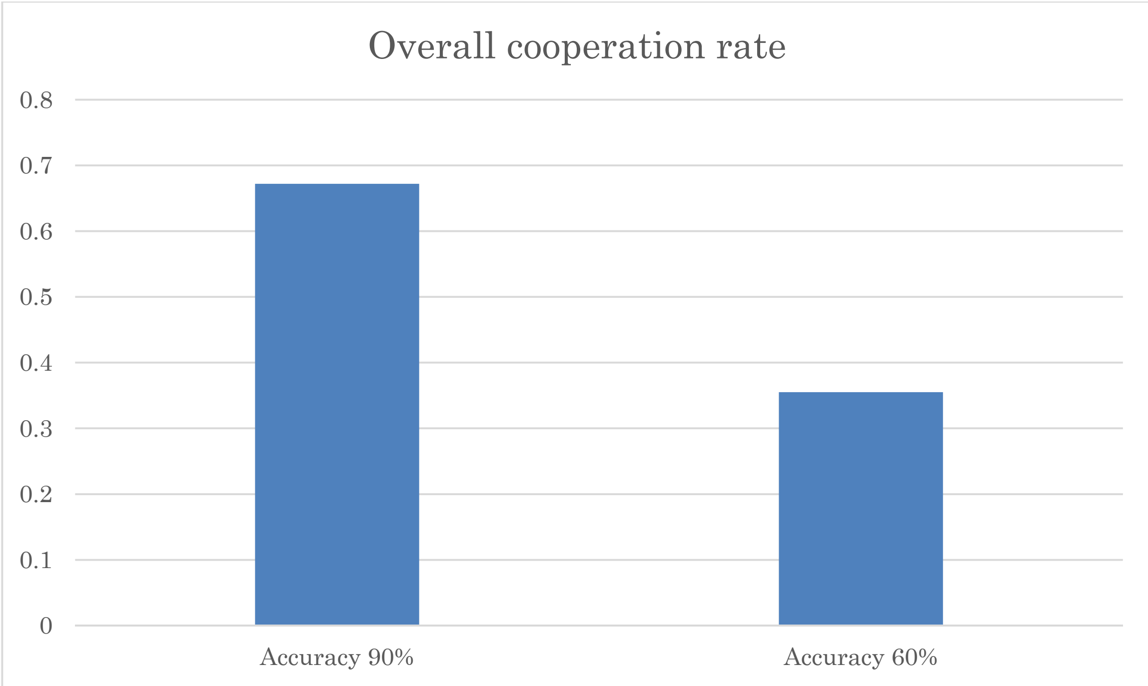
Maximal Likelihood Estimation
リストアップした戦略のうち各被験者の行動はどれに近似できる？
Cooperative Strategy はどの程度ある？
g-TFT をプレイしている？
Lenience は？

Aggregate Behaviors

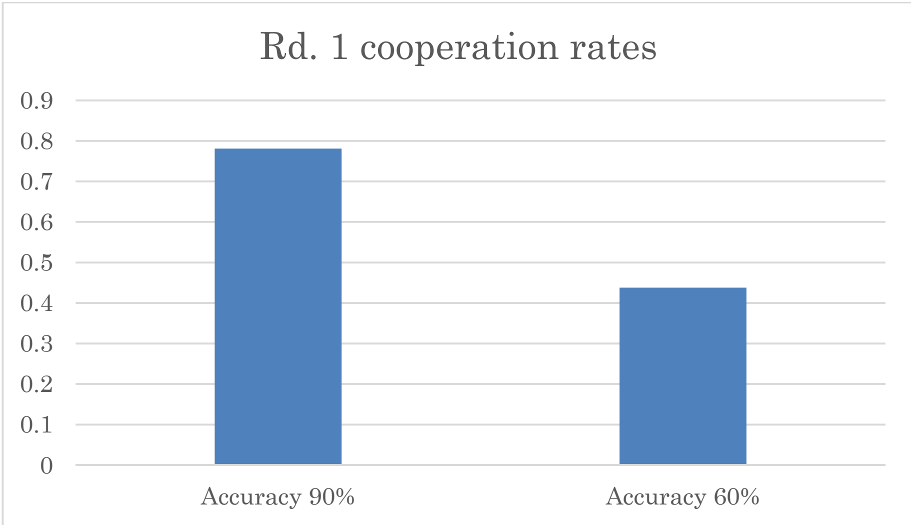
結果 1

High Accuracy（$\boldsymbol{p = 0.9}$）の方が **Low Accuracy**（$\boldsymbol{p = 0.6}$）より協力（**C**）頻度が高い

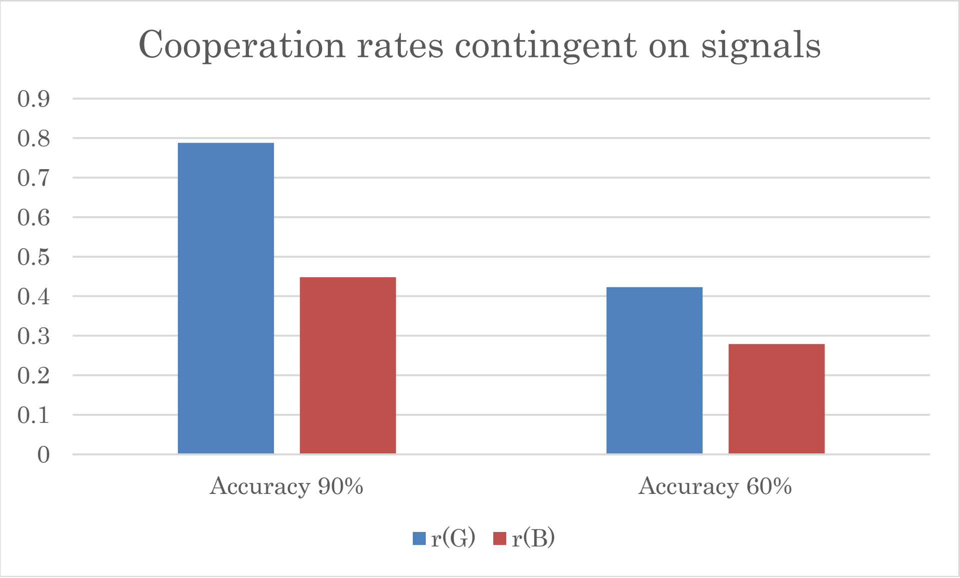
Overall Cooperation Rates



Rd. 1 Cooperation Rates



Cooperation Rates Contingent on Signals
(Subject-Wise Means)



Aggregate Behaviors

結果２

High Accuracy（ $p = 0.9$ ）の方が **Low Accuracy**（ $p = 0.6$ ）より
RI（**Retaliation Intensity,** 報復強度）が高い。
つまり理論値と真逆。

High Accuracy（ $p = 0.9$ ）における **RI** は
理論値 $w(0.9) = 0.235$ より高い
（強すぎる報復態度）

Low Accuracy（ $p = 0.6$ ）における **RI** は
理論値 $w(0.6) = 0.94$ より（かなり）低い
にもかかわらず **Lenience** していない
（弱すぎる報復態度）

RI, Retaliation Intensity (1)

（ここでは$\boldsymbol{a}=\boldsymbol{G}$、$\boldsymbol{b}=\boldsymbol{B}$と表記している）

	Mean	S.E.	p-value
$r(a;0.9)-r(b;0.9)$	**0.508**	**0.028**	**< 0.001***
Subject-wise means	**0.352**	**0.028**	**< 0.001***
$r(a;0.6)-r(b;0.6)$	**0.165**	**0.025**	**< 0.001***
Subject-wise means	**0.144**	**0.023**	**< 0.001***
$(r(a;0.9)-r(b;0.9))-(r(a;0.6)-r(b;0.6))$	**0.344**	**0.031**	**< 0.001**
Subject-wise means	**0.208**	**0.029**	**< 0.001**

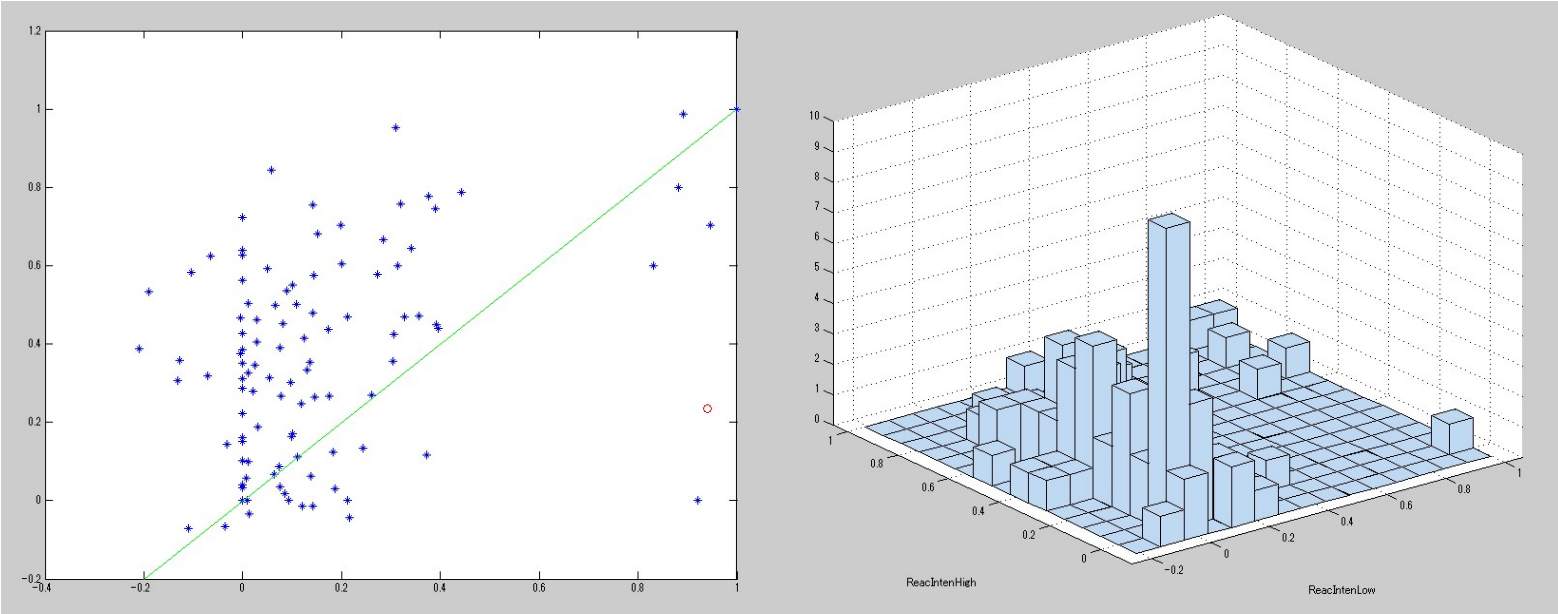
***Hypothesis tests on theoretically predicted values of cooperative equilibrium**

$\boldsymbol{p=0.9}$: $\boldsymbol{w(p)=0.235}$

$\boldsymbol{p=0.6}$: $\boldsymbol{w(p)=0.94}$

$*\,w(p)$ is decreasing in p.

RI, Retaliation Intensity (2)



X axis: RI in the low accuracy treatment
Y axis: RI in the high accuracy treatment
Red plot: Theoretically implied values of RI $(w(0.6), w(0.9)) = (0.94, 0.235)$

Lenience

	p = 0.9	P = 0.6
Gc	0.887*** (0.019)	0.766*** (0.038)
Gd	0.502*** (0.040)	0.255*** (0.035)
Bc	0.506*** (0.038)	0.582*** (0.049)
Bd	0.106*** (0.021)	0.110*** (0.042)
Gc after G	0.062*** (0.022)	0.037 (0.031)
Gd after G	0.017 (0.054)	-0.055 (0.034)
Bc after G	0.125*** (0.041)	-0.045 (0.046)
Bd after G	0.132*** (0.037)	0.022 (0.016)
Obs.	8,216	8,496
R2	0.816	0.531
Adjusted R2	0.816	0.531

Probabilistic linear regression on choices (Breitmoser (2013))

- **No intercept for avoiding singularity problem**
- **Individual heterogeneity controlled with random effect**

Effects of signals two period ago are exclusive in the high accuracy treatment

- **However, the coefficients are marginal**

Individual Strategies

Strategy Frequency Estimation Method (SFEM)

Existing studies empirically demonstrate that the strategies people take are substantially heterogeneous. The study do not focus on a specific strategy (generous-TFT). Rather, we estimate the distribution of various strategies. Rely on Strategy Frequency Estimation Method (SFEM): Dal Bo & Frechette (2011)

- **Maximum Likelihood Estimation on finite mixture model of pre-specified strategies**
  - ✓ **Estimate distribution of strategies**
  - ✓ **Assumption: each subject uses one strategy in the pre-specified strategies set**
- **Workhorse model in the field**
  - ✓ **The validity is verified in Fudenberg, Rand & Dreber (2012), including the case for stochastic strategies (g-TFT) – however, no detailed empirical discussion so far**

Maximum Likelihood Estimation

The choice probability (likelihood) of subject i, employing strategy s, at round r of supergame k, given the history up to the round is defined as $P_{ikr}(s) = \frac{1}{1+\exp(-1/\gamma)}$ **if the observed choice is matched with the predicted choice of strategy s. Otherwise the choice is classified with implementation error as follows.** $P_{ikr}(s) = \frac{1}{1+\exp(1/\gamma)}$ **where** γ **controls the degree of the error. The overall likelihood of subject i, employing strategy s is** $P_i(s) = \prod_k \prod_r P_{ikr}(s)$ **. Now in the SFEM framework, the likelihood of subject i over all strategies is a finite mixture of the above likelihood. The log likelihood of the whole subjects is** $LH = \sum_i \ln \prod_s P(s) P_i(s)$ **where** $P(s)$ **is the frequency of the occurrence of strategy s. Maximize the Likelihood w.r.t.** $P(s)$ **and** γ **. In the case that the strategy s is stochastic, substitute choice probabilities with the following mixture of choice probabilities;** $P_{ikr}(s) = s_{ikr}\left(\frac{1}{1+\exp(-1/\gamma)}\right) + (1-s_{ikr})\left(\frac{1}{1+\exp(1/\gamma)}\right)$ **if the observed choice c,** $P_{ikr}(s) = (1-s_{ikr})\left(\frac{1}{1+\exp(-1/\gamma)}\right) + s_{ikr}\left(\frac{1}{1+\exp(1/\gamma)}\right)$ **if the observed choice d. Here** s_{ikr} **is the probability of choosing c in the stochastic strategy s, given the history up to the round. Observe that the above formulations of the choice probabilities are reduced to the previous definitions when** s_{ikr} **takes either zero or one.**

19 Strategies Set in our estimation

	Description
TFT	Tit for Tat
g-TFT-25f/50f/75f	Generous-Tit for Tat, retaliation with 75%/50%/25% (forgiving a bad signal with 25%/50%/75%, no retaliation for good signal)
TF2T	Tit for Two Tat (retaliate if signal B in all of the last two rounds)
g-TF2T-25f/50f/75f	Generous-Tit for Two Tat, retaliation with 75%/50%/25%
TF3T	Tit for Three Tat (retaliate if signal B in all of the last three rounds)
2TFT	Two Tit for Tat (retaliate twice convectively if signal B occur)
g-2TFT-25f/50f/75f	Generous-Two Tit for Tat, retaliation with 75%/50%/25% in the second punishing phase
2TF2T	Two Tit for Two Tat (retaliate twice convectively if signal B in all of the last two rounds)
All-C	Always choose c
All-D	Always choose d (<u>the only non-cooperative strategy in the set)</u>
Grim	Choose c until either player chose d (signal B), keep choosing d after that
Grim-2	Choose c until the case happens twice in a row, in which either player chose d (signal B), keep choosing d after that
Grim-3	Choose c until the case happens three times in a row, in which either player chose d (signal B), keep choosing d after that

どれがあてはまりいい？

Estimation Results Overview
(結構バラバラ、しかし……)

	p = 0.9	p = 0.6	(ctd.)	p = 0.9	p = 0.6
TFT	0.026* (0.016)	0.058*** (0.019)	2TF2T	0.041* (0.041)	0.027 (0.020)
g-TFT-25f	0.117** (0.053)	0.056* (0.033)	All-C	0.057* (0.003)	0.039** (0.017)
g-TFT-50f	0.187*** (0.056)	0.062 (0.040)	All-D	0.081*** (0.024)	0.425*** (0.056)
g-TFT-75f	0 (0.012)	0.042 (0.032)	Grim	0.012 (0.012)	0.030 (0.024)
TF2T	0.014 (0.022)	0 (0.004)	Grim-2	0.029 (0.017)	0.040 (0.024)
g-TF2T-25f	0.074 (0.059)	0.013 (0.011)	Grim-3	0.064** (0.027)	0.111*** (0.036)
g-TF2T-50f	0.055 (0.047)	0.011 (0.023)			
g-TF2T-75f	0 (0.013)	0 (0.006)			
TF3T	0.113** (0.047)	0.012 (0.013)			
2TFT	0.029* (0.017)	0 (0.009)			
g-2TFT-25f	0.003 (0.017)	0.051 (0.034)			
g-2TFT-50f	0.050 (0.034)	0.021 (0.032)			
g-2TFT-75f	0.046 (0.048)	0 (0)	Gamma	0.478*** (0.022)	0.630*** (0.043)

Individual Strategies

結果１

High Accuracy（ $\boldsymbol{p = 0.9}$）、**Low Accuracy**（ $\boldsymbol{p = 0.6}$）ともに
利己的動機に基づくナッシュ均衡
（**High accuracy：g-TFT-75f**、**Low accuracy：TFT**）
をプレイしている人はいない

High Accuracy（ $\boldsymbol{p = 0.9}$）では **g-TFT-50f**
Low Accuracy（ $\boldsymbol{p = 0.6}$）では **All-D**
がもっとも高いシェア

High Accuracy（ $\boldsymbol{p = 0.9}$）では **92%**、**Low Accuracy**（ $\boldsymbol{p = 0.6}$）**57%**が
Cooperative Strategies（**not All-D**）
をプレイ

Cooperative (other than All-D) v.s. Non-cooperative (All-D)

	p = 0.9	p = 0.6	(ctd.)	p = 0.9	p = 0.6
TFT	0.026*	0.058***	**2TF2T**	0.041*	0.027
	(0.016)	(0.019)		(0.041)	(0.020)
g-TFT-25f	0.117**	0.056*	**All-C**	0.057*	0.039**
	(0.053)	(0.033)		(0.003)	(0.017)
g-TFT-50f	0.187***	0.062	**All-D**	0.081***	0.425***
	(0.056)	(0.040)		(0.024)	(0.056)
g-TFT-75f	0	0.042	**Grim**	0.012	0.030
	(0.012)	(0.032)		(0.012)	(0.024)
TF2T	0.014	0	**Grim-2**	0.029	0.040
	(0.022)	(0.004)		(0.017)	(0.024)
g-TF2T-25f	0.074	0.013	**Grim-3**	0.064**	0.111***
	(0.059)	(0.011)		(0.027)	(0.036)
g-TF2T-50f	0.055	0.011			
	(0.047)	(0.023)			
g-TF2T-75f	0	0			
	(0.013)	(0.006)			
TF3T	0.113**	0.012			
	(0.047)	(0.013)			
2TFT	0.029*	0			
	(0.017)	(0.009)			
g-2TFT-25f	0.003	0.051			
	(0.017)	(0.034)			
g-2TFT-50f	0.050	0.021			
	(0.034)	(0.032)			
g-2TFT-75f	0.046	0	**Gamma**	0.478***	0.630***
	(0.048)	(0)		(0.022)	(0.043)

Individual Strategies

結果２（重要）

High Accuracy（ $\boldsymbol{p = 0.9}$）では、半数近くが
g-TFT（**One-shot or Longer**）をプレイ

High Accuracy（ $\boldsymbol{p = 0.9}$）では、半数近くが
あきらかに理論値より強い報復（**RI**）をプレイし、
理論値より低い可能性がある割合は
25％にとどまる

TFT, a family of g-TFT, Memory-1 strategies

	p = 0.9	p = 0.6	(ctd.)	p = 0.9	p = 0.6
TFT	0.026* (0.016)	0.058*** (0.019)	**2TF2T**	0.041* (0.041)	0.027 (0.020)
g-TFT-25f	0.117** (0.053)	0.056* (0.033)	All-C	0.057* (0.003)	0.039** (0.017)
g-TFT-50f	0.187*** (0.056)	0.062 (0.040)	All-D	0.081*** (0.024)	0.425*** (0.056)
g-TFT-75f	0 (0.012)	0.042 (0.032)	Grim	0.012 (0.012)	0.030 (0.024)
TF2T	0.014 (0.022)	0 (0.004)	Grim-2	0.029 (0.017)	0.040 (0.024)
g-TF2T-25f	0.074 (0.059)	0.013 (0.011)	Grim-3	0.064** (0.027)	0.111*** (0.036)
g-TF2T-50f	0.055 (0.047)	0.011 (0.023)			
g-TF2T-75f	0 (0.013)	0 (0.006)			
TF3T	0.113** (0.047)	0.012 (0.013)			
2TFT	0.029* (0.017)	0 (0.009)			
g-2TFT-25f	0.003 (0.017)	0.051 (0.034)			
g-2TFT-50f	0.050 (0.034)	0.021 (0.032)			
g-2TFT-75f	0.046 (0.048)	0 (0)	Gamma	0.478*** (0.022)	0.630*** (0.043)

The family of g-TFT – 30.4% (A family of TFT – 33.0%)
Memory-1 (+ a family of g-2TFT) – 47.1%

Strategies with stronger retaliations (p = 0.9)

	p = 0.9	p = 0.6	(ctd.)	p = 0.9	p = 0.6
TFT	0.026*	0.058***	2TF2T	0.041*	0.027
	(0.016)	(0.019)		(0.041)	(0.020)
g-TFT-25f	0.117**	0.056*	All-C	0.057*	0.039**
	(0.053)	(0.033)		(0.003)	(0.017)
g-TFT-50f	0.187***	0.062	All-D	0.081***	0.425***
	(0.056)	(0.040)		(0.024)	(0.056)
g-TFT-75f	0	0.042	Grim	0.012	0.030
	(0.012)	(0.032)		(0.012)	(0.024)
TF2T	0.014	0	Grim-2	0.029	0.040
	(0.022)	(0.004)		(0.017)	(0.024)
g-TF2T-25f	0.074	0.013	Grim-3	0.064**	0.111***
	(0.059)	(0.011)		(0.027)	(0.036)
g-TF2T-50f	0.055	0.011			
	(0.047)	(0.023)			
g-TF2T-75f	0	0			
	(0.013)	(0.006)			
TF3T	0.113**	0.012			
	(0.047)	(0.013)			
2TFT	0.029*	0			
	(0.017)	(0.009)			
g-2TFT-25f	0.003	0.051			
	(0.017)	(0.034)			
g-2TFT-50f	0.050	0.021			
	(0.034)	(0.032)			
g-2TFT-75f	0.046	0	Gamma	0.478***	0.630***
	(0.048)	(0)		(0.022)	(0.043)

Stronger punishment than g-TFT N.E. - 45.8%, including more than one-shot (g-2TFT)-12.8%
Seemingly weaker strategies (TF2T, g-TF2T, TF3T) – 25.6%, significantly smaller with $p < 0.05$

Individual Strategies

結果３（マイナー）

Low Accuracy（$\boldsymbol{p} = \mathbf{0.6}$）では、**Long Retaliation** をしている割合が
10%程度にとどまる
これは弱い報復を裏付ける

Lenient (Review) Strategies をプレイしている人が少しはいる
しかし、むしろ、**High Accuracy**（$\boldsymbol{p} = \mathbf{0.9}$）において、より割合が高い。

Strategies with long-term punishment (p = 0.6)

	p = 0.9	p = 0.6	(ctd.)	p = 0.9	p = 0.6
TFT	0.026*	0.058***	**2TF2T**	0.041*	0.027
	(0.016)	(0.019)		(0.041)	(0.020)
g-TFT-25f	0.117**	0.056*	**All-C**	0.057*	0.039**
	(0.053)	(0.033)		(0.003)	(0.017)
g-TFT-50f	0.187***	0.062	**All-D**	0.081***	0.425***
	(0.056)	(0.040)		(0.024)	(0.056)
g-TFT-75f	0	0.042	**Grim**	0.012	0.030
	(0.012)	(0.032)		(0.012)	(0.024)
TF2T	0.014	0	**Grim-2**	0.029	0.040
	(0.022)	(0.004)		(0.017)	(0.024)
g-TF2T-25f	0.074	0.013	**Grim-3**	0.064**	0.111***
	(0.059)	(0.011)		(0.027)	(0.036)
g-TF2T-50f	0.055	0.011			
	(0.047)	(0.023)			
g-TF2T-75f	0	0			
	(0.013)	(0.006)			
TF3T	0.113**	0.012			
	(0.047)	(0.013)			
2TFT	0.029*	0			
	(0.017)	(0.009)			
g-2TFT-25f	0.003	0.051			
	(0.017)	(0.034)			
g-2TFT-50f	0.050	0.021			
	(0.034)	(0.032)			
g-2TFT-75f	0.046	0	**Gamma**	0.478***	0.630***
	(0.048)	(0)		(0.022)	(0.043)

Lenient strategies

	p = 0.9	p = 0.6	(ctd.)	p = 0.9	p = 0.6
TFT	0.026*	0.058***	**2TF2T**	0.041*	0.027
	(0.016)	(0.019)		(0.041)	(0.020)
g-TFT-25f	0.117**	0.056*	**All-C**	0.057*	0.039**
	(0.053)	(0.033)		(0.003)	(0.017)
g-TFT-50f	0.187***	0.062	**All-D**	0.081***	0.425***
	(0.056)	(0.040)		(0.024)	(0.056)
g-TFT-75f	0	0.042	**Grim**	0.012	0.030
	(0.012)	(0.032)		(0.012)	(0.024)
TF2T	0.014	0	**Grim-2**	0.029	0.040
	(0.022)	(0.004)		(0.017)	(0.024)
g-TF2T-25f	0.074	0.013	**Grim-3**	0.064**	0.111***
	(0.059)	(0.011)		(0.027)	(0.036)
g-TF2T-50f	0.055	0.011			
	(0.047)	(0.023)			
g-TF2T-75f	0	0			
	(0.013)	(0.006)			
TF3T	0.113**	0.012			
	(0.047)	(0.013)			
2TFT	0.029*	0			
	(0.017)	(0.009)			
g-2TFT-25f	0.003	0.051			
	(0.017)	(0.034)			
g-2TFT-50f	0.050	0.021			
	(0.034)	(0.032)			
g-2TFT-75f	0.046	0	**Gamma**	0.478***	0.630***
	(0.048)	(0)		(0.022)	(0.043)

High accuracy: Memory-2 or more (lenient strategies) - 39.1%

Low accuracy: - 21.3%, significantly smaller (p = 0.014) – Grim 3 (less apparent previously)

（ここからは **Discussion**）

かなりの割合の被験者が **g-TFT** をプレイしている。

しかし、**RI**（報復強度）は **Accuracy** が高いとより強くなっている。理論とは真逆。

実験結果から推測される行動を、均衡理論として説明したい。

そのためには、利己的動機だけでなく、別の行動仮説を導入する必要あり

「行動ゲーム理論」のナッシュ均衡分析！

実験結果に即して

$r(G;p)$、$r(B;p)$、および$r(G;p)-r(B;p)$はみな
pの増加関数とする

各 $p\in[0,1]$について
$(r(G;p),r(B;p))$がナッシュ均衡となる
行動ゲーム理論のモデルを構築する

三つの動機付け要因

利己的動機
Naiveté
Reciprocity (Social Preference)

Naiveté（Randomness, Tremble-Hand）

観察内容とは無関係に

確率$\mathbf{2\varepsilon(p)}$で　　　$\boldsymbol{C}$と$\boldsymbol{D}$をランダムチョイス

確率$\mathbf{1-2\varepsilon(p)}$で　　　合理的にチョイス（ただし「利己的動機」＋「**Social Preference**」）

Social Preferences

Either

Positive Reciprocity： **Good** を観察すると相手をほめたくなる
$\boldsymbol{d}$ をとると心理的コスト $\boldsymbol{Co(G;p)}$ 円

$$Y - Co(G;p) \quad \gtrless \quad \delta Z(2p-1)\{r(G;p) - r(B;p)\}$$

or

Null Reciprocity： 心理的コストなし

$$Y \quad \gtrless \quad \delta Z(2p-1)\{r(G;p) - r(B;p)\}$$

or

Negative Reciprocity： **Bad** を観察すると相手を責めたくなる
$\boldsymbol{c}$ をとると心理的コスト $\boldsymbol{Co(B;p)}$ 円

$$Y \quad \gtrless \quad \delta Z(2p-1)\{r(G;p) - r(B;p)\} - Co(B;p)$$

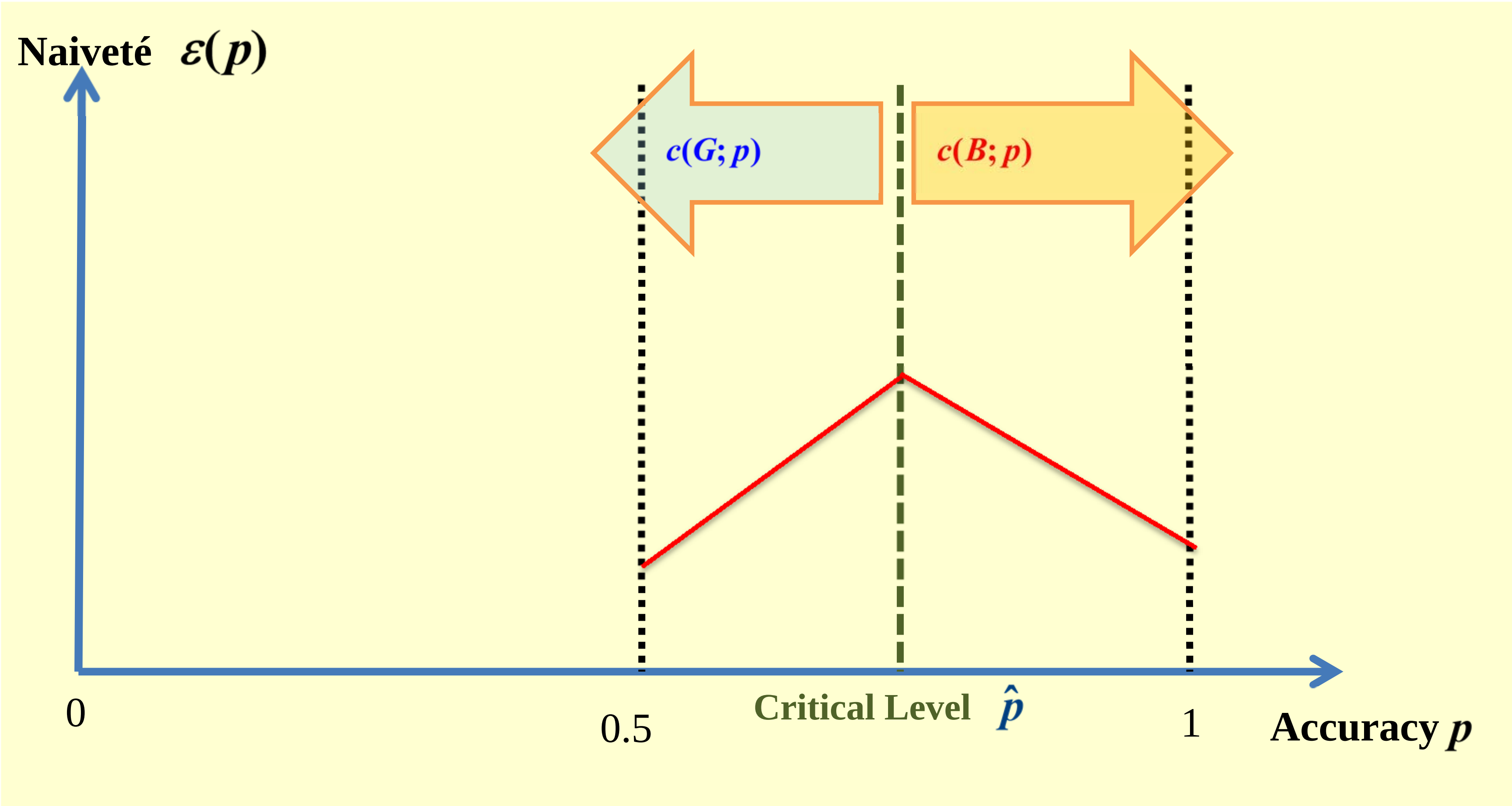
定理：　$(r(G;p), r(B;p))$がナッシュ均衡となる
$\varepsilon(p)$、$Co(G;p)$、$Co(B;p)$は一意に存在する

Critical Accuracy Level $\hat{p} \in [0,1]$が存在して

$\varepsilon(p)$は$\hat{p}$を頂点として **Single-Peaked**

$Co(G;p)$は$p < \hat{p}$において正、pについて減少

$Co(G;p)$は$p > \hat{p}$において正、pについて増加



Naiveté $\varepsilon(p)$
$c(G;p)$
$c(B;p)$
0
0.5
Critical Level $\hat{p}$
1
Accuracy p

二つのトレードオフ（クラウドアウト）

「Social Preference」　vs　「Naiveté」

Social Preference が強くなると

Naiveté (Random, Tremble-Hand)なくなる
More Conscious

「Stick and Carrot」　vs　「Kindness」

＊**Accuracy** が高まるつまり「**Stick and Carrot**」が使えるようになると

相手に対するやさしさや寛容さがなくなる
相手に厳しくなる
Positive Reciprocity → **Null Reciprocity**（$\hat{p}$）→ **Negative Reciprocity**

最後に…脱線

愛情のある社会の樹立について

不完全モニタリングの繰り返しゲームを社会全体に拡張すると......

「飴とムチ（**Stick and Carrot**）」⇒ 「村八分（**Ostracism**）」原理
悪い評判うわさがついてまわる

不完全モニタリング下では
差別偏見との境界はあいまいに（雇用差別、罪のなすりあい）
⇒ 愛情のある社会は可能か？

村八分以外の社会システムの導入
Matsushima (90) "Long-Term Partnership in a Repeated Prisoner's Dilemma with Random Matching"
実験による再考察中

以上