Collusion-Proof Dynamic Mechanisms

Heng Liu
University of Michigan

July 10, 2018
IMS, Dynamic Models in Economics
NUS

Motivation

- Most of the dynamic mechanism design literature focuses on the "truthtelling" equilibrium (e.g. Bergemann and Välimäki (2010), Athey and Segal (2013)).

Motivation

- Most of the dynamic mechanism design literature focuses on the "truthtelling" equilibrium (e.g. Bergemann and Välimäki (2010), Athey and Segal (2013)).

- Agents have more opportunities to coordinate or collude in dynamic settings.

Motivation

- Most of the dynamic mechanism design literature focuses on the "truthtelling" equilibrium (e.g. Bergemann and Välimäki (2010), Athey and Segal (2013)).
- Agents have more opportunities to coordinate or collude in dynamic settings.
- Propose a framework to address the possibility of collusion in dynamic mechanisms.
- Main question: Which dynamic mechanisms are immune to collusion?

Motivation

- Most of the dynamic mechanism design literature focuses on the "truthtelling" equilibrium (e.g. Bergemann and Välimäki (2010), Athey and Segal (2013)).
- Agents have more opportunities to coordinate or collude in dynamic settings.
- Propose a framework to address the possibility of collusion in dynamic mechanisms.
- Main question: Which dynamic mechanisms are immune to collusion?
- Define collusion-proofness in dynamic settings.
- Construct collusion-proof dynamic mechanisms.
- Characterize collusion-proofness in stationary settings.

Related Literature

- Collusion-proof static mechanisms: Laffont and Martimort (1997, 2000), Che and Kim (2006), Safronov (2017)
- Efficient dynamic mechanisms: Bergemann and Välimäki (2010), Athey and Segal (2013), Skrzypacz and Toikka (2015)
- Optimal dynamic mechanisms: Pavan, Segal and Toikka (2014), Pavan (2016), Bergemann and Välimäki (2017)
- Collusion with persistence private info.: Athey and Bagwell (2001, 2008), Miller (2012)
- Repeated implementation: Jackson and Sonnenschein (2007), Lee and Sabourian (2009, 2013), Renou and Mezzetti (2017), Renou and Tomala (2015), Chassang and Ortner (2015)

Model: IPV w/ transfers

- Time: $t = 1, 2, \ldots, T$ ($T \leq \infty$).
- Agents: $i \in \{1, 2, \ldots, N\} = \mathcal{N}$. $N \geq 2$.
- Private type: $\forall\, t \geq 1$, $\theta_t^i \in \Theta^i$. $\theta_t \triangleq (\theta_t^1, \ldots, \theta_t^N) \in \prod_i \Theta^i \triangleq \Theta$.
- Allocations: $a_t \in A$.
- Flow payoff: $u^i(a_t, \theta_t^i) - p_t^i$. ("private values")
- Discounted payoff:

$$\mathbb{E}\left\{\sum_{t\geq 1} \delta^{t-1}\left[u^i(a_t, \theta_t^i) - p_t^i\right]\right\}.$$

- Common prior: $\mu_1^i(\cdot) \in \Delta(\Theta^i)$. ("independence")
- Markov transition: $\mu^i(\cdot\,|a_{t-1}, \theta_{t-1}^i) \in \Delta(\Theta^i)$.

Dynamic Mechanisms

To simplify notations, consider public mechanisms where all the past reported types are public to all agents.

A dynamic mechanism is $M = (a_t, p_t)_{t \geq 1}$ where $\forall t \geq 1$,

- allocations: $a_t : \Theta^{t-1} \times A^{t-1} \times \Theta \to \Delta(A)$
- transfers: $p_t = (p_t^i)_{i \in \mathcal{N}}$ with $p_t^i : \Theta^{t-1} \times A^{t-1} \times \Theta \to \mathbb{R}$

Dynamic Mechanisms

To simplify notations, consider public mechanisms where all the past reported types are public to all agents.

A dynamic mechanism is $M = (a_t, p_t)_{t\geq 1}$ where $\forall\, t \geq 1$,

- allocations: $a_t : \Theta^{t-1} \times A^{t-1} \times \Theta \to \Delta(A)$
- transfers: $p_t = (p_t^i)_{i\in\mathcal{N}}$ with $p_t^i : \Theta^{t-1} \times A^{t-1} \times \Theta \to \mathbb{R}$

Given M, a strategy $\sigma^i = (\sigma_t^i)_{t\geq 1}$ of agent i is

$$\sigma_t^i : \Theta^{t-1} \times A^{t-1} \times {\Theta^i}^t \to \Delta(\Theta^i).$$

Agent i's expected payoff under M and strategy profile $\sigma = (\sigma^i)$ is

$$\mathbb{E}_{M,\mu,\sigma}\left[\sum_{t\geq 1} \delta^{t-1}\left(u^i(\tilde{a}_t, \tilde{\theta}_t^i) - \tilde{p}_t^i\right)\right].$$

IC, IR, & BB

Truthtelling strategy $\sigma^{i*} = (\sigma_t^{i*})_{t\geq 1}$: $\forall t,\ \theta^{t-1}, a^{t-1}, \theta_t^i$,

$$\sigma_t^{i*}(\theta^{t-1}, a^{t-1}, \theta_t^i) = \mathbf{1}_{\{\theta_t^i\}}.$$

- per-period interim IC (IC): truthtelling is a wPBE
- per-period ex post IC (epIC): per-period ex post eq.

IC, IR, & BB

Truthtelling strategy $\sigma^{i*} = (\sigma_t^{i*})_{t\geq 1}$: $\forall t,\ \theta^{t-1}, a^{t-1}, \theta_t^i$,

$$\sigma_t^{i*}(\theta^{t-1}, a^{t-1}, \theta_t^i) = \mathbf{1}_{\{\theta_t^i\}}.$$

- per-period interim IC (IC): truthtelling is a wPBE
- per-period ex post IC (epIC): per-period ex post eq.
- ex ante IR (IR_0): ex ante payoff under truthtelling $\geq \bar{U}^i$
- per-period interim IR (IR): interim payoff $\geq \bar{U}^i(\theta_t^i)$
- per-period ex post IR (epIR): ex post payoff $\geq \bar{U}^i(\theta_t)$

IC, IR, & BB

Truthtelling strategy $\sigma^{i*} = (\sigma_t^{i*})_{t\geq 1}$: $\forall t,\ \theta^{t-1}, a^{t-1}, \theta_t^i,$

$$\sigma_t^{i*}(\theta^{t-1}, a^{t-1}, \theta_t^i) = \mathbf{1}_{\{\theta_t^i\}}.$$

- per-period interim IC (IC): truthtelling is a wPBE
- per-period ex post IC (epIC): per-period ex post eq.
- ex ante IR (IR_0): ex ante payoff under truthtelling $\geq \bar{U}^i$
- per-period interim IR (IR): interim payoff $\geq \bar{U}^i(\theta_t^i)$
- per-period ex post IR (epIR): ex post payoff $\geq \bar{U}^i(\theta_t)$
- ex post budget balance (BB): $\sum_i p_t^i = 0,\ \forall t \geq 1$

Efficiency

An allocation $a^* = (a_t^*)$ w/ $a_t^* : \Theta \to \Delta(A)$ is **efficient** if it solves

$$\max_{(a_t)} \mathbb{E}\left[\sum_{t \geq 1} \delta^{t-1} \sum_{i \in \mathcal{N}} u^i(a_t(\tilde{\theta}_t), \tilde{\theta}_t^i)\right]$$

Efficiency

An allocation $a^* = (a_t^*)$ w/ $a_t^* : \Theta \to \Delta(A)$ is **efficient** if it solves

$$\max_{(a_t)} \mathbb{E}\left[\sum_{t\geq 1} \delta^{t-1} \sum_{i\in\mathcal{N}} u^i(a_t(\tilde{\theta}_t), \tilde{\theta}_t^i)\right]$$

An allocation $\bar{a}_t : \Theta \to \Delta(A)$ is **incentive efficient** if

$$\mathbb{E}\left[\sum_{t\geq 1} \delta^{t-1} \sum_{i\in\mathcal{N}} u^i(\bar{a}_t(\tilde{\theta}_t), \tilde{\theta}_t^i)\right] \geq \mathbb{E}\left[\sum_{t\geq 1} \delta^{t-1} \sum_{i\in\mathcal{N}} u^i(\bar{a}_t(\gamma_t(\tilde{\theta}_t, \tilde{\theta}^{t-1}, \tilde{a}^{t-1})), \tilde{\theta}_t^i\right.$$

for all $\gamma = (\gamma_t)_{t\geq 1}$ where $\gamma_t : \Theta \times \Theta^{t-1} \times A^{t-1} \to \Delta(\Theta)$.

Efficiency

An allocation $a^* = (a^*_t)$ w/ $a^*_t : \Theta \to \Delta(A)$ is **efficient** if it solves

$$\max_{(a_t)} \mathbb{E}\left[\sum_{t\geq 1} \delta^{t-1} \sum_{i\in\mathcal{N}} u^i(a_t(\tilde{\theta}_t), \tilde{\theta}^i_t)\right]$$

An allocation $\bar{a}_t : \Theta \to \Delta(A)$ is **incentive efficient** if

$$\mathbb{E}\left[\sum_{t\geq 1} \delta^{t-1} \sum_{i\in\mathcal{N}} u^i(\bar{a}_t(\tilde{\theta}_t), \tilde{\theta}^i_t)\right] \geq \mathbb{E}\left[\sum_{t\geq 1} \delta^{t-1} \sum_{i\in\mathcal{N}} u^i(\bar{a}_t(\gamma_t(\tilde{\theta}_t, \tilde{\theta}^{t-1}, \tilde{a}^{t-1})), \tilde{\theta}^i_t\right.$$

for all $\gamma = (\gamma_t)_{t\geq 1}$ where $\gamma_t : \Theta \times \Theta^{t-1} \times A^{t-1} \to \Delta(\Theta)$.

A mechanism $(\bar{a}, p)$ is **incentive efficient** if $\forall \gamma$,

$$\mathbb{E}\left[\sum_{t\geq 1} \delta^{t-1} \sum_{i\in\mathcal{N}} \left(u^i(\bar{a}_t(\tilde{\theta}_t), \tilde{\theta}^i_t) - p^i_t(\tilde{\theta}^{t-1}, \tilde{a}^{t-1}, \tilde{\theta}_t)\right)\right]$$

$$\geq \mathbb{E}\left[\sum_{t\geq 1} \delta^{t-1} \sum_{i\in\mathcal{N}} \left(u^i(\bar{a}_t(\gamma_t(\tilde{\theta}_t), \tilde{\theta}^i_t) - p^i_t(\tilde{\theta}^{t-1}, \tilde{a}^{t-1}, \gamma_t(\tilde{\theta}_t)\right)\right]$$

Modeling Collusion

Given a dynamic mechanism M, a mediator can coordinate collusion among (subgroups of) agents: $\forall t$

- collect reports from agents then jointly report to the designer
- make balanced transfers among agents

Modeling Collusion

Given a dynamic mechanism M, a mediator can coordinate collusion among (subgroups of) agents: $\forall t$

- collect reports from agents then jointly report to the designer
- make balanced transfers among agents

Formally, M induces a dynamic game G_M among agents with outside options. Given G_M, a collusion scheme (among all agents) $\Gamma = (\gamma, q)$ is a mediated game (or mechanism)

- $\gamma = (\gamma_t)_{t\geq 1}$, $\gamma_t : \Theta \times \Theta^{t-1} \times A^{t-1} \to \Delta(\Theta)$
- $q = (q_t^i)_{i,t}$, $q_t^i : \Theta \times \Theta^{t-1} \times A^{t-1} \to \mathbb{R}$ & $\sum_i q_t^i = 0$.

Given M & Γ, agents play $G_{M\Gamma}$ (with outside options).

Focus on all IC (& IR) Γ's.

Collusion-Proofness

A dynamic mechanism M is collusion-proof if the expected payoffs of all agents under all IC Γ's are the same as the expected payoffs in M under truthtelling.

Collusion-proofness: the set of equilibrium payoff vectors under mediation in G_M is a singleton, which equals the payoff vector from truthtelling in G_M.

Collusion-Proofness

A dynamic mechanism M is collusion-proof if the expected payoffs of all agents under all IC Γ's are the same as the expected payoffs in M under truthtelling.

Collusion-proofness: the set of equilibrium payoff vectors under mediation in G_M is a singleton, which equals the payoff vector from truthtelling in G_M.

Collusion-proofness is defined without referring to IR.

- can define IR w.r.t. G_M or outside option
- beliefs after rejecting a collusion scheme

Remarks

The dynamic pivot mechanism (Bergemann & Välimäki, 2010) is not collusion-proof.

The balanced-team mechanism (Athey & Segal, 2013) is collusion-proof when $N = 2$ but not when $N \geq 3$.

A mechanism with a constant allocation rule is collusion-proof.

Results

$\star$ incentive efficiency $\Rightarrow$ collusion-proofness

Proposition 1

If $\bar{a}$ is incentive efficient, then $\exists$ a BB transfer p s.t. $(\bar{a}, p)$ is IC & collusion-proof.

Sketch of Proof: i.i.d. case

Bound the minmax payoff under truthtelling: $\forall \sigma_{-i}$

$$\mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}^i, \sigma_{-i}(\tilde{\theta}^{-i})), \tilde{\theta}^i) - p^i(\tilde{\theta}^i, \sigma_{-i}(\tilde{\theta}^{-i}))\right] \geq \mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}^i, \tilde{\theta}^{-i}), \tilde{\theta}^i)\right]$$

Sketch of Proof: i.i.d. case

Bound the minmax payoff under truthtelling: $\forall \sigma_{-i}$

$$\mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}^i, \sigma_{-i}(\tilde{\theta}^{-i})), \tilde{\theta}^i) - p^i(\tilde{\theta}^i, \sigma_{-i}(\tilde{\theta}^{-i}))\right] \geq \mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}^i, \tilde{\theta}^{-i}), \tilde{\theta}^i)\right]$$

$\Rightarrow$

$$\mathbb{E}\left[p^i(\tilde{\theta}^i, \sigma_{-i}(\tilde{\theta}^{-i}))\right] \leq \mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}^i, \sigma_{-i}(\tilde{\theta}^{-i})), \tilde{\theta}^i) - u^i(\bar{a}(\tilde{\theta}^i, \tilde{\theta}^{-i}), \tilde{\theta}^i)\right]$$

Sketch of Proof: i.i.d. case

Bound the minmax payoff under truthtelling: $\forall \sigma_{-i}$

$$\mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}^i, \sigma_{-i}(\tilde{\theta}^{-i})), \tilde{\theta}^i) - p^i(\tilde{\theta}^i, \sigma_{-i}(\tilde{\theta}^{-i}))\right] \geq \mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}^i, \tilde{\theta}^{-i}), \tilde{\theta}^i)\right]$$

$\Rightarrow$

$$\mathbb{E}\left[p^i(\tilde{\theta}^i, \sigma_{-i}(\tilde{\theta}^{-i}))\right] \leq \mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}^i, \sigma_{-i}(\tilde{\theta}^{-i})), \tilde{\theta}^i) - u^i(\bar{a}(\tilde{\theta}^i, \tilde{\theta}^{-i}), \tilde{\theta}^i)\right]$$

$$+ \underbrace{\sum_{j \neq i} \mathbb{E}\left[\mathbb{E}_{\tilde{\theta}^i}[u^j(\bar{a}(\tilde{\theta}^i, \sigma_{-i}(\tilde{\theta}^{-i})), \sigma_j(\tilde{\theta}^j))] - u^j(\bar{a}(\tilde{\theta}^i, \sigma_{-i}(\tilde{\theta}^{-i})), \sigma_j(\tilde{\theta}^j))\right]}_{=0}$$

Sketch of Proof: i.i.d. case

$N = 2$:

$$p^i(\hat{\theta}^i, \hat{\theta}^j) = \mathbb{E}[u^i(\bar{a}(\tilde{\theta}^i, \hat{\theta}^j), \tilde{\theta}^i)] - \mathbb{E}[u^i(\bar{a}(\tilde{\theta}^i, \tilde{\theta}^j), \tilde{\theta}^i)]$$

$$+\mathbb{E}[u^j(\bar{a}(\tilde{\theta}^i, \tilde{\theta}^j), \tilde{\theta}^j)] - \mathbb{E}[u^j(\bar{a}(\hat{\theta}^i, \tilde{\theta}^j), \tilde{\theta}^j)]$$

Sketch of Proof: i.i.d. case

$N = 2$:

$$p^i(\hat{\theta}^i, \hat{\theta}^j) = \mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}^i, \hat{\theta}^j), \tilde{\theta}^i)\right] - \mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}^i, \tilde{\theta}^j), \tilde{\theta}^i)\right]$$

$$+\mathbb{E}\left[u^j(\bar{a}(\tilde{\theta}^i, \tilde{\theta}^j), \tilde{\theta}^j)\right] - \mathbb{E}\left[u^j(\bar{a}(\hat{\theta}^i, \tilde{\theta}^j), \tilde{\theta}^j)\right]$$

$N > 2$: decompose

$$\mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}^i, \hat{\theta}^{-i}), \tilde{\theta}^i)\right] - \mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}^i, \tilde{\theta}^{-i}), \tilde{\theta}^i)\right]$$

into (e.g. $-i = \{j, k\}$)

Sketch of Proof: i.i.d. case

$N = 2$:

$$p^i(\hat{\theta}^i, \hat{\theta}^j) = \mathbb{E}[u^i(\bar{a}(\tilde{\theta}^i, \hat{\theta}^j), \tilde{\theta}^i)] - \mathbb{E}[u^i(\bar{a}(\tilde{\theta}^i, \tilde{\theta}^j), \tilde{\theta}^i)]$$

$$+\mathbb{E}[u^j(\bar{a}(\tilde{\theta}^i, \tilde{\theta}^j), \tilde{\theta}^j)] - \mathbb{E}[u^j(\bar{a}(\hat{\theta}^i, \tilde{\theta}^j), \tilde{\theta}^j)]$$

$N > 2$: decompose

$$\mathbb{E}[u^i(\bar{a}(\tilde{\theta}^i, \hat{\theta}^{-i}), \tilde{\theta}^i)] - \mathbb{E}[u^i(\bar{a}(\tilde{\theta}^i, \tilde{\theta}^{-i}), \tilde{\theta}^i)]$$

into (e.g. $-i = \{j, k\}$)

$$\underbrace{\mathbb{E}[u^i(\bar{a}(\tilde{\theta}^i, \hat{\theta}^j, \hat{\theta}^k), \tilde{\theta}^i)] - \mathbb{E}[u^i(\bar{a}(\tilde{\theta}^i, \hat{\theta}^j, \tilde{\theta}^k), \tilde{\theta}^i)]}_{i \to k}$$

$$+\underbrace{\mathbb{E}[u^i(\bar{a}(\tilde{\theta}^i, \hat{\theta}^j, \tilde{\theta}^k), \tilde{\theta}^i)] - \mathbb{E}[u^i(\bar{a}(\tilde{\theta}^i, \tilde{\theta}^{-i}), \tilde{\theta}^i)]}_{i \to j}$$

Sketch of Proof: I

Proposition 1

$\bar{a}$ incentive efficient $\Rightarrow$ $\exists$ BB transfer p s.t. $(\bar{a}, p)$ is IC & collusion-proof.

Define

$$V_{\bar{a}}^i = \mathbb{E}\left[\sum_{t\geq 1} \delta^{t-1} u^i(\bar{a}(\tilde{\theta}_t), \tilde{\theta}_t^i)\right]$$

$$V_{\bar{a}}^i(\theta_t) = u^i(\bar{a}(\theta_t), \theta_t^i) + \delta \mathbb{E}\left[V^i(\tilde{\theta}_{t+1}; \bar{a}) | \bar{a}(\theta_t), \theta_t\right]$$

Aim: For each i, agent i can guarantee an ex ante expected payoff $V_{\bar{a}}^i + \kappa_i$ by truthtelling, regardless of others' strategies, where $\sum_i \kappa_i = 0$.

The sum of "minmax" payoffs is at least $\sum_i V_{\bar{a}}^i$ in G_M.

Sketch of Proof: II

Fix any order of agents, wlog, $1 \succ \cdots \succ N$.

Sketch of Proof: II

Fix any order of agents, wlog, $1 \succ \cdots \succ N$.

The change in agent j's expected continuation payoff caused by agent i's report:

$$\psi_t^{ij}(\hat{\theta}_t, \hat{\theta}_{t-1}) = \mathbb{E}\left[V_{\bar{a}}^j(\hat{\theta}_t^1, \ldots, \hat{\theta}_t^{i-1}, \hat{\theta}_t^i, \tilde{\theta}_t^{i+1}, \ldots, \tilde{\theta}_t^N) | \bar{a}(\hat{\theta}_{t-1}), \hat{\theta}_{t-1}\right]$$
$$- \mathbb{E}\left[V_{\bar{a}}^j(\hat{\theta}_t^1, \ldots, \hat{\theta}_t^{i-1}, \tilde{\theta}_t^i, \tilde{\theta}_t^{i+1}, \ldots, \tilde{\theta}_t^N) | \bar{a}(\hat{\theta}_{t-1}), \hat{\theta}_{t-1}\right]$$

Sketch of Proof: II

Fix any order of agents, wlog, $1 \succ \cdots \succ N$.

The change in agent j's expected continuation payoff caused by agent i's report:

$$\psi_t^{ij}(\hat{\theta}_t, \hat{\theta}_{t-1}) = \mathbb{E}\left[V_{\bar{a}}^j(\hat{\theta}_t^1, \ldots, \hat{\theta}_t^{i-1}, \hat{\theta}_t^i, \tilde{\theta}_t^{i+1}, \ldots, \tilde{\theta}_t^N) | \bar{a}(\hat{\theta}_{t-1}), \hat{\theta}_{t-1}\right]$$
$$- \mathbb{E}\left[V_{\bar{a}}^j(\hat{\theta}_t^1, \ldots, \hat{\theta}_t^{i-1}, \tilde{\theta}_t^i, \tilde{\theta}_t^{i+1}, \ldots, \tilde{\theta}_t^N) | \bar{a}(\hat{\theta}_{t-1}), \hat{\theta}_{t-1}\right]$$

Define the BB transfer as

$$p_t^i(\hat{\theta}_t, \hat{\theta}_{t-1}) = -\sum_{j \neq i}\left[\psi_t^{ij}(\hat{\theta}_t, \hat{\theta}_{t-1}) - \psi_t^{ji}(\hat{\theta}_t, \hat{\theta}_{t-1})\right] - \mathbf{1}_{\{t=1\}} \cdot \kappa_i.$$

Sketch of Proof: II

Fix any order of agents, wlog, $1 \succ \cdots \succ N$.

The change in agent j's expected continuation payoff caused by agent i's report:

$$\psi_t^{ij}(\hat{\theta}_t, \hat{\theta}_{t-1}) = \mathbb{E}\left[V_{\bar{a}}^j(\hat{\theta}_t^1, \ldots, \hat{\theta}_t^{i-1}, \hat{\theta}_t^i, \tilde{\theta}_t^{i+1}, \ldots, \tilde{\theta}_t^N) | \bar{a}(\hat{\theta}_{t-1}), \hat{\theta}_{t-1}\right]$$
$$- \mathbb{E}\left[V_{\bar{a}}^j(\hat{\theta}_t^1, \ldots, \hat{\theta}_t^{i-1}, \tilde{\theta}_t^i, \tilde{\theta}_t^{i+1}, \ldots, \tilde{\theta}_t^N) | \bar{a}(\hat{\theta}_{t-1}), \hat{\theta}_{t-1}\right]$$

Define the BB transfer as

$$p_t^i(\hat{\theta}_t, \hat{\theta}_{t-1}) = -\sum_{j \neq i}\left[\psi_t^{ij}(\hat{\theta}_t, \hat{\theta}_{t-1}) - \psi_t^{ji}(\hat{\theta}_t, \hat{\theta}_{t-1})\right] - \mathbf{1}_{\{t=1\}} \cdot \kappa_i.$$

$\forall t$, agent i

- pays j the change in i's expected continuation payoff caused by j's report
- is paid by j the change in j's expected continuation payoff caused by i's report

Sketch of Proof: III

For simplicity, consider a two-period & two-agent ($i \succ j$) setting.
Given $(\bar{a}, p)$, suppose agent i always reports truthfully. $\forall \sigma^j$,

- $\forall t$, j's expected payment to i, $\mathbb{E}\left[\psi_t^{ij}\right] = 0$. (independence)
- i's expected payoff from allocations ($\mathbb{E}$: prior)

$$u^i(\bar{a}(\tilde{\theta}_1^i, \sigma^j(\tilde{\theta}_1^j)), \tilde{\theta}_1^i) + \delta \mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}_2^i, \sigma^j(\tilde{\theta}_1, \tilde{\theta}_2^j)), \tilde{\theta}_2^i) | \bar{a}(\tilde{\theta}_1^i, \sigma^j(\tilde{\theta}_1^j)), \tilde{\theta}_1\right]$$

Sketch of Proof: III

For simplicity, consider a two-period & two-agent ($i \succ j$) setting.

Given $(\bar{a}, p)$, suppose agent i always reports truthfully. $\forall \sigma^j$,

- $\forall t$, j's expected payment to i, $\mathbb{E}\left[\psi_t^{ij}\right] = 0$. (independence)
- i's expected payoff from allocations ($\mathbb{E}$: prior)

$$u^i(\bar{a}(\tilde{\theta}_1^i, \sigma^j(\tilde{\theta}_1^j)), \tilde{\theta}_1^i) + \delta\mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}_2^i, \sigma^j(\tilde{\theta}_1, \tilde{\theta}_2^j)), \tilde{\theta}_2^i) | \bar{a}(\tilde{\theta}_1^i, \sigma^j(\tilde{\theta}_1^j)), \tilde{\theta}_1\right]$$

In $t = 2$, i pays j $(\delta\mathbb{E}\left[\cdot | \bar{a}(\tilde{\theta}_1^i, \sigma^j(\tilde{\theta}_1^j)), \tilde{\theta}_1\right])$

$$-u^i(\bar{a}(\tilde{\theta}_2^i, \sigma^j(\tilde{\theta}_1, \tilde{\theta}_2^j)), \tilde{\theta}_2^i) + \mathbb{E}_{\tilde{\theta}_2^j}\left[u^i(\bar{a}(\tilde{\theta}_2^i, \tilde{\theta}_2^j), \tilde{\theta}_2^i) | \bar{a}(\tilde{\theta}_1^i, \sigma^j(\tilde{\theta}_1^j)), \tilde{\theta}_1^i, \sigma^j(\tilde{\theta}_1^j)\right]$$

Sketch of Proof: III

For simplicity, consider a two-period & two-agent ($i \succ j$) setting. Given $(\bar{a}, p)$, suppose agent i always reports truthfully. $\forall \sigma^j$,

- $\forall t$, j's expected payment to i, $\mathbb{E}\left[\psi_t^{ij}\right] = 0$. (independence)
- i's expected payoff from allocations ($\mathbb{E}$: prior)

$$u^i(\bar{a}(\tilde{\theta}_1^i, \sigma^j(\tilde{\theta}_1^j)), \tilde{\theta}_1^i) + \delta \mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}_2^i, \sigma^j(\tilde{\theta}_1, \tilde{\theta}_2^j)), \tilde{\theta}_2^i) | \bar{a}(\tilde{\theta}_1^i, \sigma^j(\tilde{\theta}_1^j)), \tilde{\theta}_1\right]$$

In $t = 2$, i pays j $(\delta \mathbb{E}\left[\cdot | \bar{a}(\tilde{\theta}_1^i, \sigma^j(\tilde{\theta}_1^j)), \tilde{\theta}_1\right])$

$$-u^i(\bar{a}(\tilde{\theta}_2^i, \sigma^j(\tilde{\theta}_1, \tilde{\theta}_2^j)), \tilde{\theta}_2^i) + \mathbb{E}_{\tilde{\theta}_2^j}\left[u^i(\bar{a}(\tilde{\theta}_2^i, \tilde{\theta}_2^j), \tilde{\theta}_2^i) | \bar{a}(\tilde{\theta}_1^i, \sigma^j(\tilde{\theta}_1^j)), \tilde{\theta}_1^i, \sigma^j(\tilde{\theta}_1^j)\right]$$

In $t = 1$, i pays j ($\mathbb{E}$: prior)

$$-u^i(\bar{a}(\tilde{\theta}_1^i, \sigma^j(\tilde{\theta}_1^j)), \tilde{\theta}_1^i) - \delta \mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}_2^i, \tilde{\theta}_2^j), \tilde{\theta}_2^i) | \bar{a}(\tilde{\theta}_1^i, \sigma^j(\tilde{\theta}_1^j)), \tilde{\theta}_1^i, \sigma^j(\tilde{\theta}_1^j)\right]$$

$$+\mathbb{E}_{\tilde{\theta}_1^j}\left[u^i(\bar{a}(\tilde{\theta}_1^i, \tilde{\theta}_1^j), \tilde{\theta}_1^i) + \delta \mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}_2^i, \tilde{\theta}_2^j), \tilde{\theta}_2^i) | \bar{a}(\tilde{\theta}_1^i, \tilde{\theta}_1^j), \tilde{\theta}_1^i, \tilde{\theta}_1^j\right]\right]$$

Sketch of Proof: IV

Summing up and canceling terms, i's ex ante expected payoff is

$$\mathbb{E}_{\tilde{\theta}_1}\left[\mathbb{E}_{\tilde{\theta}_1^j}\left[u^i(\bar{a}(\tilde{\theta}_1^i,\tilde{\theta}_1^j),\tilde{\theta}_1^i)+\delta\mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}_2^i,\tilde{\theta}_2^j),\tilde{\theta}_2^i)|\bar{a}(\tilde{\theta}_1^i,\tilde{\theta}_1^j),\tilde{\theta}_1^i,\tilde{\theta}_1^j\right]\right]\right]$$

$$= V_{\bar{a}}^i.$$

Similarly, j's ex ante expected payoff is $V_{\bar{a}}^j$ (despite $i \succ j$).

By incentive efficiency of $\bar{a}$, the sum is $V_{\bar{a}}^i + V_{\bar{a}}^j$ the maximum ex ante expected payoff.

Sketch of Proof: IV

Summing up and canceling terms, i's ex ante expected payoff is

$$\mathbb{E}_{\tilde{\theta}_1}\left[\mathbb{E}_{\tilde{\theta}_1^j}\left[u^i(\bar{a}(\tilde{\theta}_1^i,\tilde{\theta}_1^j),\tilde{\theta}_1^i)+\delta\mathbb{E}\left[u^i(\bar{a}(\tilde{\theta}_2^i,\tilde{\theta}_2^j),\tilde{\theta}_2^i)|\bar{a}(\tilde{\theta}_1^i,\tilde{\theta}_1^j),\tilde{\theta}_1^i,\tilde{\theta}_1^j\right]\right]\right]$$

$$= V_{\bar{a}}^i.$$

Similarly, j's ex ante expected payoff is $V_{\bar{a}}^j$ (despite $i \succ j$).

By incentive efficiency of $\bar{a}$, the sum is $V_{\bar{a}}^i + V_{\bar{a}}^j$ the maximum ex ante expected payoff.

- The argument extends to arbitrary T and N.
- When $N \geq 3$, possible joint deviations.
- The order $\succ$ takes care of this possibility.
- The order can be history-dependent too.
- $(\bar{a}, p)$ is also IC. (similar to Athey & Segal (2013))

Results

Corollary 1

If a mechanism $(\bar{a}, p)$ is incentive efficient, then $\exists$ another collusion-proof & IC mechanism $(\bar{a}, q)$.

Results

Corollary 1

If a mechanism $(\bar{a}, p)$ is incentive efficient, then $\exists$ another collusion-proof & IC mechanism $(\bar{a}, q)$.

Proposition 2

Suppose μ^i is ergodic under any allocation rule. If $\bar{a}$ is incentive efficient & strict IR_0 under null transfers, then $\exists \bar{\delta} \in (0, 1)$ s.t. $\forall \delta \in (\bar{\delta}, 1)$, $\exists$ a BB transfer p such that the mechanism $(\bar{a}, p)$ is collusion-proof, IC & IR.

Adding IR

- If $(\bar{a}, \mathbf{0})$ is strictly IR_0, so is $(\bar{a}, p)$.
- Under ergodicity and patience, private information in any period has a vanishing impact on total expected payoffs, which implies $(\bar{a}, p)$ is IR.

Results

collusion-proofness $\Rightarrow$ incentive efficiency

Proposition 3

Suppose μ^i is ergodic under any allocation rule. $\forall$ IC & BB mechanism $(\bar{a}, p)$ where $\bar{a}$ is not incentive efficient, $\exists \bar{\delta} \in (0,1)$ s.t. $\forall \delta \in (\bar{\delta}, 1)$, $(\bar{a}, p)$ is not collusion-proof.

Conversely

Suppose $\bar{a}$ is not incentive efficient.

- If $M = (\bar{a}, q)$ is IC and BB, under ergodicity and patience, construct a collusive (and efficient) equilibrium in G_M.
- Similar construction if $M = (\bar{a}, q)$ is IC and IR.

Extensions

- Optimal collusion-proof dynamic mechanisms.
- Property rights in the presence of collusion.
- Collusion with limited transfers.
- Collusion with correlated information.

Related Literature

- Collusion-proof static mechanisms: Laffont and Martimort (1997, 2000), Che and Kim (2006), Safronov (2017)
- Efficient dynamic mechanisms: Bergemann and Välimäki (2010), Athey and Segal (2013), Skrzypacz and Toikka (2015)
- Optimal dynamic mechanisms: Pavan, Segal and Toikka (2014), Pavan (2016), Bergemann and Välimäki (2017)
- Collusion with persistence private info.: Athey and Bagwell (2001, 2008), Miller (2012)
- Repeated implementation: **Jackson and Sonnenschein (2007)**, Lee and Sabourian (2009, 2013), Renou and Mezzetti (2017), Renou and Tomala (2015), Chassang and Ortner (2015)