Preparing for the Worst but Hoping for the Best: Robust (Bayesian) Persuasion

Piotr Dworczak & Alessandro Pavan
(Northwestern University)

July 12-14, 2021

Stony Brook Workshop on Strategic Communication and Learning

Motivation

- In the canonical Bayesian persuasion model, the Sender

Motivation

- In the canonical Bayesian persuasion model, the Sender
  - knows Receivers' sources of information,

Motivation

- In the canonical Bayesian persuasion model, the Sender
  - knows Receivers' sources of information,
  - trusts her ability to coordinate Receivers on actions most favorable to her;

Motivation

- In the canonical Bayesian persuasion model, the Sender
  - knows Receivers' sources of information,
  - trusts her ability to coordinate Receivers on actions most favorable to her;
  - optimal information structure **sensitive to fine details of the Sender's knowledge about the strategic and informational environment**

Motivation

- In many problems of interest,
  - knows Receivers' sources of information,
  - trusts her ability to coordinate Receivers on actions most favorable to her;
  - optimal information structure **sensitive to fine details of the Sender's knowledge about the strategic and informational environment**

Motivation

- In many problems of interest,
  - Receivers' sources of information may be unknown,
  - trusts her ability to coordinate Receivers on actions most favorable to her;
  - optimal information structure **sensitive to fine details of the Sender's knowledge about the strategic and informational environment**

Motivation

- In many problems of interest,
  - Receivers' sources of information may be unknown,
  - Sender may not trust her ability to coordinate Receivers;
  - optimal information structure **sensitive to fine details of the Sender's knowledge about the strategic and informational environment**

Motivation

- In many problems of interest,
  - Receivers' sources of information may be unknown,
  - Sender may not trust her ability to coordinate Receivers;
  - Quest for robustness

Motivation

- In many problems of interest,
  - Receivers' sources of information may be unknown,
  - Sender may not trust her ability to coordinate Receivers;
  - Quest for robustness
- We propose a robust solution concept for Bayesian persuasion that

Motivation

- In many problems of interest,
  - Receivers' sources of information may be unknown,
  - Sender may not trust her ability to coordinate Receivers;
  - Quest for robustness
- We propose a robust solution concept for Bayesian persuasion that
  - incorporates a concern for robustness without abandoning the Bayesian model completely;

Motivation

- In many problems of interest,
  - Receivers' sources of information may be unknown,
  - Sender may not trust her ability to coordinate Receivers;
  - Quest for robustness
- We propose a robust solution concept for Bayesian persuasion that
  - incorporates a concern for robustness without abandoning the Bayesian model completely;
  - suggests an approach to robust design that could be used outside of information design contexts.

Motivating example

- Sender = Prosecutor, Receiver = Judge

Motivating example

- Sender = Prosecutor, Receiver = Judge
- $\Omega = \{$**i**(nnocent), **m**(isdemeanor), **f**(elony)$\}$, $\mu_0 = (1/2, 1/4, 1/4)$.

Motivating example

- Sender = Prosecutor, Receiver = Judge
- $\Omega = \{$**i**(nnocent), **m**(isdemeanor), **f**(elony)$\}$, $\mu_0 = (1/2, 1/4, 1/4)$.
- Judge convicts if she believes the suspect to be guilty with probability $\geq 2/3$.

Motivating example

- Sender = Prosecutor, Receiver = Judge
- $\Omega = \{$**i**(nnocent), **m**(isdemeanor), **f**(elony)$\}$, $\mu_0 = (1/2, 1/4, 1/4)$.
- Judge convicts if she believes the suspect to be guilty with probability $\geq 2/3$.
- Conditional on convicting, the Judge chooses a sentence $x \in [\underline{x}, \bar{x}]$.

Motivating example

- Sender = Prosecutor, Receiver = Judge
- $\Omega = \{$**i**(nnocent), **m**(isdemeanor), **f**(elony)$\}$, $\mu_0 = (1/2, 1/4, 1/4)$.
- Judge convicts if she believes the suspect to be guilty with probability $\geq 2/3$.
- Conditional on convicting, the Judge chooses a sentence $x \in [\underline{x}, \bar{x}]$.
- Sentence is linearly increasing in the conditional probability of f, up to a cap of $\bar{x}$ reached at probability $1/2$.

Motivating example

- Sender = Prosecutor, Receiver = Judge
- $\Omega = \{$**i**(nnocent), **m**(isdemeanor), **f**(elony)$\}$, $\mu_0 = (1/2,\, 1/4,\, 1/4)$.
- Judge convicts if she believes the suspect to be guilty with probability $\geq 2/3$.
- Conditional on convicting, the Judge chooses a sentence $x \in [\underline{x},\, \bar{x}]$.
- Sentence is linearly increasing in the conditional probability of f, up to a cap of $\bar{x}$ reached at probability $1/2$.
- Prosecutor wants to maximize expected sentence:

$$\widehat{V}(\mu) = \mathbf{1}_{\{\mu(m)+\mu(f)\geq \frac{2}{3}\}} \min\{\bar{x},\, \underline{x} + 2\frac{\mu(f)}{\mu(f)+\mu(m)}(\bar{x}-\underline{x})\}.$$

Motivating example

Bayesian solution:

- The Prosecutor induces the posterior belief:

Motivating example

Bayesian solution:

- The Prosecutor induces the posterior belief:
  - $(\mu(i), \mu(m), \mu(f)) = (1, 0, 0)$ with probability $1/4$
    by saying "innocent" with probability $1/2$ conditional on state i;

Motivating example

Bayesian solution:

- The Prosecutor induces the posterior belief:
  - $(\mu(i), \mu(m), \mu(f)) = (1, 0, 0)$ with probability $1/4$
    by saying "innocent" with probability $1/2$ conditional on state i;
  - $(\mu(i), \mu(m), \mu(f)) = (1/3, 1/3, 1/3)$ with probability $3/4$
    by saying "guilty" in all other cases.

Motivating example

Bayesian solution:

- The Prosecutor induces the posterior belief:
  - $(\mu(i),\, \mu(m),\, \mu(f)) = (1, 0, 0)$ with probability $1/4$
    by saying "innocent" with probability $1/2$ conditional on state i;
  - $(\mu(i),\, \mu(m),\, \mu(f)) = (1/3,\, 1/3,\, 1/3)$ with probability $3/4$
    by saying "guilty" in all other cases.
- The expected payoff is $(3/4)\bar{x}$.

Motivating example

Bayesian solution:

- The Prosecutor induces the posterior belief:
  - $(\mu(i),\, \mu(m),\, \mu(f)) = (1, 0, 0)$ with probability $1/4$
    by saying "innocent" with probability $1/2$ conditional on state i;
  - $(\mu(i),\, \mu(m),\, \mu(f)) = (1/3,\, 1/3,\, 1/3)$ with probability $3/4$
    by saying "guilty" in all other cases.
- The expected payoff is $(3/4)\bar{x}$.
- **Implicit assumption:** The Prosecutor believes ("conjectures") the she is the sole provider of information.

Motivating example

Bayesian solution:

- The Prosecutor induces the posterior belief:
  - $(\mu(i),\, \mu(m),\, \mu(f)) = (1, 0, 0)$ with probability $1/4$
    by saying "innocent" with probability $1/2$ conditional on state i;
  - $(\mu(i),\, \mu(m),\, \mu(f)) = (1/3,\, 1/3,\, 1/3)$ with probability $3/4$
    by saying "guilty" in all other cases.
- The expected payoff is $(3/4)\bar{x}$.
- **Implicit assumption:** The Prosecutor believes ("conjectures") the she is the sole provider of information.
- But what if the Judge calls a Witness?

Motivating example

Robustness:

- The Prosecutor might not know the likelihood of the Witness appearing in court, the amount of information that the she has, or the Witness' motives.

Motivating example

Robustness:

- The Prosecutor might not know the likelihood of the Witness appearing in court, the amount of information that the she has, or the Witness' motives.
- Assumption: The Sender is concerned about the worst-case scenario.

Motivating example

Robustness:

- The Prosecutor might not know the likelihood of the Witness appearing in court, the amount of information that the she has, or the Witness' motives.
- Assumption: The Sender is concerned about the worst-case scenario.
- In the worst case, the Sender cannot do better than the full-disclosure payoff, $(1/4)\underline{x} + (1/4)\bar{x}$.

Motivating example

Robustness:

- The Prosecutor might not know the likelihood of the Witness appearing in court, the amount of information that the she has, or the Witness' motives.
- Assumption: The Sender is concerned about the worst-case scenario.
- In the worst case, the Sender cannot do better than the full-disclosure payoff, $(1/4)\underline{x} + (1/4)\bar{x}$.
    - She clearly cannot do strictly better;

Motivating example

Robustness:

- The Prosecutor might not know the likelihood of the Witness appearing in court, the amount of information that the she has, or the Witness' motives.
- Assumption: The Sender is concerned about the worst-case scenario.
- In the worst case, the Sender cannot do better than the full-disclosure payoff, $(1/4)\underline{x} + (1/4)\bar{x}$.
  - She clearly cannot do strictly better;
  - She can achieve that payoff by disclosing the state herself.

Motivating example

Robustness:

- The Prosecutor might not know the likelihood of the Witness appearing in court, the amount of information that the she has, or the Witness' motives.
- Assumption: The Sender is concerned about the worst-case scenario.
- In the worst case, the Sender cannot do better than the full-disclosure payoff, $(1/4)\underline{x} + (1/4)\bar{x}$.
  - She clearly cannot do strictly better;
  - She can achieve that payoff by disclosing the state herself.
- This was already observed by Hu and Weng (2019).

Motivating example

Robustness:

- The Prosecutor might not know the likelihood of the Witness appearing in court, the amount of information that the she has, or the Witness' motives.
- Assumption: The Sender is concerned about the worst-case scenario.
- In the worst case, the Sender cannot do better than the full-disclosure payoff, $(1/4)\underline{x} + (1/4)\bar{x}$.
  - She clearly cannot do strictly better;
  - She can achieve that payoff by disclosing the state herself.
- This was already observed by Hu and Weng (2019).
- Key idea of this paper:

  The Sender should **not** fully disclose the state in this case!

Motivating example

Superior policy:

- The Prosecutor reveals the state i but pools together the states m and f.

Motivating example

Superior policy:

- The Prosecutor reveals the state i but pools together the states m and f.
- In the worst case:

Motivating example

Superior policy:

- The Prosecutor reveals the state i but pools together the states m and f.
- In the worst case:
  - When the state is i, the Witness has no more information to reveal.

Motivating example

Superior policy:

- The Prosecutor reveals the state i but pools together the states m and f.
- In the worst case:
  - When the state is i, the Witness has no more information to reveal.
  - When the state is m or f, the Witness reveals the state.

Motivating example

Superior policy:

- The Prosecutor reveals the state i but pools together the states m and f.
- In the worst case:
  - When the state is i, the Witness has no more information to reveal.
  - When the state is m or f, the Witness reveals the state.
- Thus, in the worst case, the payoff is **exactly the same** as under full disclosure by the Sender.

Motivating example

Superior policy:

- The Prosecutor reveals the state i but pools together the states m and f.
- In the worst case:
  - When the state is i, the Witness has no more information to reveal.
  - When the state is m or f, the Witness reveals the state.
- Thus, in the worst case, the payoff is **exactly the same** as under full disclosure by the Sender.
- If the conjecture happens to be right:

Motivating example

Superior policy:

- The Prosecutor reveals the state i but pools together the states m and f.
- In the worst case:
  - When the state is i, the Witness has no more information to reveal.
  - When the state is m or f, the Witness reveals the state.
- Thus, in the worst case, the payoff is **exactly the same** as under full disclosure by the Sender.
- If the conjecture happens to be right:
  - When the state is m or f, the Judge's belief is $(0, 1/2, 1/2)$;

Motivating example

Superior policy:

- The Prosecutor reveals the state i but pools together the states m and f.
- In the worst case:
  - When the state is i, the Witness has no more information to reveal.
  - When the state is m or f, the Witness reveals the state.
- Thus, in the worst case, the payoff is **exactly the same** as under full disclosure by the Sender.
- If the conjecture happens to be right:
  - When the state is m or f, the Judge's belief is $(0, 1/2, 1/2)$;
  - The expected payoff is $(1/2)\bar{x} > (1/4)\underline{x} + (1/4)\bar{x}$.

Motivating example

Superior policy:

- The Prosecutor reveals the state i but pools together the states m and f.
- In the worst case:
  - When the state is i, the Witness has no more information to reveal.
  - When the state is m or f, the Witness reveals the state.
- Thus, in the worst case, the payoff is **exactly the same** as under full disclosure by the Sender.
- If the conjecture happens to be right:
  - When the state is m or f, the Judge's belief is $(0,\, 1/2,\, 1/2)$;
  - The expected payoff is $(1/2)\bar{x} > (1/4)\underline{x} + (1/4)\bar{x}$.
- This policy is just as good as full disclosure in the worst-case but strictly better when the Sender's conjecture turns out to be right.

Motivating example

Robust solution:

- The Prosecutor reveals the state i but pools together the states m and f.
- In the worst case:
  - When the state is i, the Witness has no more information to reveal.
  - When the state is m or f, the Witness reveals the state.
- Thus, in the worst case, the payoff is **exactly the same** as under full disclosure by the Sender.
- If the conjecture happens to be right:
  - When the state is m or f, the Judge's belief is $(0, 1/2, 1/2)$;
  - The expected payoff is $(1/2)\bar{x} > (1/4)\underline{x} + (1/4)\bar{x}$.
- This policy is just as good as full disclosure in the worst-case but strictly better when the Sender's conjecture turns out to be right.

Motivating example

Why not use a Bayesian solution?

- In the Bayesian solution, the Prosecutor sometimes says "guilty" which induces a posterior belief $(1/3,\, 1/3,\, 1/3)$.

Motivating example

Why not use a Bayesian solution?

- In the Bayesian solution, the Prosecutor sometimes says "guilty" which induces a posterior belief $(1/3, 1/3, 1/3)$.
- Adversarial scenario:

Motivating example

Why not use a Bayesian solution?

- In the Bayesian solution, the Prosecutor sometimes says “guilty” which induces a posterior belief $(1/3, 1/3, 1/3)$.
- Adversarial scenario:
  - The Witness tells the truth, and she knows the state f with probability ϵ (and nothing otherwise).

Motivating example

Why not use a Bayesian solution?

- In the Bayesian solution, the Prosecutor sometimes says "guilty" which induces a posterior belief $(1/3,\, 1/3,\, 1/3)$.
- Adversarial scenario:
  - The Witness tells the truth, and she knows the state f with probability ϵ (and nothing otherwise).
  - With conditional probability $1-\epsilon$, the Judge's belief becomes $(1/(3-\epsilon),\, 1/(3-\epsilon),\, (1-\epsilon)/(3-\epsilon))$.

Motivating example

Why not use a Bayesian solution?

- In the Bayesian solution, the Prosecutor sometimes says "guilty" which induces a posterior belief $(1/3, 1/3, 1/3)$.
- Adversarial scenario:
  - The Witness tells the truth, and she knows the state f with probability ϵ (and nothing otherwise).
  - With conditional probability $1 - \epsilon$, the Judge's belief becomes $(1/(3-\epsilon), 1/(3-\epsilon), (1-\epsilon)/(3-\epsilon))$.
  - This induces the Judge to acquit the suspect.

Motivating example

Why not use a Bayesian solution?

- In the Bayesian solution, the Prosecutor sometimes says "guilty" which induces a posterior belief $(1/3,\, 1/3,\, 1/3)$.
- Adversarial scenario:
  - The Witness tells the truth, and she knows the state f with probability ϵ (and nothing otherwise).
  - With conditional probability $1-\epsilon$, the Judge's belief becomes $(1/(3-\epsilon),\, 1/(3-\epsilon),\, (1-\epsilon)/(3-\epsilon))$.
  - This induces the Judge to acquit the suspect.
- The expected payoff from the Bayesian solution is 0 in the worst case.

Motivating example

Why not use a Bayesian solution?

- In the Bayesian solution, the Prosecutor sometimes says "guilty" which induces a posterior belief $(1/3,\, 1/3,\, 1/3)$.
- Adversarial scenario:
  - The Witness tells the truth, and she knows the state f with probability ϵ (and nothing otherwise).
  - With conditional probability $1-\epsilon$, the Judge's belief becomes $(1/(3-\epsilon),\, 1/(3-\epsilon),\, (1-\epsilon)/(3-\epsilon))$.
  - This induces the Judge to acquit the suspect.
- The expected payoff from the Bayesian solution is 0 in the worst case.

Motivating example

Why not use a Bayesian solution?

- In the Bayesian solution, the Prosecutor sometimes says "guilty" which induces a posterior belief $(1/3,\, 1/3,\, 1/3)$.
- Adversarial scenario:
  - The Witness tells the truth, and she knows the state f with probability ϵ (and nothing otherwise).
  - With conditional probability $1-\epsilon$, the Judge's belief becomes $(1/(3-\epsilon),\, 1/(3-\epsilon),\, (1-\epsilon)/(3-\epsilon))$.
  - This induces the Judge to acquit the suspect.
- The expected payoff from the Bayesian solution is 0 in the worst case.

A Bayesian solution need not be robust to misspecifications in the Sender's beliefs.

Our approach

- The Sender has a conjecture but is concerned about the robustness of her chosen policy.

Our approach

- The Sender has a conjecture but is concerned about the robustness of her chosen policy.
- **Robust solutions:**

Our approach

- The Sender has a conjecture but is concerned about the robustness of her chosen policy.

- **Robust solutions:**
  - Step 1: "Preparing for the worst"

Our approach

- The Sender has a conjecture but is concerned about the robustness of her chosen policy.
- **Robust solutions:**
  - Step 1: "Preparing for the worst"
    - Optimal in case Nature provides information and coordinates play to minimize the Sender's payoff (worst-case scenario);

Our approach

- The Sender has a conjecture but is concerned about the robustness of her chosen policy.

- **Robust solutions:**
  - Step 1: "Preparing for the worst"
    - ⋆ Optimal in case Nature provides information and coordinates play to minimize the Sender's payoff (worst-case scenario);
  - Step 2: "Hoping for the best"

Our approach

- The Sender has a conjecture but is concerned about the robustness of her chosen policy.

- **Robust solutions:**
  - Step 1: "Preparing for the worst"
    - Optimal in case Nature provides information and coordinates play to minimize the Sender's payoff (worst-case scenario);
  - Step 2: "Hoping for the best"
    - Among all worst-case optimal policies, optimal when Nature plays according to the Sender's conjecture (best-case scenario)

Our approach

- The Sender has a conjecture but is concerned about the robustness of her chosen policy.
- **Robust solutions:**
  - Step 1: "Preparing for the worst"
    - Optimal in case Nature provides information and coordinates play to minimize the Sender's payoff (worst-case scenario);
  - Step 2: "Hoping for the best"
    - Among all worst-case optimal policies, optimal when Nature plays according to the Sender's conjecture (best-case scenario)
- Robust solutions are defined as policies that are best-case optimal among worst-case optimal ones.

General framework

- Any Bayesian persuasion setting as long as the Sender uses public signals.

General framework

- Any Bayesian persuasion setting as long as the Sender uses public signals.
- Nature chooses an arbitrary signal (that can depend on the Sender's signal realization) and selects the equilibrium (consistent with the assumed solution concept).

General framework

- Any Bayesian persuasion setting as long as the Sender uses public signals.
- Nature chooses an arbitrary signal (that can depend on the Sender's signal realization) and selects the equilibrium (consistent with the assumed solution concept).
- Arbitrary conjecture of the Sender about Nature's choice.

General framework

- Any Bayesian persuasion setting as long as the Sender uses public signals.
- Nature chooses an arbitrary signal (that can depend on the Sender's signal realization) and selects the equilibrium (consistent with the assumed solution concept).
- Arbitrary conjecture of the Sender about Nature's choice.
- A robust solution maximizes the Sender's payoff against the conjecture among all policies that are optimal in the worst-case scenario (when Nature chooses the information structure and equilibrium to minimize the Sender's payoff).

Separation Theorem

Theorem (KG, 2011)

Distribution $\rho \in \Delta\Delta\Omega$ *is a* **Bayesian solution** *iff it maximizes*

$$\int \widehat{V}(\mu) d\rho(\mu)$$

over all Bayes-plausible distributions over posterior beliefs $\rho \in \Delta\Delta\Omega$.

Separation Theorem

Theorem (Separation Theorem)

There exists $\mathcal{F} \subset \mathcal{P}(\Omega)$ *such that* $\rho \in \Delta\Delta\Omega$ *is a* **robust solution** *iff it maximizes*

$$\int \widehat{V}(\mu) d\rho(\mu)$$

over all Bayes-plausible distributions over posterior beliefs $\rho \in \Delta\Delta\Omega$ *such that*

$$supp(\mu) \in \mathcal{F}, \forall \mu \in supp(\rho).$$

Separation Theorem

Theorem (Separation Theorem)

There exists $\mathcal{F} \subset \mathcal{P}(\Omega)$ *such that* $\rho \in \Delta\Delta\Omega$ *is a* **robust solution** *iff it maximizes*

$$\int \widehat{V}(\mu) d\rho(\mu)$$

over all Bayes-plausible distributions over posterior beliefs $\rho \in \Delta\Delta\Omega$ *such that*

$$supp(\mu) \in \mathcal{F}, \forall \mu \in supp(\rho).$$

Moreover,

$$\mathcal{F} \equiv \{B \subseteq \Omega : \underline{V}(\mu) \geq \underline{V}_{\mathit{full}}(\mu), \forall \mu \in \Delta B\}.$$

Corollaries of the Separation Theorem

1. A robust solution exists (under standard assumptions).

Corollaries of the Separation Theorem

1. A robust solution exists (under standard assumptions).
2. For any Bayesian solution, there exists a robust solution that either Blackwell dominates it or is incomparable in the Blackwell order.

Corollaries of the Separation Theorem

1. A robust solution exists (under standard assumptions).
2. For any Bayesian solution, there exists a robust solution that either Blackwell dominates it or is incomparable in the Blackwell order.
3. Robust solutions can be found by concavifying a modified objective function.

Corollaries of the Separation Theorem

1. A robust solution exists (under standard assumptions).
2. For any Bayesian solution, there exists a robust solution that either Blackwell dominates it or is incomparable in the Blackwell order.
3. Robust solutions can be found by concavifying a modified objective function.
4. Under regularity conditions, robust solutions are exactly limits of solutions to a weighted problem, as the weight on the worst-case scenario converges to 1.

Corollaries of the Separation Theorem

1. A robust solution exists (under standard assumptions).
2. For any Bayesian solution, there exists a robust solution that either Blackwell dominates it or is incomparable in the Blackwell order.
3. Robust solutions can be found by concavifying a modified objective function.
4. Under regularity conditions, robust solutions are exactly limits of solutions to a weighted problem, as the weight on the worst-case scenario converges to 1.
5. When the conjecture satisfies a permissive condition, all robust solutions are undominated.

Literature

- **Bayesian persuasion**
  - Kamenica and Gentzkow (2011), surveys: Bergemann and Morris (2019), Kamenica (2019)
- **Information design with adversarial coordination**
  - Inostroza and Pavan (2018), Mathevet, Perego, Taneva (2019), Morris, Oyama, and Takahashi, (2019), Ziegler (2019), Li, Song, and Zhao (2020)
- **Persuasion with unknown beliefs**
  - Non-Bayesian uncertainty: Hu and Weng (2019), Kosterina (2019);
  - Bayesian uncertainty: Kolotilin, Mylovanov, Zapechelnyuk, and Li (2017), Laclau and Renou (2017), Guo and Schmaya (2018)
- **Competition in persuasion**
  - Simultaneous: Gentzkow and Kamenica (2016, 2017) , Board and Lu (2018), Au and Kawai (2018), Cui and Ravindran (2020);
  - Sequential: Li and Norman (2019), and Wu (2018)
- **Max-max over max-min design**
  - Börgers (2017)

Conclusions

- Bayesian persuasion when Sender does not trust her conjecture about
  - Receivers' exogenous information;
  - Equilibrium selection.
- Robust solutions
  - Maximize the payoff under the conjecture but on the set of policies that guarantee the best payoff in the worst case;
  - Weakly more tractable than Bayesian solutions;
  - Never disclose less information;
  - Are guaranteed to be undominated.
- Future work:
  - Analogous approach in mechanism design
  - Decision-theoretic foundations for this solution concept