Constructive Vs Toxic Argumentation in Debates

Andriy Zapechelnyuk | Tymofiy Mylovanov

University of Edinburgh KSE and Pittsburgh Universtity

2023

1 Motivation

Two key ingredients of debates:

• **Disclosure** (constructive argumentation) – informing the audience in a controlled way

• **Obfuscation** (toxic argumentation) – reducing the impact the opponent's arguments on the audience

Question: How and to what extent the nature of arguments affects truth discovery by the audience?

Answering this question will allow us to contribute to the policy discussion of whether debates should adhere to the principle of freedom of speech or whether they should be moderated

2 Model in Nutshell

Two debaters

- sequentially choose information disclosure strategies

- of an uncertain state of the world

- to influence the choice of a heterogeneous audience.

We compare two cases:

- sequential disclosure: second mover reveals additional information about the state

- sequential obfuscation: second mover obfuscates the information revealed by the first mover

3 Why is This a Puzzle?

Disclosure and obfuscation (garbling) are seen as synonyms in the literature.

"Two senders sequentially **obfuscate** an **initially revealed** state of the world."

is symmetric to

"Two senders sequentially **disclose** information about **initially hidden** state of the world."

There is no difference when there is only one sender.

What difference does it makes to the strategic interaction of two senders?

4 Model

- Two parties, A and R, are engaged in a debate on some issue
- The truth about the issue is summarized by state $\theta \in \Theta = [0, 1]$
- The public is a continuum of citizens indexed by type $t \in T = [0, 1]$

  - θ and t are independent, with distributions F and G.

- Each citizen chooses whether to support party A or party R.
- Given a posterior expected state x, the utility of citizen with type t is

$$\begin{cases} x - t & \text{if supports } A, \\ 0 & \text{if supports } R. \end{cases}$$

- $G(x)$ and $1 - G(x)$ are total supports of parties A and R

5 Preferences and Strategies

- Parties are expected utility maximizers
  - Let q_i be the support of party $i \in \{A, R\}$, so $q_A + q_R = 1$
  - Utility $u_i(q_i)$, twice continuously differentiable and strictly increasing in q_i
- Let M_i be a set of messages of party i.
  - M_A and M_R are rich enough: $\Theta \subseteq M_A$ and $\Theta \times M_A \subseteq M_R$
- Strategy of party A: conditional distribution $\phi_A(\cdot|\theta)$ over M_A.
- Strategy of party R: conditional distribution $\phi_R(\cdot|\theta, m_A)$ over M_R.

6 Timing

1. Parties A and R choose their strategies sequentially

2. State θ realizes

3. Message m_A is generated according to party A's strategy

4. Message m_R is generated according to party R's strategy.

5. Citizens observe their types, strategies of the parties, and message m_R of party R (but **not** message m_A of party A!)

6. Citizens derive the posterior state x and choose which party to support.

7 Distribution of of Posterior Expected State

• Parties' utilities depend only on posterior expected state x conditional on some message m

• Distribution of posterior expected state summarizes information for citizens

• Given a pair of strategies (ϕ_A, ϕ_R), let

- $\mu_A(\phi_A) \in \Delta(\Theta)$

- $\mu_R(\phi_A, \phi_R) \in \Delta(\Theta)$

be the distributions of the expected state induced by observation of messages of party A and R, respectively

• compare distributions by their Blackwell informativeness for the citizens

- μ' is *more informative* than μ'', denoted by $\mu' \succeq \mu''$, if μ' is a m.p.s. of μ''.

8 Sequential Disclosure and Sequential Obfuscation

• Sequential disclosure: Party R reveals information in addition to what has been revealed by party A's message m_A.

- I.e., m_R contains finer the information than m_A

- Formally, party R's strategy ϕ_R must satisfy

$$F \succeq \mu_R(\phi_A, \phi_R) \succeq \mu_A(\phi_A).$$

• Sequential obfuscation: Party R obfuscates (or garbles) information revealed by party A's message m_A.

- I.e., m_R contains coarser information than m_A.

- Formally, party R's strategy ϕ_R must satisfy

$$\mu_A(\phi_A) \succeq \mu_R(\phi_A, \phi_R).$$

9 Outcomes

• Given strategies (ϕ_A, ϕ_R), the induced outcome is the distribution of posterior expected state induced by messages of party R,

$$\mu = \mu_R(\phi_A, \phi_R)$$

• Let $V_A(x) = u_A(G(x))$ and $V_R(x) = u_R(1 - G(x))$.

• Let $V_i(\mu) = E_\mu[V_i(x)]$ expected payoff under μ

• Let $\mathcal{M}$ be the set of feasible outcomes:

$$\mathcal{M} = \{\mu \in \Delta(\Theta) : F \succeq \mu\}.$$

• Outcome $\mu \in \mathcal{M}$ is unimprovable by disclosure for party R if

$$V_R(\mu) \geq V_R(\mu') \text{ for all } \mu' \in \mathcal{M} \text{ such that } \mu' \succeq \mu.$$

• Outcome $\mu \in \mathcal{M}$ is unimprovable by obfuscation for party R if

$$V_R(\mu) \geq V_R(\mu') \text{ for all } \mu' \in \mathcal{M} \text{ such that } \mu' \preceq \mu.$$

10 Simplifying the Problem

• $\mathcal{M}_R^D$ and $\mathcal{M}_R^O$ be the sets of feasible outcomes that are unimprovable by disclosure and obfuscation, respectively, for party R.

• Consider two problems:

$$\max_{\mu \in \mathcal{M}_R^D} V_A(\mu), \tag{$\mathsf{P_D}$}$$

$$\max_{\mu \in \mathcal{M}_R^O} V_A(\mu). \tag{$\mathsf{P_O}$}$$

Observation 1. An outcome $\mu \in \Delta(\Theta)$ is an equilibrium outcome of sequential disclosure (sequential obfuscation) if and only if it is a solution of problem ($\mathsf{P_D}$) (respectively, ($\mathsf{P_O}$)).

Observation 2. For generic utility functions u_A and u_R, the games of sequential disclosure and sequential obfuscation have unique equilibrium outcomes.

11 Comparison

• Sequential disclosure restricts party A's choice to outcomes that are sufficiently **revealing** from party R's perspective

• Sequential obfuscation restricts party A's choice to outcomes that are sufficiently **unrevealing** from party R's perspective

• Party A optimizes on two essentially disjoint sets in the two problems, one clearly favoring more information disclosure than the other.

Proposition 1. Let μ^D and μ^O be equilibrium outcomes of sequential disclosure and sequential obfuscation, respectively, and suppose that the parties' expected utilities are not identical,

$$\left(V_A(\mu^D), V_R(\mu^D)\right) \neq \left(V_A(\mu^O), V_R(\mu^O)\right).$$

Then μ^O cannot be more informative than μ^D.

12 Full Disclosure and No Disclosure

Proposition 2. In the sequential disclosure game:

(i) If full disclosure is Pareto undominated, then it is an equilibrium outcome.

(ii) If no disclosure is Pareto dominant, then it is an equilibrium outcome. If no disclosure is not Pareto dominant, then, for generic preferences, it is not an equilibrium outcome.

Proposition 3. In the sequential obfuscation game:

(i) If no disclosure is Pareto undominated, then it is an equilibrium outcome.

(ii) If full disclosure is Pareto dominant, then it is an equilibrium outcome. If full disclosure is not Pareto dominant, then, for generic preferences, it is not an equilibrium outcome.

13 Log-concave preferences

For several results, we will assume that

- Distribution of citizens' types G has a strictly log-concave density g:

$$\begin{array}{l} G \text{ admits a continuously differentiable density } g, \text{ and} \\ \ln g(\cdot) \text{ is strictly concave.} \end{array} \qquad (A_1)$$

- Marginal utilities of both parties are log-concave:

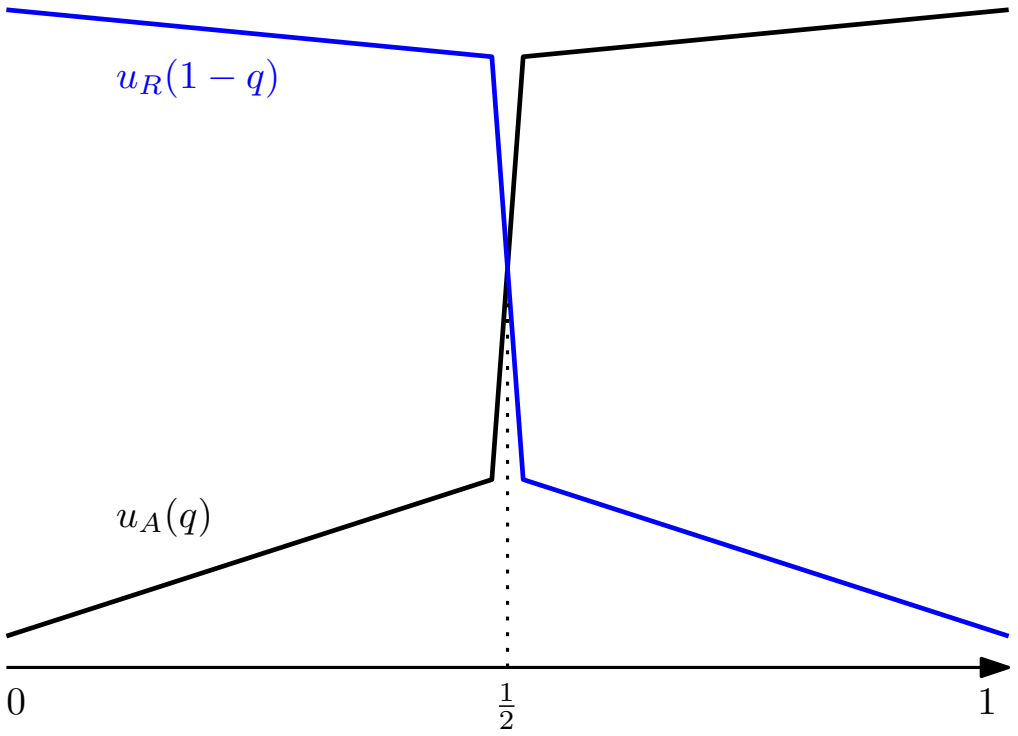
$$\ln(u'_A(\cdot)) \text{ and } \ln(u'_R(\cdot)) \text{ are concave} \qquad (A_2)$$

14 Constant marginal utility ratio

- Let $\dfrac{u'_A(q)}{u'_R(1-q)}$ be constant.
- Equivalently, $u_R(1-q) = b - cu_A(q)$ for some $b \in \mathbb{R}$ and $c > 0$.
  - zero-sum or constant-sum
  - risk neutrality

Theorem 1. Suppose that $u'_A(q)/u'_R(1-q)$ is constant. Then full disclosure (no disclosure) is an equilibrium outcome of sequential disclosure (sequential obfuscation, respectively). Moreover, these equilibrium outcomes are unique in the respective games if assumptions (A_1) and (A_2) are satisfied.

15 Decreasing marginal utility ratio.



- Let $\frac{u'_A(q)}{u'_R(1-q)}$ be decreasing (includes risk averse preferences)

- In the case of democratic regimes where the minority party stands to gain more from increasing its support than the majority party stands to lose:

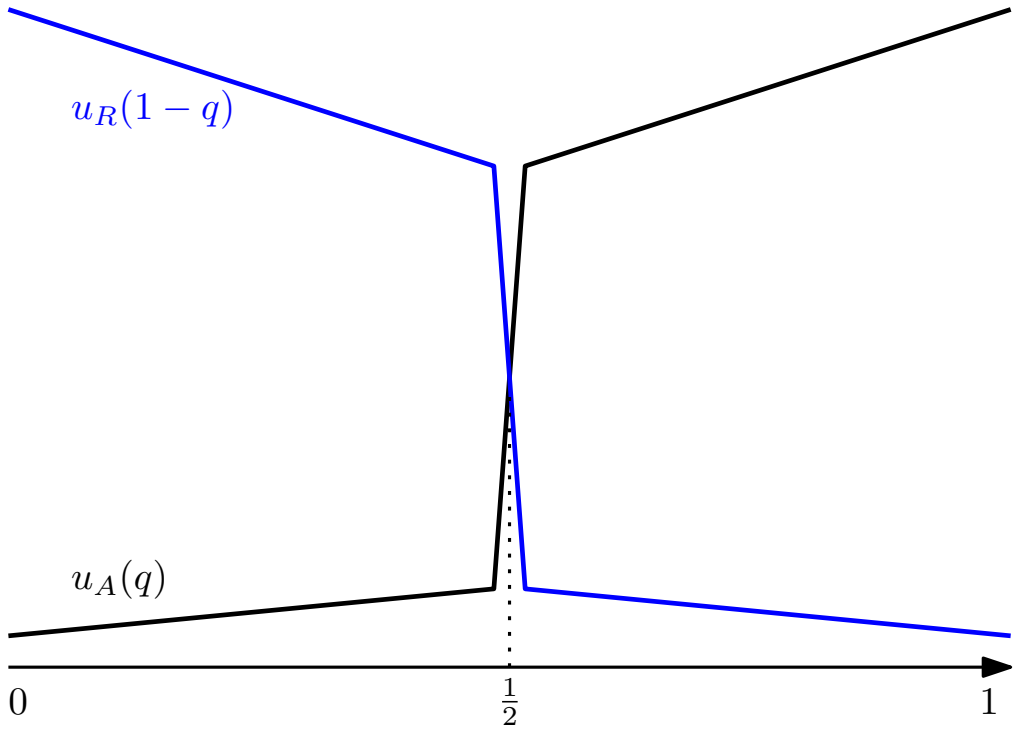
16 Decreasing marginal utility ratio.

Theorem 2. Suppose that $u'_A(q)/u'_R(1-q)$ is decreasing. Then no disclosure is an equilibrium outcome of sequential obfuscation. Moreover, it is the unique equilibrium outcome if either $u'_A(q)/u'_R(1-q)$ is strictly decreasing, or assumptions (A_1) and (A_2) are satisfied.

- toxic debates reveal nothing about the state,

- constructive debates reveal some information, except when no disclosure is Pareto dominant.

17 Increasing marginal utility ratio



- Let $\frac{u'_A(q)}{u'_R(1-q)}$ be increasing (includes risk seeking preferences)

- In the case of authoritarian regimes where the majority party stands to gain more from squashing the minority opposition than the opposition stands to lose:

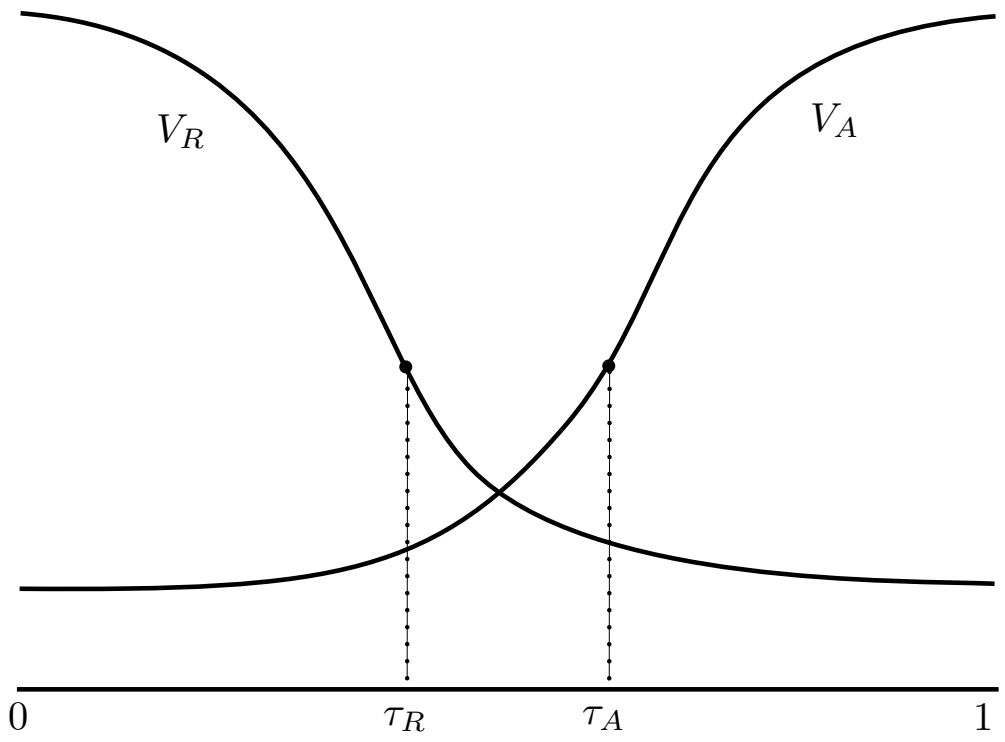
18 Increasing marginal utility ratio

Theorem 2. Suppose that $u'_A(q)/u'_R(1-q)$ is increasing and assumptions (A_1) and (A_2) are satisfied. Then full disclosure is the unique equilibrium outcome of sequential disclosure.

- constructive debates reveal the state,

- toxic debates do not, except when full disclosure is Pareto dominant.

19 Intuition



Increasing marginal utility ratio $u'_A(q)/u'_R(1-q)$

20 Assumption: Ex Ante Commitment

• Ex ante, the parties and citizens are symmetrically informed about the state

• Communication is not cheap talk

- the parties can hide or garble information, but they cannot outright lie (lies are costly to reputation)

- Bayesian persuasion results are robust to minor departures from full commitment t (Lipnowski, Ravid, and Shishkin, 2021; Guo and Shmaya, 2021; Min, 2021; Eilat and Neeman, 2023).

• The parties do not have private information when choosing their information strategies.

- contribution to information design literature

- a pooling equilibrium of the signaling game

- full disclosure and no disclosure equilibria are robust to parties' private info

21 Assumption: Sequential moves

• The parties choose their information strategies sequentially.

- plausible that the respondent chooses her strategy only after she has seen the choice of the accuser.

- full disclosure and no disclosure equilibria are robust to the sequence of moves

• Sequential moves allow us to substantially reduce the set of equilibria, and to obtain a unique equilibrium outcome for generic preferences.

- In contrast, simultaneous disclosure/obfuscation leads to a plethora of equilibria

22 Related Literature

• Competition in persuasion (simultaneous disclosure): Gentzkow and Kamenica (2017); Li and Norman (2018); Ravindran and Cui (2020)

- with private states in Boleslavsky and Cotton (2018); Au and Kawai (2020).

• Sequential disclosure:

- multiple senders: Li and Norman (2021); Wu (2021)

- two senders alternating indefinitely: Koessler, Laclau, Renault, and Tomala (2022)

• Sequential obfuscation: Lipnowski, Mathevet, and Wei (2020); Arieli, Babichenko, and Sandomirskiy (2022)

THANK YOU!