Expectation Conformity in Strategic Cognition

Alessandro Pavan Jean Tirole

Motivation

- Players' understanding of strategic situation often **endogenous**
- Information acquisition
  - payoffs
  - other players' view of the game
- Memory management
- Endogenous depth of reasoning

Motivation

- Cognition
  - **self-directed** (affecting player's own view of the game)
  - **manipulative** (affecting other players' view of the game)
- Manipulative cognition
  - framing and signal jamming
  - disclosure/noisy communication
  - defensive measures (counter-intelligence)

This Paper

- How cognitive choices depend on
  - type of strategic interaction (e.g., complements vs substitutes)
  - beliefs over selected cognition
- **Expectation conformity**
  - Cognitive choice reflects *beliefs* over
    - opponents' expectation about own's cognition
    - opponents' actual cognition
- ...use decomposition to shed light on
  - determinacy of equilibria
  - relation between strategic interaction and cognitive posture

Literature – Incomplete

- **Information acquisition**
  - auctions:...Persico (2000)...
  - mechanism design:...Bergemann and Välimäki (2002)...
  - contracting:...Crémer and Khalil (1992, 1994), Crémer et al. (1998a,b), Dang (2008), Tirole (2009), Bolton and Faure-Grimaud (2010), Pavan and Tirole (2021a,b)...
  - security design:...Farhi and Tirole (2015), Dang et al (2017), Yang (2020)...
  - linear-quadratic + global games:...Hellwig and Veldkamp (2009), Myatt and Wallace (2012), Colombo, Femminis, Pavan (2014), Szkup and Trevino (2015), Pavan (2016), Morris and Yang (2019), Denti (2020), Liang and Mu (2020), Banerjee et al. (2021)...
- **Interpreting other players' views + noisy communication**
  - ...Dewatripont and Tirole (2005), Che and Kartik (2009), Calvo-Armengol et al. (2015), Sethi and Yildiz (2016, 2018), Kozlovskaya (2018), Adriani and Sonderegger (2020)...
- **Signal jamming and framing**
  - ...Fudenberg and Tirole (1986), Holmström (1999), Dewatripont et al. (1999), Salant and Siegel (2018), Horner and Lambert (2019)...
- **Sparsity and endogenous depth of reasoning**
  - ...Gabaix (2014), Alaoui and Penta (2016, 2017, 2018)...
- **Psychological games**
  - ...Geanakoplos et al., (1988), Battigalli and Dufwenberg (2009)...

Plan

1. Introduction
2. Model
3. Expectation conformity
4. Equilibrium determinacy
5. Sparsity
6. Espionage and counter-espionage

Model

Model

- **Primitive game**
  - n players: $i \in I$
  - A_i: action set
  - $u_i(\alpha_i, \alpha_{-i}, \omega)$: gross payoff
  - mixed actions: α_i, α_{-i}
  - payoff-relevant state: $\omega \in \Omega$ (prior F)

Model

- **Cognition**
  - $\rho = (\rho_i, \rho_{-i})$: cognitive profile
  - $C_i(\rho_i)$: cognitive cost
- **Info acquisition**
  - $Q(s|\omega, \rho) \in \Delta(S)$: "signals/beliefs" distribution
    - $S = \mathsf{X}_{i \in I} S_i$
    - Self-directed cognition: $Q(s|\omega, \rho) = \otimes_{i \in I} Q_i(s_i|\omega, \rho_i)$

Model

- Stage-2 strategies
  - $\sigma_i : S_i \to \Delta(A_i)$
  - σ^ρ: stage-2 BNE given cognition ρ

Model

- Ex-ante gross payoff

$$U_i(\sigma;\rho) \equiv \int_\omega \Big[\int_s u_i(\sigma_i(s_i), \sigma_{-i}(s_{-i}), \omega) \mathrm{d}Q(s|\omega,\rho) \Big] \mathrm{d}F(\omega)$$

- Value function

$$V_i(\rho_i';\rho) \equiv \sup_{\sigma_i \in \Delta(A_i)^{S_i}} U_i(\sigma_i, \sigma_{-i}^{\rho}; \rho_i', \rho_{-i})$$

- Net ex-ante payoff

$$V_i(\rho_i';\rho) - C_i(\rho_i')$$

Plan

1. Introduction
2. Model
3. Expectation conformity
4. Equilibrium determinacy
5. Sparsity
6. Espionage and counter-espionage

Expectation Conformity

Expectation Conformity

Definition 1

Expectation conformity (EC) holds for ρ and $\hat{\rho}$ if, for all i,

$$\Gamma_i^{EC}(\rho, \hat{\rho}) \equiv \Big[V_i\big(\hat{\rho}_i; \hat{\rho}_i\,, \hat{\rho}_{-i}\big) - V_i\big(\rho_i; \hat{\rho}_i\,, \hat{\rho}_{-i}\big) \Big] - \Big[V_i\big(\hat{\rho}_i; \rho_i\,, \rho_{-i}\big) - V_i\big(\rho_i; \rho_i\,, \rho_{-i}\big) \Big] \geq 0$$

Within and Across Synergies

Definition 2

Unilateral expectation conformity (UEC) holds for ρ and $\hat{\rho}$ if, for all i,

$$\Gamma_i^{UEC}(\rho,\hat{\rho}) \equiv \Big[V_i(\hat{\rho}_i;\hat{\rho}_i\,,\rho_{-i}) - V_i(\rho_i;\hat{\rho}_i\,,\rho_{-i})\Big] - \Big[V_i(\hat{\rho}_i;\rho_i\,,\rho_{-i}) - V_i(\rho_i;\rho_i\,,\rho_{-i})\Big] \geq 0$$

Definition 3

Increasing differences holds for ρ and $\hat{\rho}$ if, for all i,

$$\Gamma_i^{ID}(\rho,\hat{\rho}) \equiv \Big[V_i(\hat{\rho}_i;\hat{\rho}_i\,,\hat{\rho}_{-i}) - V_i(\rho_i;\hat{\rho}_i\,,\hat{\rho}_{-i})\Big] - \Big[V_i(\hat{\rho}_i;\hat{\rho}_i\,,\rho_{-i}) - V_i(\rho_i;\hat{\rho}_i\,,\rho_{-i})\Big] \geq 0$$

Plan

1. Introduction
2. Model
3. Expectation conformity
4. Equilibrium determinacy
5. Sparsity
6. Espionage and counter-espionage

Equilibrium Determinacy

Equilibrium Determinacy

Proposition 1

If EC holds for ρ and $\hat{\rho}$, there exist $(C_i)_{i \in I}$ s.t. ρ and $\hat{\rho}$ are eq. profiles.

If cognition self-directed and totally ordered, $(C_i)_{i \in I}$ monotone.

If EC not satisfied for any ρ and $\hat{\rho}$, then unique eq., irrespective of $(C_i)_{i \in I}$.

(Eq-det-proof)

Cognitive Traps

Definition 4

Suppose cognition self-directed and cognitive profiles totally ordered. Players exposed to **cognitive trap** if there exist ρ and $\hat{\rho}$ s.t.

(i) ρ and $\hat{\rho}$ are equilibria

(ii) for all i s.t. $\hat{\rho}_i \neq \rho_i$, $\hat{\rho}_i$ Blackwell more informative than ρ_i and

$$V_i(\hat{\rho}_i; \hat{\rho}) - C_i(\hat{\rho}_i) < V_i(\rho_i; \rho) - C_i(\rho_i)$$

Two-person Constant-Sum Games

- For all $(\alpha_i, \alpha_j, \omega)$
$$u_i(\alpha_i, \alpha_j, \omega) + u_j(\alpha_i, \alpha_j, \omega) = k(\omega).$$

Proposition 2

For all $(\rho, \hat{\rho})$,
$$\Sigma_i \, \Gamma_i^{EC}(\rho, \hat{\rho}) \leq 0.$$
If there are multiple equilibria, in none can player have strict preference for her eq.

cognition over her cognition in any other eq.

Plan

1. Introduction
2. Model
3. Expectation conformity
4. Equilibrium determinacy
5. Sparsity
6. Espionage and counter-espionage

Sparsity

Sparsity

- Features of sparsity:
  - rich state space
  - attention to subset of dimensions
  - other dimensions “as if” did not exist
- Typically: bounded rationality
- Here: rational players

Sparsity

- For simplicity: 2 players
- Payoffs:
$$u_i(a_i, a_j, \omega) = -(1-\beta)(a_i - g(\omega))^2 - \beta(a_i - a_j)^2$$
- $a_i, a_j \in \mathbb{R}$
- $\omega \equiv (\omega^k)_{k=1}^K$, ω^k drawn independently from F^k, $\mathbb{E}[\omega^k] = 0$, $Var[\omega^k] = \sigma_k^2$
- $g(\omega) = (1 + \Sigma_{k=1}^K \omega^k)/(1-\beta)$

Sparsity

- Natural progression: dimension k explored only if all $k' < k$ also explored
  - can be microfounded
- **Cognition: number of dimensions $\rho_i \in \mathbb{N}$ explored**
  - self-directed
  - ordered
- Player i's signal: $s_i = (\omega^1, ..., \omega^{\rho_i})$

Sparsity

- Player 1: *follower* $(\rho_1 \leq \rho_2)$
- Player 2: *leader*
- Eq. actions, given $\rho = (\rho_1, \rho_2)$

$$a_1^{\rho}(s_1) = \frac{1 + \Sigma_{k=1}^{\rho_1} \omega^k}{1 - \beta}$$

$$\begin{aligned} a_2^{\rho}(s_2) &= \tfrac{1 + \Sigma_{k=1}^{\rho_1} \omega^k}{1 - \beta} + \Sigma_{k=\rho_1+1}^{\rho_2} \omega^k \\ &= a_1^{\rho}(s_1) + \Sigma_{k=\rho_1+1}^{\rho_2} \omega^k \end{aligned}$$

- Features:
  - **unexplored dimensions treated "as if" did not exist**
  - **leader predicts perfectly follower's beliefs (and actions)**
  - **follower reasons "as if" leader's knowledge same as hers**

Sparsity

Proposition 3

Let $\hat{\rho}$ and ρ be s.t. $\hat{\rho}_2 > \rho_2 \geq \hat{\rho}_1 > \rho_1$:

- UEC holds strictly for 1 (follower), weakly for 2 (leader)

- ID holds as equality for both players

- Follower's action invariant to number of dimensions explored solely by leader
- Follower does not benefit from surprising leader
  - leader responds more to dimensions commonly explored when $\beta > 0$
  - leader responds less to dimensions commonly explored when $\beta < 0$
- Usefulness: determinacy of *asymmetric equilibria*

Sparsity

Proposition 4

Let $\hat{\rho}$ and ρ be s.t. $\hat{\rho}_2 = \hat{\rho}_1 > \rho_2 = \rho_1$:

- UEC holds as equality

- ID holds if $\beta > 0$ but not if $\beta < 0$

- Exploring dimensions jointly
  - more valuable when $\beta > 0$
  - less valuable when $\beta < 0$
- Result suggests (symmetric C_i)
  - unique symm eq. w. substitutes
  - multiple symm eq. with complements

Sparsity: Strategic Complements

- $C_i(\rho_i) \equiv \Sigma_{k=1}^{\rho_i} c_k$
  - σ_k^2 / c_k decreasing
  - $\lim_{k \to K} \sigma_k^2 / c_k = 0$

Sparsity: Strategic Complements

Proposition 5

All (pure-strategy) equilibria symmetric:

- *Any* $k^* \in [\underline{k}, \bar{k}(\beta)]$ *part of symmetric (pure-strategy) eq., with*

$$\underline{k} \equiv \min \left\{ k \middle| \sigma_k^2 \leq c_k \right\} \qquad \text{and} \qquad \bar{k}(\beta) \equiv \max \left\{ k \middle| \frac{\sigma_k^2}{(1-\beta)^2} \geq c_k \right\}.$$

- *Equilibria Pareto ranked: players' net payoff increasing in* $k^* \in [\underline{k}, \bar{k}(\beta)]$
- *Equilibria robust to endogenous order*

Sparsity: Strategic Substitutes

Proposition 6

Symmetric (pure-strategy) eq. exists iff there is $k^* \in \mathbb{N}$ s.t.

$$\frac{\sigma^2_{k^*+1}}{c_{k^*+1}} \leq 1 \leq \frac{\sigma^2_{k^*}}{(1-\beta)^2 c_{k^*}}$$

- *At most one symmetric (pure-strategy) eq.*
- *Asymmetric (pure-strategy) eq. may exist.*
- *Follower' eq. payoff*
  - *increasing in own cognition*
  - *invariant in leader's cognition*
- *Leader's eq. payoff*
  - *decreasing in follower's cognition*
- *Sum of eq. payoffs maximal when follower's cognition lowest*

Plan

1. Introduction
2. Model
3. Expectation conformity
4. Equilibrium determinacy
5. Sparsity
6. Espionage and counter-espionage

Espionage and counter-espionage

Self-directed cognition: Espionage

- Payoffs:

$$u_i(a_i, a_j, \omega) = -(1-\beta)(a_i - g(\omega))^2 - \beta(a_i - a_j)^2$$

- $\omega \in \mathbb{R}$

- $g(\omega) = \omega$

- Signals:

  - primary (exogenous): $s_i^P = \omega + \varepsilon_i$, with $\varepsilon_i \sim N(0, 1)$

  - secondary (endogenous): $s_i^S = s_j^P + \eta_i$, with $\eta_i \sim N(0, \rho_i^{-1})$

Self-directed cognition: Espionage

Proposition 7

UEC always holds. ID holds iff $\beta\left(\hat{\rho}_i - \rho_i\right)\left(\hat{\rho}_j - \rho_j\right) \leq 0$.

- UEC: when j expects i to "spy" more, a_j more sensitive to s_j^P if $\beta > 0$ (less sensitive if $\beta < 0$) $\Rightarrow$ higher value for i to spy
- ID: more spying by $j \Rightarrow$ lower sensitivity of a_j to $s_j^P \Rightarrow$ lower value for i to spy when $\beta > 0$ (higher value when $\beta < 0$)

Manipulative cognition: Counter-espionage

- Payoffs:

$$u_i(a_i, a_j, \omega) = -(1-\beta)(a_i - \omega) - \beta(a_i - a_j)^2$$

- Signals:
  - primary (exogenous): $s_i^P = \omega + \varepsilon_i$, with $\varepsilon_i \sim N(0, 1)$
  - secondary (endogenous): $s_i^S = s_j^P + \gamma_j$, with $\gamma_j \sim N(0, \rho_j^{-1})$
  - counter-espionage or noisy info sharing

Manipulative cognition: Counter-espionage

Proposition 8

UEC holds iff $\beta > 0$*. ID holds iff* $(\hat{\rho}_i - \rho_i)(\hat{\rho}_j - \rho_j) \leq 0$

- UEC: j expects more precise $s_j^S \Rightarrow a_j$ more sensitive to $s_j^S \Rightarrow$ more value to i to make s_j^S precise when $\beta > 0$ (lower value when $\beta < 0$)
- ID: when j sends more accurate signal, a_j more sensitive to s_j^P when $\beta > 0$ (less sensitive when $\beta < 0$) $\Rightarrow$ lower value for i to reciprocate by sending more precise secondary signal

Additional Material

- Paper (+ Supplement) also cover
  - noisy information acquisition about exogenous payoff states
    - case most studied in literature
  - other modes of manipulative cognition:
    - generalized career-concerns models
    - framing and defensive memory management
  - endogenous depth of reasoning

Conclusions

- Expectation conformity in strategic reasoning
    - UEC (beliefs over opponents' expectation about own cognition)
    - ID (beliefs over opponents' actual cognition)
- EC helps predicting/interpreting
    - equilibrium determinacy
    - sensitivity of cognition to downstream strategic interaction
    - cognitive choices...without fixed points

Most Important Slide

THANKS!

Framing

- Persuasion game
  - Player 1: Sender
  - Player 2: Receiver

- Sender's payoff: $u_1(a_1, a_2, \omega) = a_2$

- Receiver's payoff: $u_2(a_2, \omega) = -(a_2 - \omega)^2$

- Receiver's knowledge of ω recalled with probability

$$r(\omega; \rho) = \begin{cases} r^-(\rho_2) & \text{if} \quad \omega < 0 \\ r^+(\rho_1, \rho_2) & \text{if} \quad \omega \geq 0 \end{cases}$$

- With prob $1 - r(\omega; \rho)$, Receiver recalls $\emptyset$
- $\bar{\omega}(\hat{\rho}_1, \hat{\rho}_2) = \mathbb{E}\left[\tilde{\omega} | \emptyset\right]$
- $\omega^- = \mathbb{E}\left[\tilde{\omega} | \tilde{\omega} < 0\right]$ and $\omega^+ = \mathbb{E}\left[\tilde{\omega} | \tilde{\omega} \geq 0\right]$
- Frames: design of contextual purchasing experience

Framing

Proposition 9

UEC holds weakly for Receiver, strictly for Sender

- *ID holds for Sender iff*

$$\left[r^+(\hat{\rho}_1,\hat{\rho}_2) - r^+(\rho_1,\hat{\rho}_2)\right]\left[\omega^+ - \bar{\omega}(\hat{\rho}_1,\hat{\rho}_2)\right] \geq$$

$$\left[r^+(\hat{\rho}_1,\rho_2) - r^+(\rho_1,\rho_2)\right]\left[\omega^+ - \bar{\omega}(\hat{\rho}_1,\rho_2)\right]$$

- Sender's UEC: $\mathbb{E}\left[\tilde{\omega}|\emptyset;\rho\right]$ smaller under less framing $\Rightarrow$ stronger incentives to frame
  - signal jamming
- ID satisfied
  - when r^+ weakly supermodular
  - $\mathbb{E}\left[\tilde{\omega}|\emptyset;\rho\right]$ weakly decreasing in ρ_2

Endogenous Depth of Reasoning

- Complete-information
- **Cognition ρ_i: steps of iterated best responses**
- Sequence of (mixed) actions $(\alpha_i^k)_k$
    - α_i^0: "anchor"
    - $\alpha_i^k = BR_i(\alpha_j^{k-1})$

Endogenous Depth of Reasoning

- Stage-2 game: Alaoui and Penta (2016, 2017,2018)

$$\sigma_i^{\rho_i';\rho} = \begin{cases} \alpha_i^{\rho_i'} & \text{if} \quad \rho_i' \leq \min\{\rho_i + 1, \rho_j\} + 1 \\ \alpha_i^{\min\{\rho_i+1,\rho_j\}+1} & \text{if} \quad \rho_i' > \min\{\rho_i + 1, \rho_j\} + 1 \end{cases}$$

- Action identified by depth of reasoning, unless capacity exceeds what necessary to predict α_j
- Players understand impact of cognition on payoffs
- ...even if unable to predict BR
- Feature: $V_i(\rho_i'; \rho)$ not monotone in ρ_i' (bounded rationality)

Arad and Rubinstein 11-20 undercutting game

- $A_i = \{11, 12, ..., 20\}$
- Payoffs

$$u_i(a_i, a_j) = \begin{cases} a_i + x & \text{if} \quad a_i = a_j - 1 \\ a_i + 10 & \text{if} \quad a_i = a_j \\ a_i & \text{otherwise} \end{cases}$$

Arad and Rubinstein 11-20 undercutting game

- Let $\hat{\rho}$ and ρ s.t.
  - $\hat{\rho}_1 > \rho_1$
  - $\rho_2 = \rho_1 + 1$
  - $\hat{\rho}_2 = \hat{\rho}_1 + 1$.

Proposition 10

$\Gamma_1^{UEC} = \Gamma_2^{UEC} = 0$ *and* $\Gamma_1^{ID} < 0 < \Gamma_2^{ID}$

- Reason for $\Gamma_i^{UEC} = 0$: players either at capacity or think their cognition suffices to predict opponent's action
- Reason for $\Gamma_2^{ID} > 0$: leader over-cuts when going deeper – smaller damage when follower also goes deeper

Arad and Rubinstein 11-20 undercutting game

- Let $\hat{\rho}$ and ρ s.t. $\hat{\rho}_2 = \hat{\rho}_1 > \rho_2 = \rho_1$

Proposition 11

$\Gamma_i^{UEC} = 0$ *whereas* $\Gamma_i^{ID} < 0$, $i = 1, 2$

- Reason for $\Gamma_2^{ID} < 0$: ability to perfectly over-cut rival diminished by rival's cognition

Equilibrium determinacy: Proof Sketch

- EC satisfied for $\rho = (\rho_i, \rho_{-i})$ and $\hat{\rho} = (\hat{\rho}_i, \hat{\rho}_{-i}) \Rightarrow$ there exist (C_i) s.t.

$$V_i(\hat{\rho}_i;\hat{\rho}) - V_i(\rho_i;\hat{\rho}) \geq C_i(\hat{\rho}_i) - C_i(\rho_i) \geq V_i(\hat{\rho}_i;\rho) - V_i(\rho_i;\rho) \qquad EC_{\{\rho,\hat{\rho}\}}$$

- Cognition self-directed and totally ordered $(\hat{\rho}_i > \rho_i)$

$$C_i(\tilde{\rho}_i) = \begin{cases} C_i(\rho_i) & \text{for} \quad \tilde{\rho}_i \leq \rho_i \\ C_i(\hat{\rho}_i) & \text{for} \quad \rho_i < \tilde{\rho}_i \leq \hat{\rho}_i \\ +\infty & \text{for} \quad \tilde{\rho}_i > \hat{\rho}_i. \end{cases}$$

- Self-directed cognition
  - covert
  - more information always beneficial: $V_i(\hat{\rho}_i;\rho) - V_i(\rho_i;\rho) \geq 0$
  - $C_i(\hat{\rho}_i) \geq C_i(\rho_i)$

- ρ and $\hat{\rho}$ are eq. profiles $\Rightarrow EC_{\{\rho,\hat{\rho}\}}$. Hence if EC holds for no pair of profiles, unique eq,. no matter (C_i)

Go back