Coalition-Proof Disclosure

Germán Gieczewski **Maria "Masha" Titova**

Princeton University Vanderbilt University

QMUL-City Workshop in Economic Theory | June 2024

Introduction: Verifiable Disclosure

two players: sender (S, she/her) and receiver (R, he/him)

Sender learns his type $\theta \in \Theta$	$\rightarrow$	Sender chooses message $m \in M(\theta)$	$\rightarrow$	Receiver observes m, takes action $a \in A$	$\rightarrow$	Payoffs $u_S(a,\theta), u_R(a,\theta)$ are realized

message mapping $M : \Theta \to 2^{\mathcal{M}} \setminus \varnothing$

- $\mathcal{M}$ is some "grand" message space

our substantive assumption: sender's preferences are type-independent

Existing Models of Verifiable Disclosure

- persuasion games (Milgrom, 1981; Grossman, 1981)
  - $M(\theta)$ is all subsets of Θ that contain θ
- Dye, 1985 evidence
  - $\Theta = \{\varnothing, \theta_1, \ldots, \theta_N\}$ and $M(\theta) = \{\varnothing, \theta\}$
- evidence as partial order on the set of types (Hart, Kremer & Perry, 2017; Ben-Porath, Dekel & Lipman, 2019)
- cheap talk (Crawford & Sobel, 1982)
  - $M(\theta) = \mathcal{M}$ for all $\theta \in \Theta$

issue: multiplicity of equilibria $\rightarrow$ literature focuses on receiver-optimal equilibrium (see also Glazer & Rubinstein 2004, 2006; Sher, 2011; Rappoport, 2022)

What We Do

We show that all PBE strategies belong to a particular class

- partitions into coalitions

We focus on **coalition-proof** equilibria

- rule out blocking coalitions of senders
  - statements made by subsets of sender types which, if credible, lead to a strictly higher than equilibrium payoff for those types
  - related to neologism proofness (Farrell, 1986) and announcement proofness (Matthews, Okuno-Fujiwara & Postlewaite, 1991) for cheap talk games

Our Contributions

Coalition-Proof Equilibria:

- algorithmic characterization
- theorems for existence
- geometric characterization (for rich enough message space)
  - inspired by and comparable to *concave closure* (Kamenica & Gentzkow, 2011) and *quasiconcave envelope* (Lipnowski & Ravid, 2020)
  - it's a tent

Model

Assumptions

[A1] — finite type space $\Theta := \{\theta_1, \dots, \theta_n\}$

[A2] — common prior $\mu^0 := (\mu^0_1, \dots, \mu^0_n) \in \Delta\Theta$

Belief-Based Approach: given R's posterior belief $\mu \in \Delta\Theta$,

- $v(\mu) := u_S(a^*(\mu))$ is Sender's payoff
  - [A3] — $v(\mu)$ is upper semicontinuous

Notation

- $\boldsymbol{M^{-1}(X)}$ is set of types with access to at least one message in $X \subseteq \mathcal{M}$
  - $M^{-1}(X) := \{\theta \in \Theta \mid M(\theta) \cap X \neq \varnothing\}$

- $\boldsymbol{\mu_C^0} \in \Delta\Theta$ is prior belief conditional $\theta \in C \subseteq \Theta$ and no other information

$$\mu_C^0(\theta) = \frac{\mu^0(\theta)}{\sum_{\theta' \in C} \mu^0(\theta')} \cdot \mathbb{1}(\theta \in C)$$

- a **restricted game** with non-empty type space $C \subseteq \Theta$ has prior μ_C^0 and message mapping $M|_C$

Coalitions and Partitions

Partition into Coalitions

We focus on a particular class of **partition** strategies

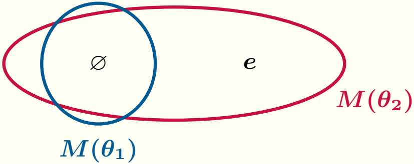
- they partition type space Θ into **coalitions**
- within a coalition, types send the same messages and attain the same payoff

Coalition

Definition: A coalition is a quadruple (C, X, σ, w), where

1. $C \subseteq \Theta$ is a non-empty set of types
2. $X \subseteq \mathcal{M}$ is a set of messages such that $M^{-1}(X) = C$
3. $\sigma : C \to \Delta\mathcal{M}$ such that $\operatorname{supp} \sigma(\cdot \mid \theta) \subseteq X \cap M(\theta)$ and $\bigcup_{\theta \in \Theta} \operatorname{supp} \sigma(\cdot \mid \theta) = X$ is sender's strategy for types in C
   - all types in C only send messages from X and all $m \in X$ are "on path"
   - does not specify what types outside C do but they do not have access to messages in X
4. $w := v(\mu(\cdot \mid m))$ for all $m \in X$ is Sender's payoff when $\theta \in C$
   - $\mu(\cdot \mid m)$ is Receiver's posterior "on path" (calculated via Bayes rule)
   - Sender's payoff is the same for all $m \in X$ and $\theta \in C$

Coalitions: Example



θ_1 *does not have evidence*, $M(\theta_1) = \varnothing$

θ_2 *has evidence*, $M(\theta_2) = \{\varnothing, \boldsymbol{e}\}$

$\underbrace{\boldsymbol{C} = M^{-1}(\boldsymbol{X})}$ $\underbrace{\boldsymbol{\sigma} \text{ such that all } \theta \in C \text{ get } \boldsymbol{w}}$

$\boldsymbol{C}$ (types)	$\boldsymbol{X}$ (messages)	$\boldsymbol{\sigma}$ (strategy)	$\boldsymbol{w}$ (payoff)
$\{\theta_2\}$	$\{\boldsymbol{e}\}$	$\sigma(\boldsymbol{e} \mid \theta_2) = 1$	$v(\mu^0_{\{\theta_2\}})$
$\{\theta_1, \theta_2\}$	$\{\varnothing\}$	$\sigma(\varnothing \mid \theta_1, \theta_2) = 1$	$v(\mu^0)$
$\{\theta_1, \theta_2\}$	$\{\varnothing, \boldsymbol{e}\}$	$\sigma(\varnothing \mid \theta_1) = 1$ $\sigma(\varnothing \mid \theta_2) = \alpha$	$v(\mu(\cdot \mid \varnothing))$ $= v(\mu(\cdot \mid \boldsymbol{e}))$

Partition

Definition: A collection $\left\{(C_t, X_t, \sigma_t, w_t)\right\}_{t=1}^{T}$ is a **partition** if it can be obtained from the following algorithm:

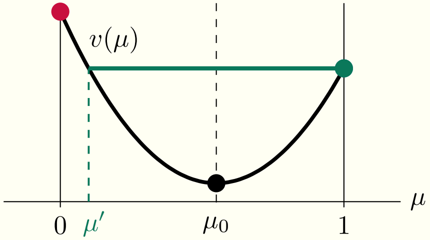
Algorithm 1: Partition Algorithm

Let $t := 1$ and $\Theta_1 := \Theta$
while $\Theta_t \neq \varnothing$
    let $(C_t, X_t, \sigma_t, w_t)$ be a coalition in restricted game with type space Θ_t
    let $\Theta_{t+1} := \Theta_t \setminus C_t$ and $t := t + 1$
end

- algorithm terminates in $T \leq |\Theta|$ steps since each C_t is non-empty
- we say that σ such that $\sigma|_{C_t} = \sigma_t$ for each t is the partition strategy associated with $\{(C_t, X_t, \sigma_t, w_t)\}_{t=1}^{T}$

Partition: Example

$\Theta = \{\theta_1, \theta_2\} \qquad M(\theta_1) = \{\varnothing\} \quad M(\theta_2) = \{\varnothing, \boldsymbol{e}\} \qquad \mu = Pr(\theta = \theta_2)$



$\{\theta_2\}$	$\{\boldsymbol{e}\}$	$\sigma_1(\boldsymbol{e} \mid \theta_2) = 1$	$v(1)$

$\{\theta_1\}$	$\{\varnothing\}$	$\sigma_2(\varnothing \mid \theta_1) = 1$	$v(0)$

$\{\theta_1, \theta_2\}$	$\{\varnothing\}$	$\sigma_1(\varnothing \mid \theta_1, \theta_2) = 1$	$v(\mu_0)$

$\{\theta_1, \theta_2\}$	$\{\varnothing, \boldsymbol{e}\}$	$\sigma_1(a \mid \theta_1) = 1,\ \sigma_1(\varnothing \mid \theta_2) = \alpha$	$v(1)$

PBE Partitions and Individual Rationality

PBE Strategy

Definition: a Sender's strategy $\sigma : \Theta \to \Delta\mathcal{M}$ is a PBE strategy if there exists a Receiver's belief system $\mu : \mathcal{M} \to \Delta\Theta$ such that

- $\forall \theta \in \Theta$, $\sigma(\cdot \mid \theta)$ is supported on $\arg\max_{m \in M(\theta)} v(\mu(\cdot \mid m))$
- μ is obtained from μ^0, given m, using Bayes' rule, for all m used with positive probability on equilibrium path

Proposition 1

σ is PBE strategy $\iff$ σ is associated with partition $\{(C_t, X_t, \sigma_t, w_t)\}_{t=1}^{T}$ such that $w_1 \geq \ldots \geq w_T$ and

$$(\mathbf{IR})\ w_t \geq \underline{v}(\theta) := \max_{m \in M(\theta)} \min_{\mu(\cdot|m) \text{ feasible}} v(\mu(\cdot|m)) \quad \text{for all } t \text{ and } \theta \in C_t$$

Coalition Proofness

Coalition Proofness

Definition: Let σ be a strategy associated with partition $\{(C_t, X_t, \sigma_t, w_t)\}_{t=1}^{T}$

- $(\widetilde{C}, \widetilde{X}, \widetilde{\sigma}, \widetilde{w})$ is a blocking coalition of σ if it is a coalition of the restricted game with type space $\bigcup_{t: w_t < \widetilde{w}} C_t$
- σ is coalition-proof if there are no blocking coalitions

coalition-proof strategies rule out *coalitional deviations*

- types in $\widetilde{C}$ announce that they switch to strategy $\widetilde{\sigma} : \widetilde{C} \to \Delta \widetilde{X}$
- if R believes this announcement, types in $\widetilde{C}$ receive $\widetilde{w}$
- types in $\widetilde{C}$ must be exactly those who have access to at least one message in $\widetilde{X}$ and benefit from the deviation

Algorithm 2: Greedy Partition Algorithm

Let $t := 1$ and $\Theta_1 := \Theta$

while $\Theta_t \neq \varnothing$

let W_t be the set of payoffs attainable by coalitions of the restricted game with type space Θ_t;

let $(C_t, X_t, \sigma_t, w_t)$ be a coalition s.t. $w_t = \max(W_t \cap [\max_{\theta \in \Theta_t} \underline{v}(\theta), w_{t-1}])$;

let $\Theta_{t+1} := \Theta_t \setminus C_t$ and $t := t + 1$;

end

Proposition 2

$\{(C_t, X_t, \sigma_t, w_t)\}_{t=1}^T$ is a greedy partition $\iff$ the associated strategy is **PBE and coalition-proof**.

Theorems for Existence of Coalition-Proof PBE

I. Quasiconcave v and Complete M

Definition: Message mapping M is complete if for any two messages $m, m' \in \mathcal{M}$, there exists $m'' \in \mathcal{M}$ such that

$$M^{-1}(\{m''\}) = M^{-1}(\{m\}) \cup M^{-1}(\{m'\})$$

- if there is a way to say "my type is in A" (message m)
- and a way to say "my type is in B" (m')
- then there is a way to say "my type is in A or B" (m'')

I: Quasiconcave v and Complete M

Theorem 1

If v is quasiconcave and M is complete, then there exists a coalition-proof PBE. If, in addition,

- v is strictly quasiconcave, then Greedy Algorithm always terminates
- v is generic (such that $v(\mu^0_C) = v(\mu^0_{C'})$ only if $C = C'$), then all coalition-proof PBE are payoff-equivalent

proof sketch:

- show (by contradiction) that at each step of Greedy Algorithm, $\max W_t \leq w_{t-1}$ — otherwise, there exists a coalition of types in $C_{t-1} \cup C_t$ that pays more than w_{t-1}

II. Betweenness of v

v satisfies betweenness if it is both quasiconcave and quasiconvex

Theorem 2

If v satisfies betweenness, then there exists a coalition-proof PBE. If, in addition,

- v satisfies strict betweenness, then Greedy Algorithm always terminates
- v is generic (such that $v(\mu_C^0) = v(\mu_{C'}^0)$ only if $C = C'$), then all coalition-proof PBE are payoff-equivalent

III: Adding Cheap Talk

Definition: Message mapping M satisfies cheap talk property if for each message $m \in \mathcal{M}$ there are at least n messages $m' \in \mathcal{M}$ (including m) such that $M^{-1}(\{m'\}) = M^{-1}(\{m\})$.

- any message mapping can be augmented with a second dimension so that Sender has access to verifiable information and cheap talk
- employ Lipnowski & Ravid (2020) — cheap talk quasiconcavifies v

Theorem 3

If M is complete and satisfies cheap talk property, then there exists a coalition-proof PBE.

Geometric Characterization: Rich Message Space

Rich Message Space

Motivation: let us find the "most" coalition-proof partition when message space is maximally rich

- for each vector of probabilities $(p_1, \ldots, p_n)$, there is a message m only accessible to fractions $p_1, \ldots, p_n$ of type $\theta_1, \ldots, \theta_n$ respectively

Result: a geometric characterization (tent) of coalition-proof PBE

- comparable to *concave closure* (Kamenica & Gentzkow, 2011) and *quasiconcave envelope* (Lipnowski & Ravid, 2020)

Rich Message Space: Coalition-Proof PBE

Let $\mu^{*1} = \arg\max_{\mu \in \Delta\Theta} v(\mu)$ (assumed unique)

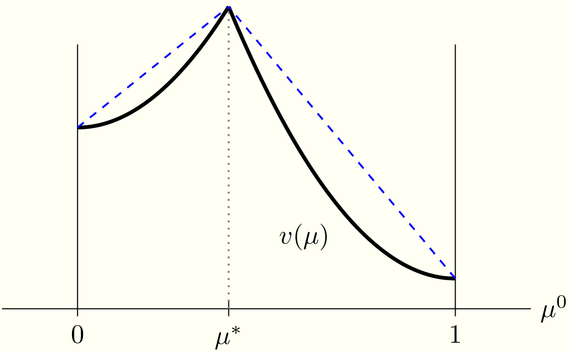
- payoff $v(\mu^{*1})$ must be attained in first coalition
- some type θ_i attains it with probability 1

Let μ^{*2} be the argmax (assumed unique) of $v(\mu)$ s.t. $\mu(\theta_i) = 0$

- payoff $v(\mu^{*2})$ is attained in second coalition
- some type θ_j attains it with probability 1

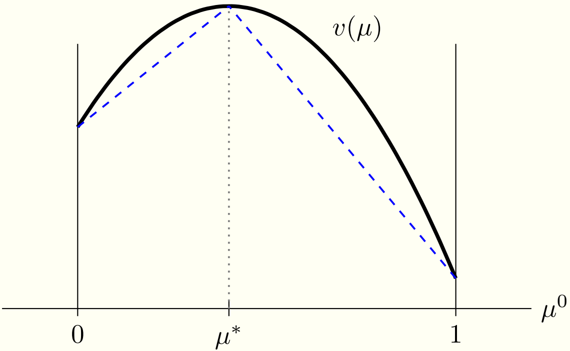
and so on

Tent: Two Types



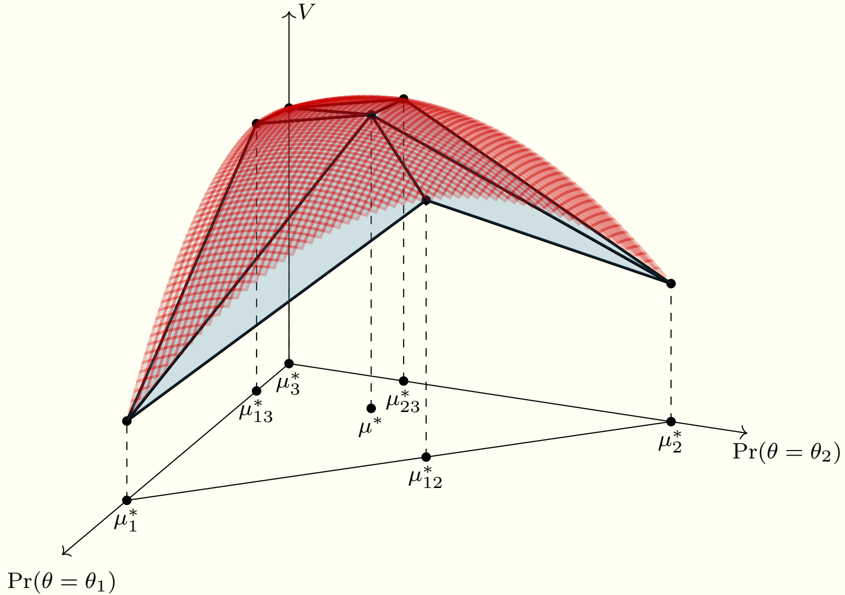
in coalition-proof PBE for this v, Sender does as well as under Bayesian Persuasion and better than under Cheap Talk

Tent: Two Types



in coalition-proof PBE for this v, Sender does worse than under Cheap Talk and BP

Tent: Three Types



Conclusion

We show that all PBE strategies in verifiable disclosure games belong to a certain class

- partitions into coalitions

We focus on coalition-proof PBEs

- algorithmic characterization
- theorems for existence
- geometric characterization when message space is rich

Thank You!