Persuasion with Verifiable Information

by

Maria (Masha) Titova

uc san diego

December 9, 2020

- persuasion games with verifiable information
  - privately informed sender
    - wants receivers to approve his proposal
    - sends verifiable messages to receivers
  - group of uninformed receivers, each choosing between
    - approving and rejecting proposal
- many applications
  - politician challenges status quo, convinces voters to elect him
  - firm convinces consumers to adopt its product
  - job market candidate convinces committee members to offer them a job

- **persuasion games with verifiable information**
  - direct implementation: can restrict attention to direct equilibria
    - sender tells each receiver what to do
  - ranking of equilibrium outcomes (ex-ante utility of sender):
    - worst: equivalent to full disclosure
    - best: commitment outcome (Kamenica and Gentzkow, 2011)
- **targeted advertising in elections**
  - challenger has positive odds of winning elections that he certainly loses with public advertising
  - more polarized electorate ⇒ higher odds of swinging elections

- **communication**:
  - **Milgrom (1981)** and **Grossman (1981)**; Crawford and Sobel (1982); Spence (1973); Kamenica and Gentzkow (2011); Alonso and Câmara (2016); Lipnowski and Ravid (2020)

    my contribution: sender reaches commitment outcome with verifiable information

- **targeted advertising in elections**:
  - Prat and Strömberg (2013); DellaVigna and Gentzkow (2010); George and Waldfogel (2006); DellaVigna and Kaplan (2007); Enikolopov et al. (2011); Oberholzer-Gee and Waldfogel (2009)

    my contribution: targeted advertising allows politicians to swing elections

MODEL WITH ONE RECEIVER

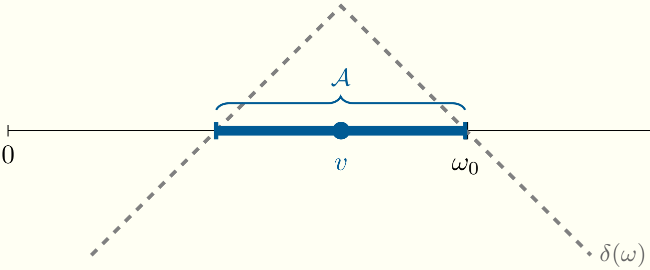
$$\Omega := [0, 1] \text{ – } \underline{\text{state space}}$$

- **sender (he)**
  - privately observes state of the world $\omega \in \Omega$
    - ω drawn from common prior $p > 0$ over Ω
  - gets 1 if receiver approves, 0 otherwise
    - state-independent preferences
  - sends verifiable message $m \in \mathbb{M}$ to receiver
    - message space $\mathbb{M} := 2^{|\Omega|}$
    - $\omega \in m$ – no lies of commission

- **receiver (she)**
  - $\delta(\omega)$ is her net payoff of approval
  - she approves in state ω if only if $\delta(\omega) \geq 0$
  - her approval set is

$$\mathcal{A} := \{\omega \in \Omega \mid \delta(\omega) \geq 0\}$$

- **receiver** has ideal position $v \in \Omega$
  - compares sender's position to status quo $\omega_0 \in (0, 1)$
  - net payoff of approval $\delta(\omega) = |v - \omega_0| - |v - \omega|$
  - approval set $\mathcal{A} = \{\omega \in \Omega \text{ s.t. } \delta(\omega) \geq 0\}$



- (Perfect Bayesian) Equilibrium (σ, a, q)
  - $\sigma : \Omega \to \Delta(\mathbb{M})$ – messaging strategy of sender
    - maximizes sender's utility $\forall \omega \in \Omega$ subject to $\omega \in m$, $\forall m \in \mathbb{M}$
  - $a : \mathbb{M} \to \{0, 1\}$ – approval strategy of receiver
    - set of approval beliefs

$$\mathcal{B} := \{q \in \Delta\Omega \mid \mathbb{E}_q[\delta(\omega)] \geq 0\}$$

    - best response $a(m) = \mathbb{1}\big(q(m) \in \mathcal{B}\big)$, $\forall m \in \mathbb{M}$
  - $q : \mathbb{M} \to \Delta\Omega$ – posterior belief of receiver
    - Bayes-rational on equilibrium path
    - $\operatorname{supp} q(m) \subseteq m$, $\forall m \in \mathbb{M}$

- in equilibrium, consider receiver's action in each state
  - let $W \subseteq \Omega$ be set of approval states
- direct implementation of W:

state	sender's message	receiver's belief	receiver's action
$\omega \in W$	W	$p(\cdot \mid W)$	approve
$\omega \in \Omega \setminus W$	$\Omega \setminus W$	$p(\cdot \mid \Omega \setminus W)$	reject

where $p(\omega \mid W) := \frac{p(\omega)}{\int_W p(\omega')d\omega'}$ is conditional prior probability

TWO CONSTRAINTS

- receiver approves after message W:

$$p(\cdot \mid W) \in \mathcal{B} \tag{obedience}$$

- sender has no incentive to deviate to full disclosure:

$$\mathcal{A} \subseteq W \tag{IC}$$

Theorem 1

The following statements about set $W \subseteq \Omega$ are equivalent:

(1) W is an equilibrium set of approval states

(2) W satisfies receiver's (obedience) and sender's (IC) constraints

(3) W is a set of approval states in direct equilibrium

- *Intuition:*
  - *(1) ⇒ (2)*: W satisfies (IC); for (obedience), implement W directly
  - *(2) ⇒ (3)*: receiver skeptical off path
  - *(3) ⇒ (1)*: direct equilibrium is an equilibrium

- **Theorem 1** allows us to restrict attention to direct equilibria
  - described by set of approval states W satisfying (obedience) and (IC)
- rank equilibria by sender's ex-ante utility
  - same as his ex-ante odds of approval
  - equals $P(W)$, measure of set of approval states under prior distribution

- sender's odds of approval are minimized across all equilibria
  - smallest (in terms of ex-ante utility) set of approval states $\underline{W}$
  - $\underline{W} = \mathcal{A}$, sender's (IC) constraint binds
- receiver makes fully informed choice
- outcome-equivalent to full disclosure
  - (Grossman, 1981); (Milgrom, 1981); (Milgrom and Roberts, 1986); reviewed by (Milgrom, 2008)

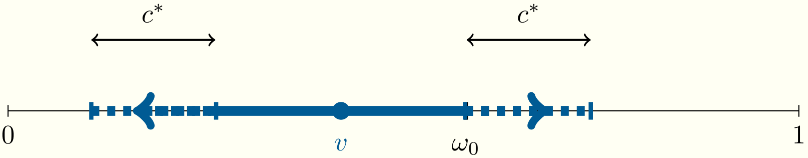
- sender's odds of approval are maximized across all equilibria
  - largest (in terms of ex-ante utility) set of approval states $\overline{W}$
  - receiver's (obedience) constraint binds

Theorem 2

Sender-preferred equilibrium outcome is a **commitment outcome**. Specifically, $\overline{W}$ is characterized by a cutoff value $c^* > 0$ such that

- sender's proposal is approved if $\delta(\omega) \geq -c^*$ and rejected if $\delta(\omega) < -c^*$
- when sender's proposal is approved, receiver's expected net payoff of approval is zero: $\mathbb{E}_p\left[\delta(\omega) \mid \overline{W}\right] = 0$

- equilibrium range
  - sender-worst equilibrium: $\underline{W} = \mathcal{A}$
  - sender-preferred equilibrium:
    - maximize $P(W)$ subject to receiver's (obedience) constraint
    - $\overline{W}$ is characterized by $c^* > 0$ that solves (obedience)



MODEL WITH MANY RECEIVERS

SETUP

$$I := \{1, \ldots, n\} \text{ – set of receivers}$$

p is common prior

- **sender**:
  - has state-independent utility $u_s : 2^I \to \mathbb{R}$
  - u_s weakly increases in every receiver's action
- **receiver** $i \in I$:
  - observes private verifiable message $m_i \in \mathbb{M}$ chosen by sender
  - solves independent problem: approves if $\omega \in \mathcal{A}_i$

Theorem 3

The following statements about the sender's ex-ante utility $\bar{u}_s$ are equivalent:

(1) $\bar{u}_s$ is reached in equilibrium

(2) $\bar{u}_s$ is reached in a direct equilibrium

(3) $\bar{u}_s$ is given by

$$\bar{u}_s = \int_\Omega u_s(\{i \in I \mid \omega \in W_i\}) \cdot p(\omega) d\omega,$$

where for every receiver $i \in I$, $W_i \subseteq \Omega$ is her set of approval states, which satisfies

- receiver's (obedience) constraint $p(\cdot \mid W_i) \in \mathcal{B}_i$
- sender's (IC) constraint $\mathcal{A}_i \subseteq W_i$

Sender-worst equilibrium is outcome-equivalent to full disclosure.

Theorem 3

Sender-preferred equilibrium outcome is a commitment outcome.

TARGETED ADVERTISING IN ELECTIONS

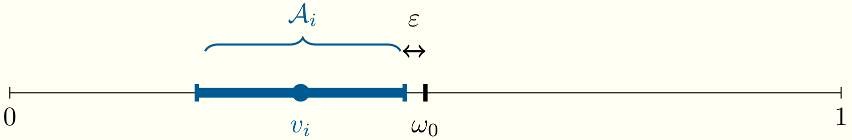
- **Targeted Advertising** was an important part of winning campaigns in recent U.S. Presidential Elections:
  - **2016 Trump**: used voter data from Cambridge Analytica
  - **2008 Obama**: first social media campaign
  - **2000 Bush**: targeting voters by mail

CAN TARGETED ADVERTISING SWING ELECTIONS? → Yes

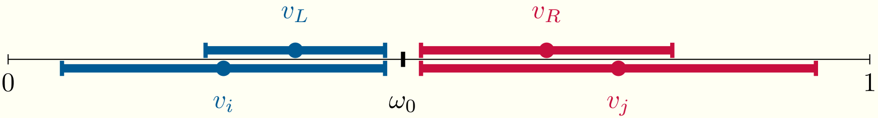
- approach: compare Targeted Advertising (**TA**) to Public Disclosure (**PD**)
  - **TA**: private messages (e.g. through Facebook)
    - application of the main model
  - **PD**: public message (e.g. debate, tweeting)
    - *common prior + common message → common posterior*

- Ω is policy space, positions range from far-left (0) to far-right (1)
- this talk: uniform prior, $p \sim U[0,1]$
- **sender**: challenger
  - receives 1 if wins election, 0 otherwise
  - any social choice rule satisfying weak monotonicity (e.g. majority, unanimity)
- focus on sender-preferred equilibrium

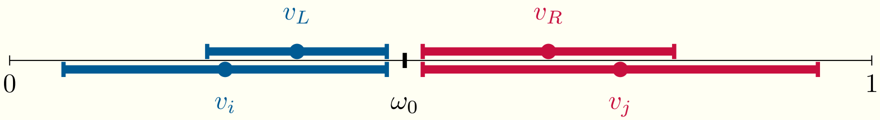
- **receivers**: sincere voters with spatial preferences:
  - ◇ approval set of voter $i \in I$ is $\mathcal{A}_i = \left\{\omega \in \Omega \text{ s.t. } |v_i - \omega| \leq |v_i - \omega_0| - \varepsilon\right\}$
    - policies that are closer to v_i than status quo by at least ε
    - $\varepsilon \in \left(0, \frac{|v_i - \omega_0|}{2}\right)$, $\forall i \in I$, is status quo bias



- $L = \arg \max_{i \in I,\ v_i < \omega_0} v_i$ is representative voter on the left
  - L is convinced $\Rightarrow$ every left voter is convinced: $\mathcal{A}_L \subseteq \mathcal{A}_i, \forall v_i < \omega_0$
- $R = \arg \min_{j \in I,\ v_j > \omega_0} v_j$ is representative voter on the right
  - R is convinced $\Rightarrow$ every right voter is convinced: $\mathcal{A}_R \subseteq \mathcal{A}_j, \forall v_j > \omega_0$



- representative voters L and R are <u>incompatible</u>
  - $\mathcal{A}_L \cap \mathcal{A}_R = \varnothing$ and $\mathcal{B}_L \cap \mathcal{B}_R = \varnothing$



- voters L and R *never* both vote for the challenger under common belief
- if voters L and R are jointly pivotal, challenger loses with probability 1

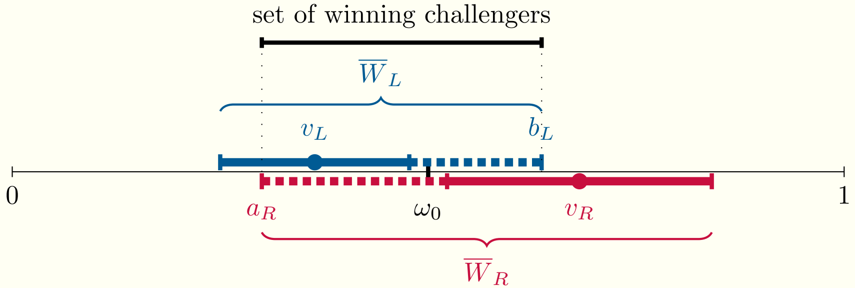
Definition

Election with representative voters L and R is <u>unwinnable</u> for the challenger under common belief, if for all $T \subseteq I$, $u_s(T) = 1$ if and only if $\{L, R\} \in T$

Theorem: Targeting in Unwinnable Elections

In the sender-preferred equilibrium of unwinnable election with representative voters L and R,

- set of winning policies is $[a_R, b_L]$ with $a_R < \omega_0 < b_L$
- challenger's odds of winning are $b_L - a_R > 0$



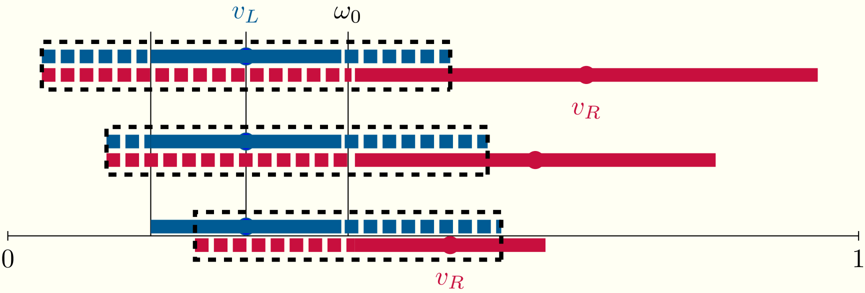
- when $v_R \uparrow$ ($v_L \downarrow$), voter R (L) becomes more persuadable
- when $v_R \uparrow$ or $v_L \downarrow$, electorate becomes more polarized

Theorem: Comparative Statics

In the sender-preferred equilibrium of an unwinnable election with representative voters L and R,

- as $v_R \uparrow$ and/or $v_L \downarrow$, challenger's odds of winning $b_L - a_R$ increase
- suppose $|v_L - \omega_0| = |v_R - \omega_0|$
  - as $v_R \uparrow$, set of winning policies shifts to the left, i.e. $a_R \downarrow$ and $b_L \uparrow$

COMPARATIVE STATICS



- I solve persuasion games with **verifiable information**
  - direct implementation: W is an equilibrium set of approval states $\iff$
    - W satisfies receiver's (obedience) and sender's (IC) constraints
  - set of equilibrium outcomes (ranked by ex-ante utility of sender):

    worst: full disclosure $\rightarrow$ best: commitment outcome

- **targeted advertising swings elections**:
  - challenger says different things to incompatible voters L and R
    - L: left + *some* right policies, left on average
    - R: right + *some* left policies, right on average
    - challenger wins if his policy is not too far from status quo
  - L and R are more polarized $\implies$ challenger wins with TA more often

Thank You!

COMMITMENT PROTOCOL

- Commitment Protocol (σ, a, q)
  - $\sigma : \Omega \to \Delta(\mathbb{M})$ – messaging strategy of sender

    maximizes sender's utility $\forall \omega \in \Omega$ subject to $\omega \in m$, $\forall m \in \mathbb{M}$
  - $a : \mathbb{M} \to \{0, 1\}$ – approval strategy of receiver
    - best response $a(m) = \mathbb{1}\big(q(m) \in \mathcal{B}\big)$, $\forall m \in \mathbb{M}$
  - $q : \mathbb{M} \to \Delta\Omega$ – posterior belief of receiver
    - Bayes-rational on equilibrium path

      $\operatorname{supp} q(m) \subseteq m$, $\forall m \in \mathbb{M}$

go back