Tailored Stories

Chiara Aina
University of Zurich

December, 2021

Motivation

People often adopt very different stories to explain the same observation

- Voters do not agree on the outcome of an election

  "The election system is fair"

  "Elections are rigged"

- Consumers differ in evaluating a company

  "This company has great corporate social responsability"

  "The company is just doing green washing"

- Feedback often misperceived in different ways

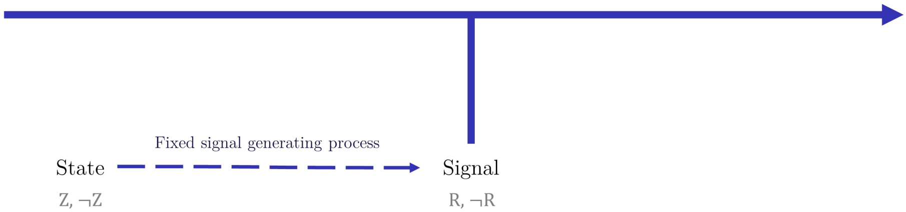
  "Grades reflect ability"

  "Grades were usually random, they don't convey much about ability"

Example

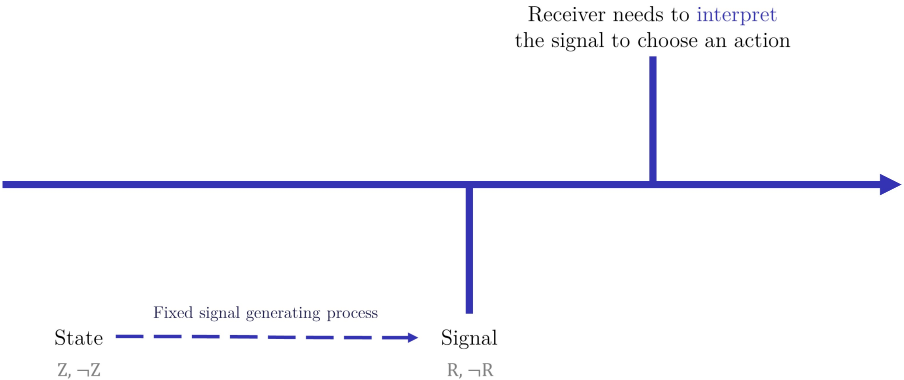

VOTE FOR ZAPHOD BEEBLEBROX
SHOULD I TRUST THE ELECTION OUTCOME ?

Timeline



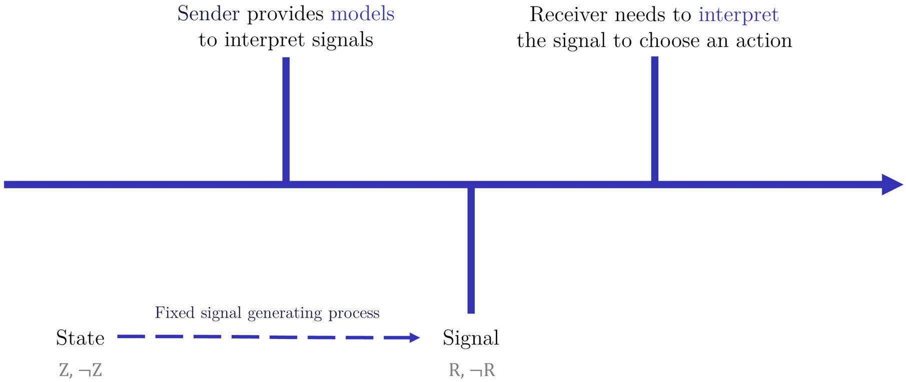
Z: Zaphod is the legitimate elected president
R: reported votes assign majority to Zaphod

Timeline



Z: Zaphod is the legitimate elected president
R: reported votes assign majority to Zaphod

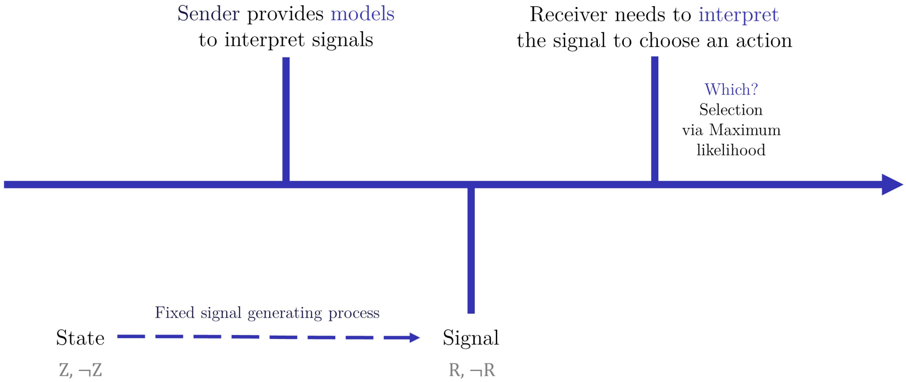
Timeline



Z: Zaphod is the legitimate elected president

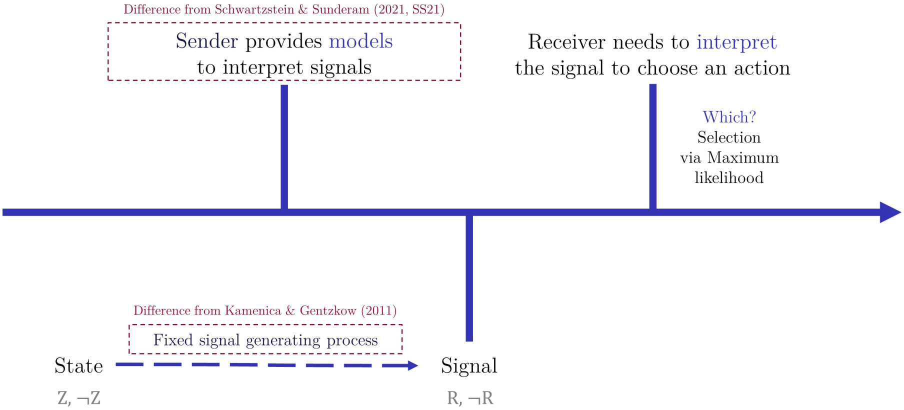
R: reported votes assign majority to Zaphod

Timeline



Z: Zaphod is the legitimate elected president
R: reported votes assign majority to Zaphod

Timeline



Z: Zaphod is the legitimate elected president
R: reported votes assign majority to Zaphod

Literature Review

Interpretations

The sender does not know the signal when he communicates the narratives

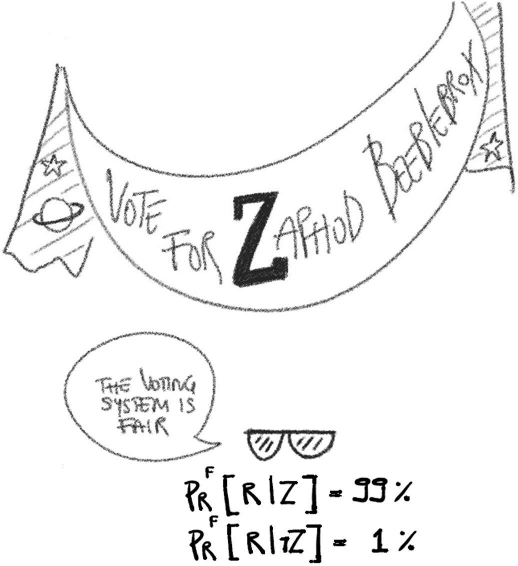
1. Temporal: the sender communicates before the signal realizes
2. Private information: the receiver knows the signal, but the sender does not

Ex-ante Persuasion

$$\text{MODEL} \begin{cases} \Pr[R \mid Z] \\ \Pr[R \mid \neg Z] \end{cases}$$

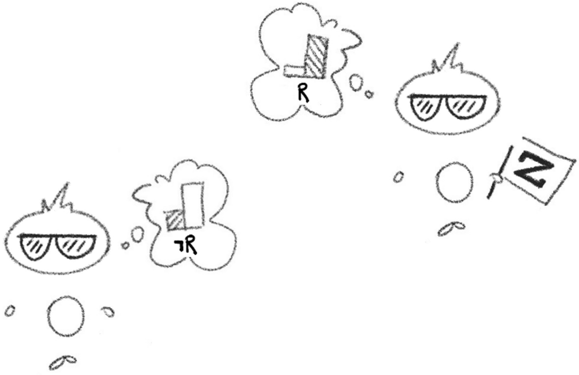
Z: Zaphod is the legitimate elected president
R: reported votes assign majority to Zaphod

One Story



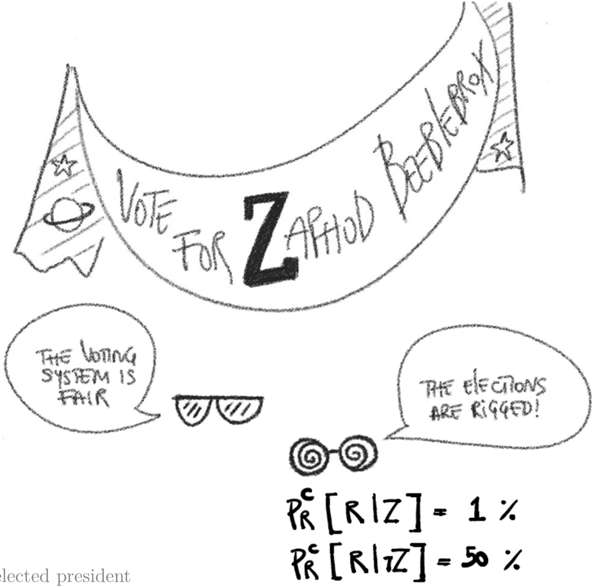
Z: Zaphod is the legitimate elected president
R: reported votes assign majority to Zaphod

One Story



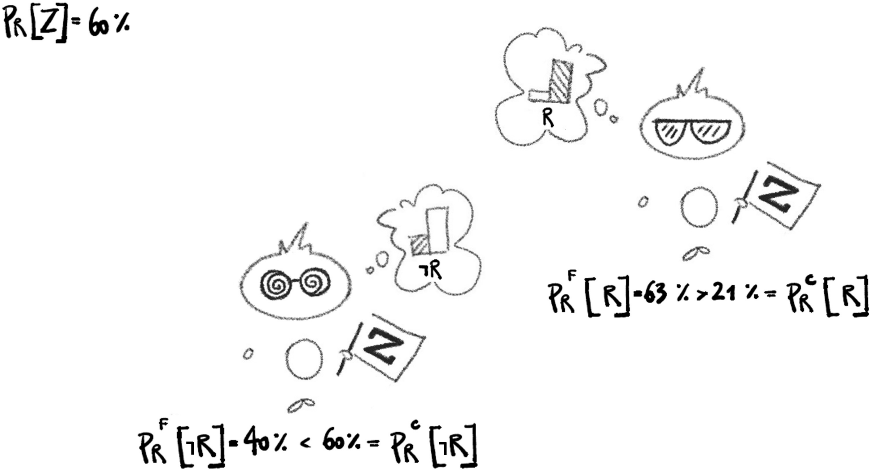
Z: Zaphod is the legitimate elected president
R: reported votes assign majority to Zaphod

Tailored Stories



Z: Zaphod is the legitimate elected president
R: reported votes assign majority to Zaphod

Tailored Stories



Z: Zaphod is the legitimate elected president
R: reported votes assign majority to Zaphod

Research Question

I study the problem of manipulating a boundedly rational agent by controlling her interpretation of signals she is about to receive

Is it possible to persuade others only by providing interpretations of future events?

- Not only it is possible, but it can also lead the receiver to hold inconsistent beliefs across observed events
  - Allowing for multiple stories to be communicated,
    I provide a disciplined relaxation of the Bayes-plausibility constraint
    In expectations, posteriors do not need to average to the prior
- Persuasion is generally limited and it depends on the initial beliefs

This Paper

- What is a story? What are its properties?
- What can the receiver be persuaded of?
- What is the optimal set of stories the sender should communicate?
- Applications:
  1. Elections
  2. Finance
  3. Nudging
  4. Intra-personal Phenomena

This Paper: Today

- What is a story? What are its properties?
- What can the receiver be persuaded of?
- What is the optimal set of stories the sender should communicate?
- Applications:
  1. Elections
  2. Finance
  3. Nudging
  4. Intra-personal Phenomena

What is a story? What are its properties?

Set Up

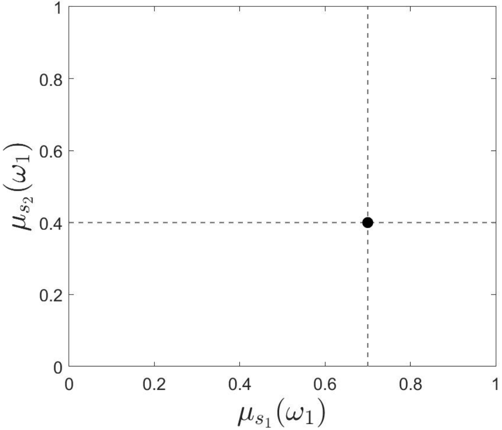
- States: $\omega \in \Omega$
- Common prior on Ω: $\mu_0 \in int(\Delta(\Omega))$
- Signals: $s \in S$
- **Model** m: $(\pi^m(s|\omega))_{s \in S, \omega \in \Omega} \in [\Delta(S)]^\Omega$
  map assigning to each state a distribution of signals conditional on that state
- Adopting model m, an agent forms beliefs conditional on signal s via Bayes rule

$$\mu_s^m = (\mu_s^m(\omega))_{\omega \in \Omega} \in \Delta(\Omega)$$

Main Object of the Analysis

- Vector of posterior beliefs: $\boldsymbol{\mu}^m = (\mu_s^m)_{s \in S} \in [\Delta(\Omega)]^S$
  array of posterior distributions conditional on each signal realization

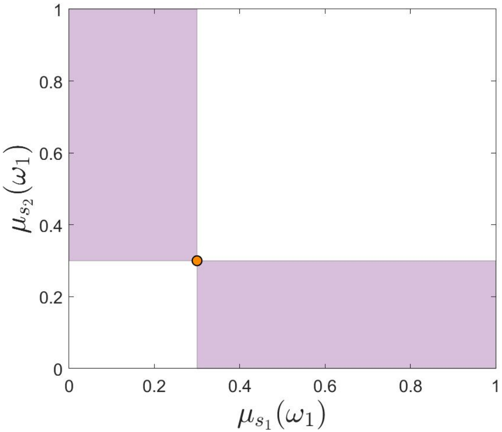
- Binary case: binary state & binary signal
- Axis: posterior of ω_1 conditional on each signal
- Point: vector of posterior beliefs



Equivalent Representation

A vector of posterior beliefs $\boldsymbol{\mu}$ is **Bayes-consistent**
if the prior μ_0 is a convex combination of the posterior across signals $(\mu_s)_{s \in S}$

- Bayes-plausibility $\Rightarrow$ Bayes-consistency
- Equivalent representation between models and Bayes-consistent vectors of posteriors
  Details
  - Orange point: prior distribution
  - Every point in the purple area is not a model but it corresponds to a model

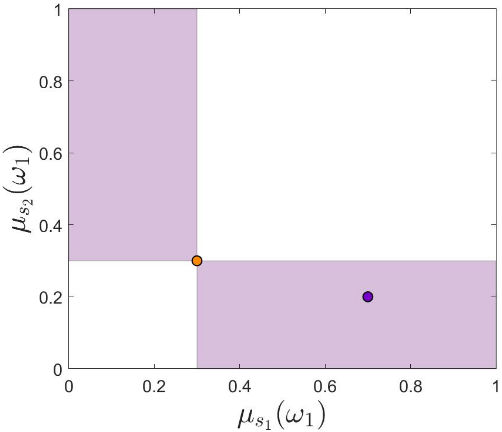


Properties: Fit

Fit of a model m conditional on the signal s: $\Pr^m(s) = \sum_{\omega \in \Omega} \mu_0(\omega)\, \pi^m(s|\omega)$

- It measures how likely a model fits the observed data

- *How to see the fit in a picture?*

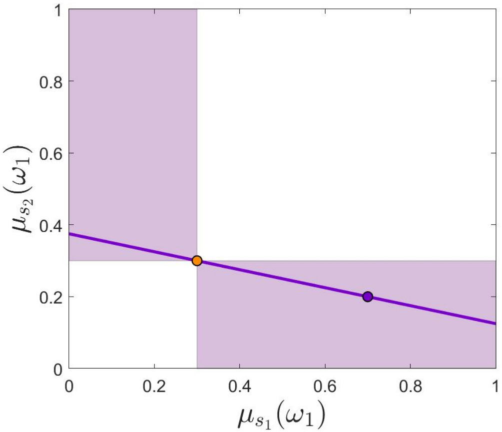


Properties: Fit

Fit of a model m conditional on the signal s: $\Pr^m(s) = \sum_{\omega \in \Omega} \mu_0(\omega)\, \pi^m(s|\omega)$

- It measures how likely a model fits the observed data

- *How to see the fit in a picture?*
    - Isofit line: all the points correspond to models that have the same fit (except the prior) More
- *What if the slope changes?*
    - The steeper, the higher fit given s_1
    - The flatter, the higher fit given s_2

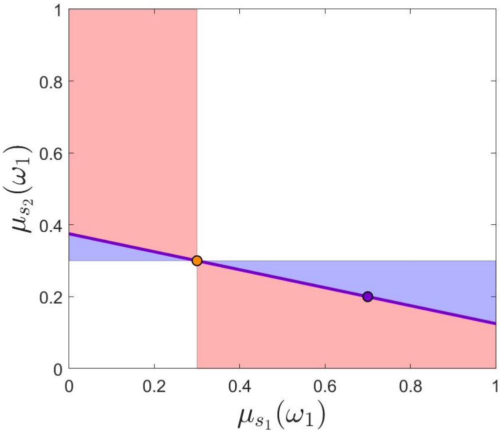


Properties: Fit

Fit of a model *m* conditional on the signal *s*: $\Pr^m(s) = \sum_{\omega \in \Omega} \mu_0(\omega)\, \pi^m(s|\omega)$

- It measures how likely a model fits the observed data

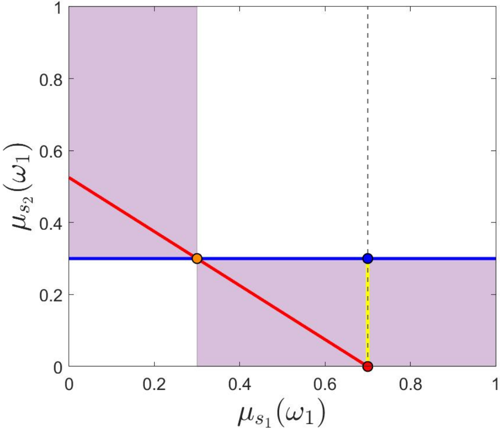
- *How to see the fit in a picture?*
  - Isofit line: all the points correspond to models that have the same fit (except the prior) More
- *What if the slope changes?*
  - The steeper, the higher fit given s_1
  - The flatter, the higher fit given s_2
- The isofit partions the set of Bayes-consistent vectors of posterior beliefs
  - Line: same fit
  - Red area: higher fit given s_1
  - Blue area: higher fit given s_2



Properties: Fit

- There is a multiplicity of models that induce the same posterior distribution conditional on a signal with different levels of fit

- Dotted line: target posterior distribution given s
- Yellow line: all models inducing the target
- Red point: model inducing the target with highest fit given $s_1 \to$ steeper isofit
- Blue point: model inducing the target with highest fit given $s_2 \to$ flatter isofit



Properties: Movement

Movement for μ_s in state ω: $$\delta(\omega;\mu_s) = \frac{\mu_s(\omega)}{\mu_0(\omega)}$$

- It is a measure of how much the target posterior is far from the prior in a state
- Maximal movement for μ_s: $\bar{\delta}(\mu_s) = \max_{\omega\in\Omega} \delta(\omega;\mu_s)$

Lemma Proof

A model inducing μ_s conditional on signal s has fit $\mathrm{Pr}^m(s) \in \left[0, \bar{\delta}(\mu_s)^{-1}\right]$

- SS21 characterizes the upper bound: the maximal fit for a target posterior coincides with the reciprocal of the maximal movement
  - Any model that leads beliefs to react a lot given a signal realization (higher movement) cannot fit the data well (lower fit)

What can the receiver be persuaded of?

Receiver's Problem

- The receiver does not know the state but she has observed a signal realization
- She needs a model to interpret the signal and update her priors
- The sender communicates a set of models $M \subseteq \mathcal{M}$
  $|M|$ is not greater than the number of models that the receiver is willing to consider

Model Adoption

$$\tilde{m}_s \in \arg\max_{m \in M} \mathrm{Pr}^m(s)$$

- Maximum likelihood selection

Action Choice

$$a^*(\mu_s) \in \arg\max_{a \in A} \mathbb{E}_{\mu_s^{\tilde{m}_s}}[U^R(a,\omega)]$$

Tie breaking rule: if indifferent, adopt the model/action maximizing the sender's expected utility

Sender's Problem

What does the sender know?

- The receiver's preferences, the (common) prior, and the number of models that the receiver is willing to consider
- The sender does not know the state, but he is endowed with a model t
  - Predictive probabilities of each signal realization $\mathrm{Pr}^t(s)$
  - Posterior induced by t conditional on each signal realization μ_s^t

Sender's Value of μ, calculated over signal and state realizations using model t

$$V(\mu) = \mathbb{E}^t[U^S(a^*(\mu_s), \omega)] = \sum_{s \in S} \mathrm{Pr}^t(s)\ \mathbb{E}_{\mu^t}\left[U^S(a^*(\mu_s), \omega)|s\right]$$

Sender's Problem

Many Models

The sender chooses the set of models M^* that maximizes his value at $\mu^M = \left(\mu_s^{\tilde{m}_s}\right)_{s \in S}$

$$M^* \in \arg\max_{M \subseteq \mathcal{M}} V(\mu^M) \quad \text{such that} \quad \tilde{m}_s \in \arg\max_{m \in M} \mathrm{Pr}^m(s)$$

One Model

If the receiver considers only one model from the sender, the problem is

$$m^* \in \arg\max_{m \in \mathcal{M}} V(\mu^m)$$

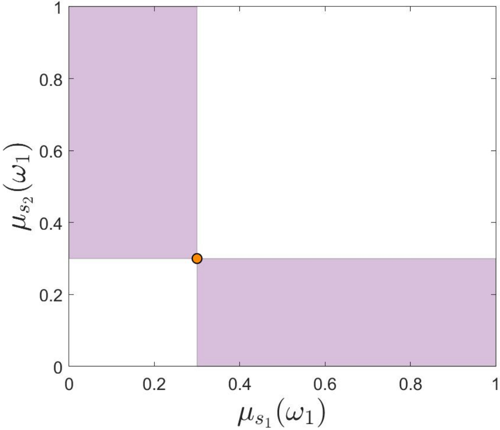
To solve these, it is enough to characterize the set of feasible vector of posterior beliefs

- Why? From the perspective of the sender, there is a fixed distribution over the signals induced by model t: $\left(\mathrm{Pr}^t(s)\right)_{s \in S}$

Set of Feasible Vectors of Posterior Beliefs: One Model

With a model, the sender can only induce Bayes-consistent vectors of posteriors

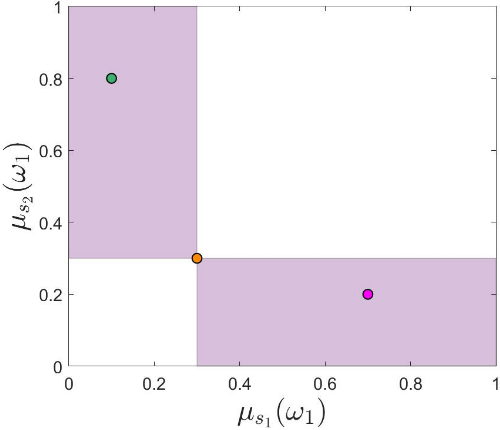
- Comparable characterizing condition to Kamenica & Gentzkow (2011): Bayes-plausibility



Many Models

The sender knows the resulting vector of posterior beliefs of the receiver

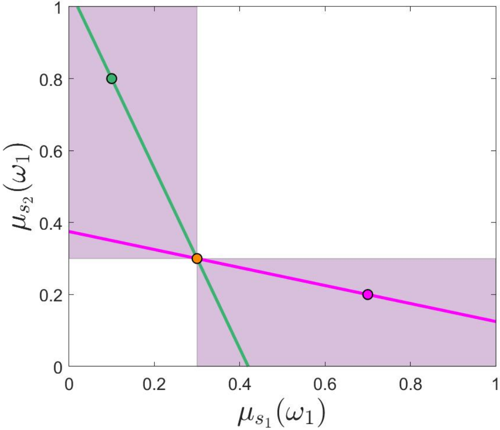
- How? He anticipates which model the receiver adopts conditional on each signal



Many Models

The sender knows the resulting vector of posterior beliefs of the receiver

- How? He anticipates which model the receiver adopts conditional on each signal

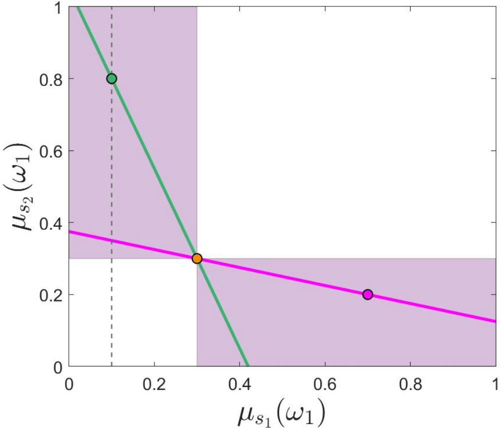


Many Models

The sender knows the resulting vector of posterior beliefs of the receiver

- How? He anticipates which model the receiver adopts conditional on each signal

- *Which model lies on the steeper isofit line?*
  This induces the posteriors given s_1

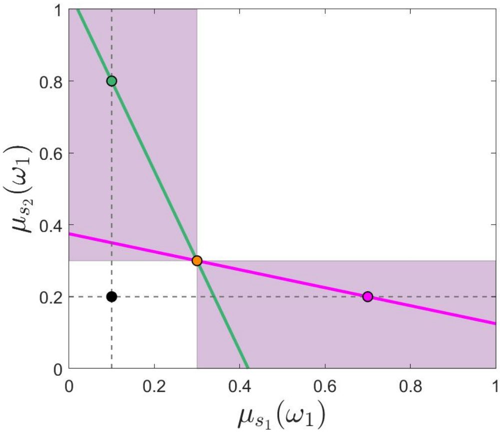


Many Models

The sender knows the resulting vector of posterior beliefs of the receiver

- How? He anticipates which model the receiver adopts conditional on each signal

- *Which model lies on the steeper isofit line?*
  This induces the posteriors given s_1
- *Which model lie on the flatter isofit line?*
  This induces the posteriors given s_2
- *What is the resulting vector of posteriors?*
  More models



Set of Feasible Vectors of Posterior Beliefs: Many Models

With more models, the sender can also induce Bayes-inconsistent vectors of posteriors

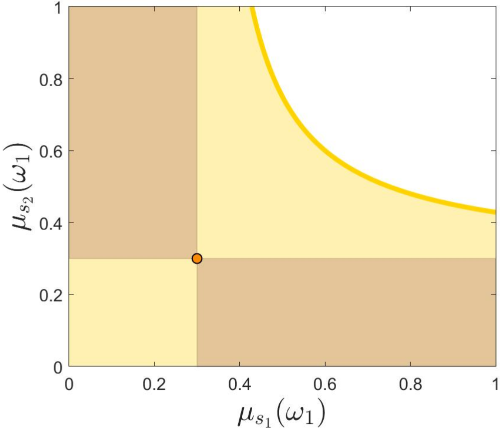
Theorem $|M| \geq |S|$ Proof

The set of feasible vectors of posterior beliefs is

$$\mathcal{F} = \left\{ \boldsymbol{\mu} \in [\Delta(\Omega)]^S : \sum_{s \in S} \bar{\delta}(\mu_s)^{-1} \geq 1 \right\}$$

To be feasible, the sum of the maximal fit levels associated with each signal realization has to exceed the unit

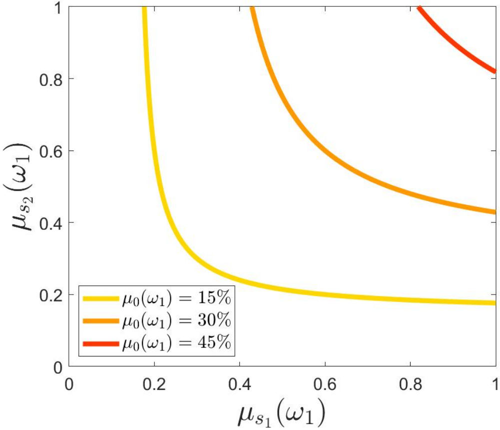
Graphical Intuition



Comparative Statics with Respect to the Priors

Generally not all vectors of posteriors are feasible, but there are exceptions

- The more uniform priors, the more belief manipulability
  More
- Binary case: the closer the receiver's priors are to 50-50, the more she can be manipulated
  Details



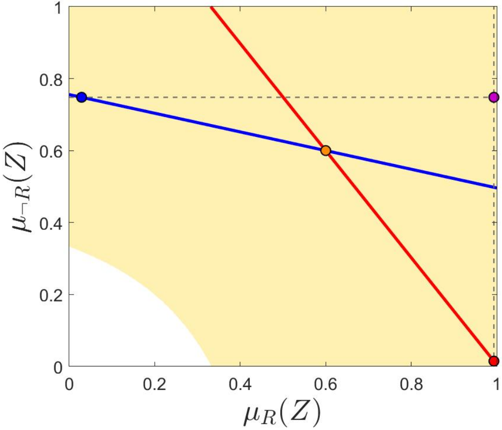
Applications

1. Firehose of Falsehood: model of Russian propaganda based on a large number of possibly contradictory and mutually inconsistent messages (Paul & Matthew, 2016)
   - Effective disinformation campaign in entertaining, confusing, and ultimately manipulating the audience
   - Coordinated operations led by official or unofficial sources
2. Finance
3. Nudging
4. Intra-personal Phenomena

Firehose of Falsehood

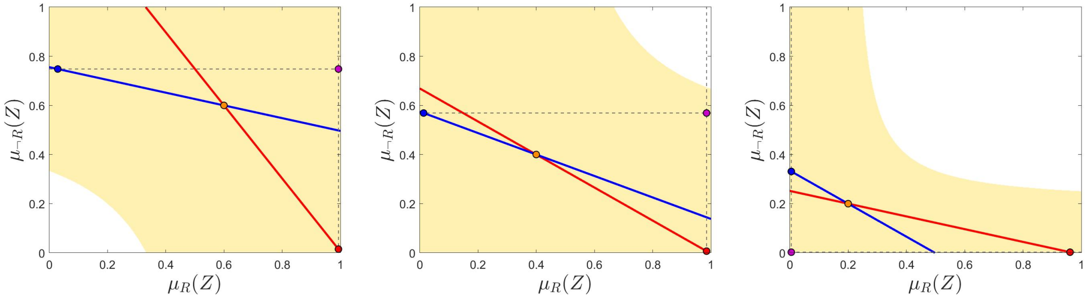
- States: $\{Z, \neg Z\}$, where "Z" is the event that Z is the legitimate elected president
- Signals: $\{R, \neg R\}$, where "R" is that the reported votes assigns majority to Z
- The voter supports Z if $\mu_s(Z) \geq 50\%$
- Z alternates two stories:

1. True model, fair system (red):
   $\pi^t(R|Z) = 99\%$ and $\pi^t(R|\neg Z) = 1\%$
2. Conspiracy theory, rigged elections (blue):
   $\pi^{m_2}(R|Z) = 1\%$ and $\pi^{m_2}(R|\neg Z) = 50\%$
   If legitimate, the votes count is reversed;
   otherwise, the votes are counted randomly



Firehose of Falsehood

Inevitable Polarization



(a) $\mu_0(Z) = 60\%$ (b) $\mu_0(Z) = 40\%$ (c) $\mu_0(Z) = 20\%$

- With conflicting narratives, belief polarization occurs
  - There is a threshold in prior such that voters with prior higher (lower) than the threshold would hold extreme high (low) posteriors regardless the election outcome

More Details

Firehose of Falsehood: 2020 US Elections

Donald J. Trump
@realDonaldTrump

RIGGED 2020 ELECTION: MILLIONS OF MAIL-IN BALLOTS WILL BE PRINTED BY FOREIGN COUNTRIES, AND OTHERS. IT WILL BE THE SCANDAL OF OUR TIMES!

7:16 AM · Jun 22, 2020

Donald J. Trump
@realDonaldTrump

With Universal Mail-In Voting (not Absentee Voting, which is good), 2020 will be the most INACCURATE & FRAUDULENT Election in history. It will be a great embarrassment to the USA. Delay the Election until people can properly, securely and safely vote???

5:46 AM · Jul 30, 2020

Donald J. Trump
@realDonaldTrump

The United States cannot have all Mail In Ballots. It will be the greatest Rigged Election in history. People grab them from mailboxes, print thousands of forgeries and "force" people to sign. Also, forge names. Some absentee OK, when necessary. Trying to use Covid for this Scam!

7:08 AM · May 24, 2020

Donald J. Trump
@realDonaldTrump

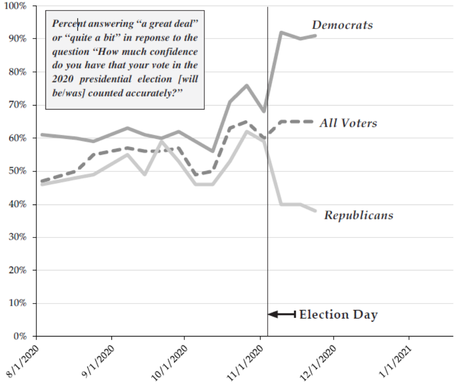
NORTH CAROLINA: To make sure your Ballot COUNTS, sign & send it in EARLY. When Polls open, go to your Polling Place to see if it was COUNTED. IF NOT, VOTE! Your signed Ballot will not count because your vote has been posted. Don't let them illegally take your vote away from you!

9:10 AM · Sep 12, 2020

Firehose of Falsehood: 2020 US Elections

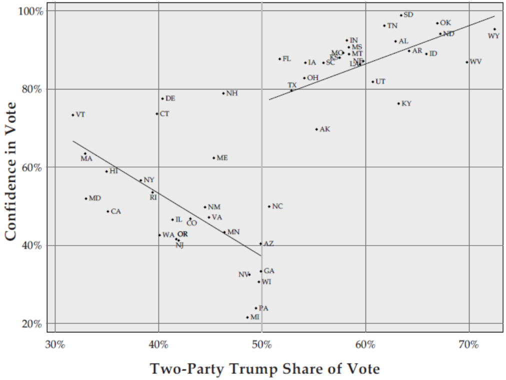
Clark and Stewart (2021)

Accuracy of Vote Count



Voters with different priors adopt different narratives once the signal realizes

Confidence in Vote Count by State, Republicans



Voters with similar priors adopt different narrative once observed different outcomes

Applications

1. Firehose of Falsehood
2. Finance: a financial advisor wants to persuade investors in invest in small stocks; however, he knows that investors' financial experience (either positive or negative) influences their beliefs on the quality of the new investment
    - The advisor does not know the investors' experience
    - Targeting the best persuasive story for each type of investor is unfeasible
    - Second-best: communicating several ways of interpreting favorably the new asset depending on experience, letting each investor then to adopt the narrative that resonate best given their experience
3. Nudging
4. Intra-personal Phenomena

Financial Application

Investor

- Allocate her 1 unit over two outcomes $\Omega = \{\omega_1, \omega_2\}$ resulting in $\alpha = (\alpha_1, \alpha_2)$

  ω_1 is the event in which the small stock outperforms the market

- Expected utility: $\mathbb{E}[U^R(\alpha)] = \sum_{i=1}^{2} \mu_i \log(\alpha_i)$ with $\alpha_i^* = \mu_s(\omega_i)$

- Past experience is good or bad $S = \{g, b\}$

Advisor

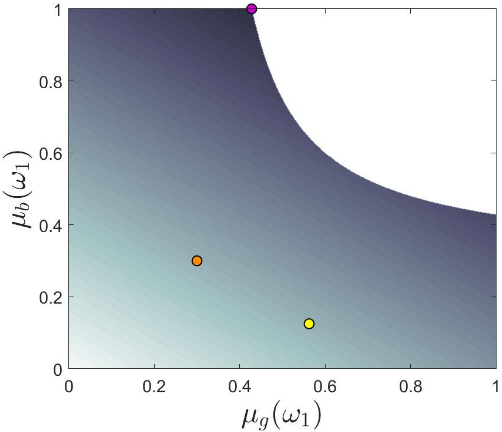
- Commission proportional to α_1

$$V(\mu) = \sum_s r \ \mu_s(\omega_1)$$

- Expect 40% (60%) good (bad) experience

- Yellow point: sender's vector of posteriors by t

- The darker, the higher the sender's value

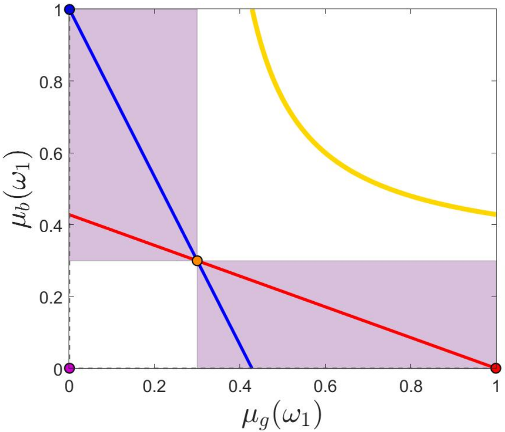


Financial Application

What if...? Perfectly revealing stories tailored for each group

1. If positive (negative) experience, the small stock will (will not) outperform the market
   $\rightarrow$ *Investors with good experience?*
2. If negative (positive) experience, the small stock will not (will) outperform the market
   $\rightarrow$ *Investors with bad experience?*

- It does not work: counterproductive!



Financial Application

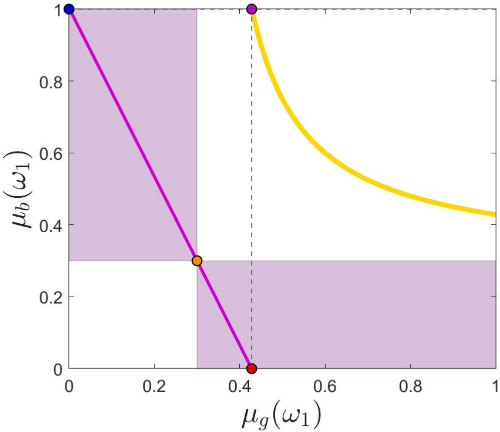
What if...? Perfectly revealing stories tailored for each group

1. If positive (negative) experience, the small stock will (will not) outperform the market
   $\rightarrow$ *Investors with good experience?*
2. If negative (positive) experience, the small stock will not (will) outperform the market
   $\rightarrow$ *Investors with bad experience?*

- It does not work: counterproductive!

Optimization

- Focus on the largest group: investors with negative experience (same as above)
- Choose the narrative adopted by the other group without being counterproductive
  bad experience occurs only if the small stock is bad, but good experience can happen in both cases



Conclusion

- I explore whether it is possible to persuade others only by providing interpretations of unknown events
- My results show that not only it is possible, but also ex-ante persuasion via storytelling can lead the receiver to hold incoherent narratives
- Allowing for multiple stories to be communicated, I provide a disciplined relaxation of the Bayes-plausibility constraint
- The paper discusses how this model uncovers a mechanism common to inter-personal (conflict of interest in financial market, polarization, and nudge) or intra-personal phenomena (commitment)

Thank you!

chiara.aina@econ.uzh.ch

Applications Default Model

Appendix

Literature Review

Literature on Narratives

- Schwartzstein & Sunderam (2021, hereafter SS21): building block of this project
- Eliaz and Spiegler (2020): formalization of narratives as causal models (directed acyclical graphs) to understand public-policy debates
- Benabou, Falk, & Tirole (2018): investigate the role of narratives and imperatives in moral reasoning
- Eliaz, Spiegler, & Thysen (2019): sender-receiver model in which persuaders seek to influence receivers' understanding of messages
- Barron and Powell (2018): theoretical analysis of markets for rhetorical services

Literature on Persuasion

- Kamenica & Gentzkow (2011): Bayesian persuasion model
- Alonso & Camara (2016), Galperti (2019): generalization of Bayesian Persuasion
  - These models are about providing information fixing a signal generating process
- Levy & Razin (2021): Persuasion with Correlation Neglect

Back

Appendix: Equivalent Representation

Back

Lemma

(i) For each vector of posterior beliefs $\boldsymbol{\mu} \in \mathcal{B}$, there exists a model that induces $\boldsymbol{\mu}$

(ii) Each model m induces a vector of posterior beliefs $\boldsymbol{\mu}^m \in \mathcal{B}$

(i) For each $\boldsymbol{\mu} \in \mathcal{B}$, there exists a model that induce $\boldsymbol{\mu}$

- Consider $\boldsymbol{\mu} \in \mathcal{B}$. Then, there exists a distribution $\sigma \in \Delta(S)$ such that $\sum_s \mu_s(\omega)\, \sigma(s) = \mu_0(\omega)$.
- For each σ, define a model such that, for each s and ω,

$$\pi^\sigma(s|\omega) = \frac{\mu_s(\omega)\ \sigma(s)}{\sum_{s'} \mu_{s'}(\omega)\ \sigma(s')}.$$

- Notice that the fit of such a model is, for each signal, $\Pr^\sigma(s) = \sigma(s)$

$$\Pr^\sigma(s) = \sum_\omega \mu_0(\omega)\ \pi^\sigma(s|\omega) = \sum_\omega \left(\sum_{s'} \mu_{s'}(\omega)\ \sigma(s')\right)\left(\frac{\mu_s(\omega)\ \sigma(s)}{\sum_{s'} \mu_{s'}(\omega)\ \sigma(s')}\right) = \sigma(s) \sum_\omega \mu_s(\omega) = \sigma(s)$$

- The posterior attached to state ω conditional on signal s induced by the model σ is

$$\mu_s^\sigma(\omega) = \frac{\mu_0(\omega)\pi^\sigma(s|\omega)}{\Pr^m(s)} = \frac{\mu_0(\omega)}{\sigma(s)}\left(\frac{\mu_s(\omega)\ \sigma(s)}{\sum_{s'} \mu_{s'}(\omega)\ \sigma(s')}\right) = \frac{\sum_{s'} \mu_{s'}(\omega)\ \sigma(s')}{\sigma(s)} \frac{\mu_s(\omega)\ \sigma(s)}{\sum_{s'} \mu_{s'}(\omega)\ \sigma(s')} = \mu_s(\omega)$$

Appendix: Equivalent Representation Back

(ii) Each model m induces a vector of posterior beliefs that is Bayes-consistent: $\boldsymbol{\mu}^m \in \mathcal{B}$

- It is enough to show that there exists a signal distribution such that the Bayes-consistency constraint holds
- Consider as the distribution of signals the fits of the model m conditional on each signal: given that $m \in [\Delta(S)]^\Omega$, it holds that it is a proper distribution with $\sum_s \Pr^m(s) = 1$
- Then, for every $\omega \in \Omega$,

$$\begin{aligned}
\sum_{s \in S} \Pr^m(s)\ \mu_s^m(\omega) &= \sum_{s \in S} \Pr^m(s)\ \frac{\mu_0(\omega)\pi^m(s|\omega)}{\Pr^m(s)} \\
&= \sum_{s \in S} \mu_0(\omega)\pi^m(s|\omega) \\
&= \mu_0(\omega) \sum_{s \in S} \pi^m(s|\omega) = \mu_0(\omega)
\end{aligned}$$

- Every vector of posterior beliefs induced by a model satisfies Bayes-consistency

Equivalent Representation: Binary Case

Corollary

In the binary signal and binary state, for each vector of posterior beliefs $\boldsymbol{\mu} \in \mathcal{B} \backslash \{\boldsymbol{\mu}^{\varnothing}\}$ with $\boldsymbol{\mu}^{\varnothing} = (\mu_0, \mu_0)$, there exists a unique model m that induces $\boldsymbol{\mu}$

Proof

- To show the uniqueness of a model associated to a Bayes-consistent vector of posterior in the binary signal and binary state case, it is enough to show that there exists only a distribution over the signal space such that a vectors of posterior is Bayes-consistent
- Let $(\sigma_{s_1}, \sigma_{s_2}) = (\sigma, 1 - \sigma)$. For each state ω, the Bayes-consistency condition implies that

  $$\mu_0(\omega) = \sigma \mu_{s_1}(\omega) + (1 - \sigma)\mu_{s_2}(\omega).$$

  Then, it holds that $\sigma = \frac{\mu_0(\omega) - \mu_{s_2}(\omega)}{\mu_{s_1}(\omega) - \mu_{s_2}(\omega)}$.
- Hence, $(\sigma_{s_1}, \sigma_{s_2})$ is a signal distribution over signals if either (i) $\mu_{s_1}(\omega) > \mu_0(\omega) > \mu_{s_2}(\omega)$, or (ii) $\mu_{s_1}(\omega) < \mu_0(\omega) < \mu_{s_2}(\omega)$. These two conditions are equivalent to $\mu \in F \backslash \{\mu^{\varnothing}\}$ in the binary case.

Back

Appendix: Isofit

Back

Isofit: set of vectors of posterior beliefs that are induced by models that have the same fit conditional on every signal realization

For each $\varphi \in \Delta(S)$,

$$I(\varphi) = \left\{ \boldsymbol{\mu} \in [\Delta(\Omega)]^S : \exists m \in \mathcal{M} \text{ such that } \boldsymbol{\mu}^m = \boldsymbol{\mu} \text{ and } \forall s \in S,\ \mathrm{Pr}^m(s) = \varphi_s \right\}$$
$$= \left\{ \boldsymbol{\mu} \in [\Delta(\Omega)]^S :\ \forall \omega \in \Omega,\ \mu_0(\omega) = \sum_{s \in S} \varphi_s \mu_s(\omega) \right\}$$

Binary case

- Consider the Bayes-consistency constraint for ω_1 for $\boldsymbol{\mu}^m$:

$$\mu_0(\omega_1) = \mathrm{Pr}^m(s_1)\ \mu^m_{s_1}(\omega_1) + \mathrm{Pr}^m(s_2)\ \mu^m_{s_2}(\omega_1)$$

- Re-arrange:

$$\mu^m_{s_2}(\omega_1) = \frac{\mu_0(\omega_1)}{\mathrm{Pr}^m(s_2)} - \frac{\mathrm{Pr}^m(s_1)}{\mathrm{Pr}^m(s_2)} \mu^m_{s_1}(\omega_1)$$

- All the models with the same fit $(\mathrm{Pr}^m(s_1), \mathrm{Pr}^m(s_2))$ corresponds to points on this line
- Slope $-\frac{\mathrm{Pr}^m(s_1)}{1-\mathrm{Pr}^m(s_1)}$: the higher $\mathrm{Pr}^m(s_1)$, the steeper the line

Proof Lemma

Back

(i) For every $p \in \left[0, \bar{\delta}_s(\mu_s)^{-1}\right]$, there exists a model inducing μ_s conditional on s with fit $\Pr^m(s) = p$

- Construct $\boldsymbol{\mu}$ such that (i) μ_s is induced conditional on s, and (ii) for each state ω, there exists $\sigma(s') \in \Delta(S)$ with the additional property $\sigma(s) = p$ such that Bayes-consistency holds:

$$\sum_{s'} \mu_{s'}(\omega)\ \sigma(s') = \mu_s(\omega)\ \sigma(s) + \sum_{s' \neq s} \mu_{s'}(\omega)\ \sigma(s') = \mu_0(\omega). \tag{a}$$

- By Lemma 1, there exists a model that induce this Bayes-consistent vector of posteriors with fit p
- Given the many degrees of freedom, there exists multiple vectors of posteriors that satisfy condition (a) as long as, for each ω,

$$\mu_0(\omega) - \mu_s(\omega)\ p = \sum_{s' \neq s} \mu_{s'}(\omega)\ \sigma(s') \geq 0. \tag{b}$$

- For instance, fix a signal $s'' \neq s$ and, for each ω, let $\mu_{s''}(\omega) = \frac{\mu_0(\omega) - p\ \mu_s(\omega)}{1-p}$
  - Condition (a) is satisfied for the distribution $\sigma(s')$ such that $\sigma(s) = p$, $\sigma(s'') = 1 - p$, and $\sigma(s') = 0$ for all the other signals
  - Condition (b) is implied by $p \in \left[0, \bar{\delta}_s(\mu_s)^{-1}\right]$
    As the condition has to hold for every state, it holds that

$$p \leq \frac{\mu_0(\omega)}{\mu_s(\omega)} \leq \left(\frac{1}{\max_\omega \frac{\mu_s(\omega)}{\mu_0(\omega)}}\right)^{-1} = \bar{\delta}_s(\mu_s)^{-1}$$

Proof Lemma

Back

(ii) Every model inducing μ_s conditional on s has fit $\Pr^m(s) \in \left[0, \bar{\delta}_s(\mu_s)^{-1}\right]$

- Consider an arbitrary model inducing μ_s conditional on s
- It follows from Bayes rule that the fit of any m inducing the target $\mu_s^m = \mu_s$ conditional on s must be such that, for every ω

$$\Pr^m(s) = \frac{\mu_0(\omega)}{\mu_s(\omega)}\pi^m(s|\omega)$$

- Notice that if $\pi^m(s|\omega) = 0$ the fit equals 0 (minimal fit). Instead, if $\pi^m(s|\omega) = 1$, it follows that

$$\Pr^m(s) \leq \frac{\mu_0(\omega)}{\mu_s(\omega)}$$

- Because this holds for every state, the maximal fit for μ_s is the minimum of the ratio across states, which equals the reciprocal of the maximal movement for μ_s:

$$\min_\omega \frac{\mu_0(\omega)}{\mu_s(\omega)} = \frac{1}{\max_\omega \frac{\mu_s(\omega)}{\mu_0(\omega)}} = \bar{\delta}_s(\mu_s)^{-1}$$

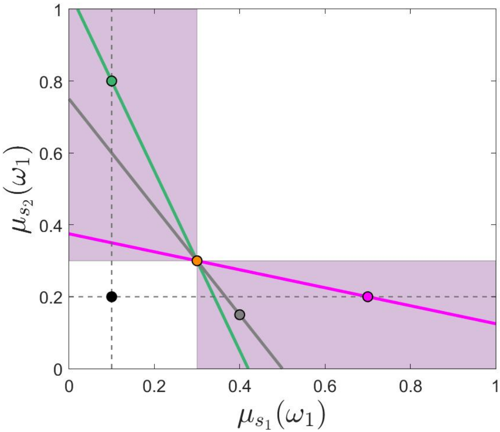
- The fit of a model that induces the target posterior can only take values in $\left[0, \bar{\delta}_s(\mu_s)^{-1}\right]$

Many Models Back

The sender knows the resulting vector of posterior beliefs of the receiver

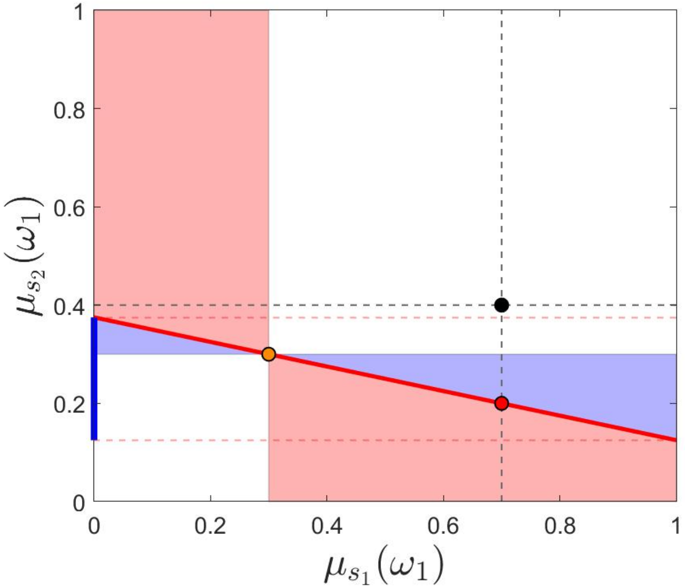
- How? He anticipates which model the receiver adopts conditional on each signal

- *Which model lies on the steeper isofit line?*
  This induces the posteriors given s_1
- *Which model lie on the flatter isofit line?*
  This induces the posteriors given s_2
- *What is the resulting vector of posteriors?*
- *What if there are more than two models?*



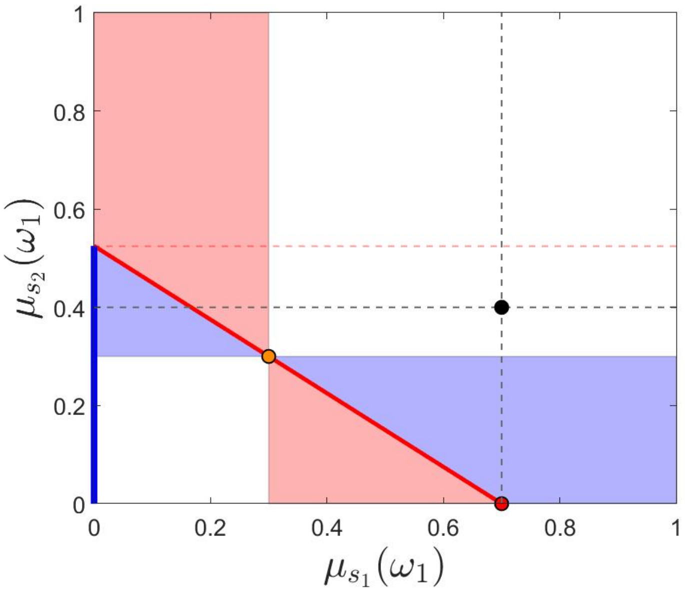
Many Models

Graphical Intuition



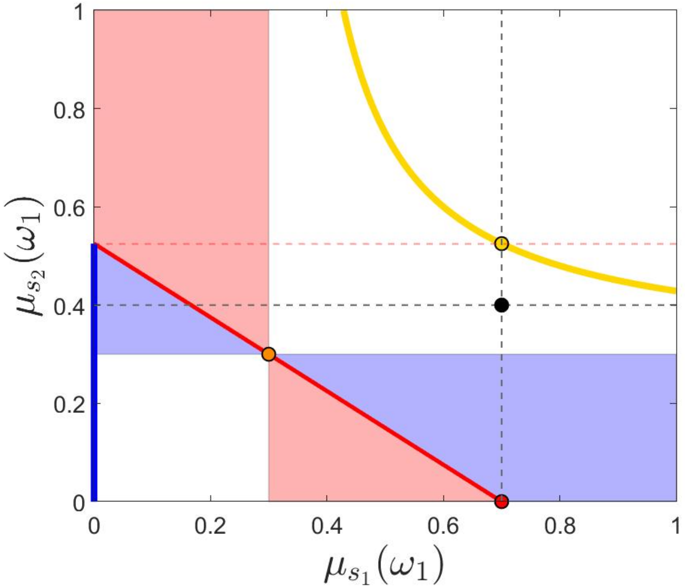
Many Models

Graphical Intuition



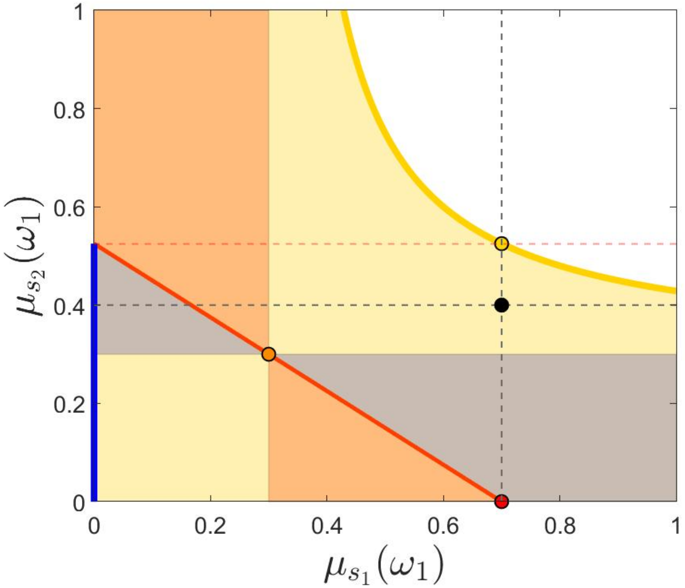
Many Models

Graphical Intuition



Many Models

Graphical Intuition



Back

Proof Theorem

Back

- Take an arbitrary vector of posterior beliefs $\boldsymbol{\mu}$
  - To induce $\boldsymbol{\mu}$, construct a set of $K = |S|$ models $(m_k)_{k=1}^{K}$ such that each model m_k is tailored to induce the target distribution μ_{s_k} conditional on the signal s_k. This implies two conditions on each m_k: (i) $\mu_{s_k}^{m_k} = \mu_{s_k}$, and (ii) $\Pr^{m_k}(s_k) \geq \Pr^{m_j}(s_k)$ for each $k \neq j$
- Assume $\boldsymbol{\mu} \in \mathcal{F}$: I show that there exists a set of models inducing $\boldsymbol{\mu}$
  - For each model m_k, I specify the vector of posteriors $\boldsymbol{\mu}^{m_k}$ and the induced fit levels $(\Pr^{m_k}(s))_{s\in S}$: the corresponding distribution of posteriors corresponds to a unique model
  - For each model m_k, specify the following posteriors and fit levels: if $s = s_k$, set $\mu_s^{m_k} = \mu_{s_k}$ and $\Pr^{m_k}(s) = \bar{\delta}(\mu_{s_k})^{-1}$; otherwise , for each signal and state, set

$$\mu_s^{m_k}(\omega) = \frac{\mu_0(\omega) - \bar{\delta}(\mu_{s_k})^{-1}\mu_{s_k}(\omega)}{1 - \bar{\delta}(\mu_{s_k})^{-1}}, \qquad \Pr^{m_k}(s) = \left(\frac{1 - \bar{\delta}(\mu_{s_k})^{-1}}{\sum_{s\neq s_k} \bar{\delta}(\mu_s)^{-1}}\right) \bar{\delta}(\mu_{s_k})^{-1}$$

    Equivalent to an information structure with binary signal s_k and s_{-k}
  - Each tailored model is chosen conditional on the signal is tailored to because

$$\Pr^{m_k}(s_k) = \bar{\delta}(\mu_{s_k})^{-1} \geq \underbrace{\left(\frac{1 - \bar{\delta}(\mu_{s_k})^{-1}}{\sum_{s\neq s_k} \bar{\delta}(\mu_s)^{-1}}\right)}_{\leq 1 \text{ since } \boldsymbol{\mu}\in\mathcal{F}} \bar{\delta}(\mu_{s_k})^{-1} = \Pr^{m_j}(s_k)$$

Proof Theorem

Back

- Assume $\boldsymbol{\mu} \notin \mathcal{F}$: $\sum_{s \in S} \bar{\delta}(\mu_s)^{-1} < 1$, equivalent to $\bar{\delta}(\mu_{s_k})^{-1} < 1 - \sum_{i \neq k} \bar{\delta}(\mu_{s_i})^{-1}, \forall k$
  - If it were to exist a set of models inducing the target $\boldsymbol{\mu}$, each tailored model m_k inducing the posterior μ_{s_k} has to be adopted conditional on s_k
  - Thus, it must hold that $\mathrm{Pr}^{m_k}(s_k) \geq \mathrm{Pr}^{m_j}(s_k)$ for each $j \neq k$
  - Notice that

$$\mathrm{Pr}^{m_j}(s_k) = 1 - \sum_{i \neq k} \mathrm{Pr}^{m_j}(s_i) \geq 1 - \sum_{i \neq k} \bar{\delta}(\mu_{s_i})^{-1},$$

    since for every other signal the fit must be lower than the maximal fit associated to the target posterior conditional on that signal, i.e. $\mathrm{Pr}^{m_j}(s_i) \leq \mathrm{Pr}^{m_i}(s_i) \leq \bar{\delta}(\mu_{s_i})^{-1}$ for every i
  - Contradiction:

$$1 - \sum_{i \neq k} \bar{\delta}(\mu_{s_i})^{-1} > \bar{\delta}(\mu_s)^{-1} \geq \mathrm{Pr}^{m_k}(s_k) \geq \mathrm{Pr}^{m_j}(s_k) \geq 1 - \sum_{i \neq k} \bar{\delta}(\mu_{s_i})^{-1}$$

  - It is not possible to construct a set of models to induce $\boldsymbol{\mu} \notin \mathcal{F}$

Comparative Statics with Respect to the Priors

Back

Proposition

If $\min_{\omega \in \Omega} \mu_0(\omega) \geq \frac{1}{|S|}$, all vectors of posterior beliefs are feasible

- The more signals, the more manipulability of the receiver's beliefs
- The more uniform the priors, the more belief manipulability
  The minimal prior across states is the lower bound for the maximal fit for any updated posteriors starting from given priors, i.e., $\bar{\delta}(\mu_s)^{-1} \leq \min_{\omega \in \Omega} \mu_0(\omega)$ for any μ_s

Corollary

If $|S| \geq |\Omega|$ and $\mu_0(\omega) = \frac{1}{|\Omega|}$ for every $\omega \in \Omega$, all vectors of posterior beliefs are feasible

Comparative Statics with Respect to the Priors: Binary Case

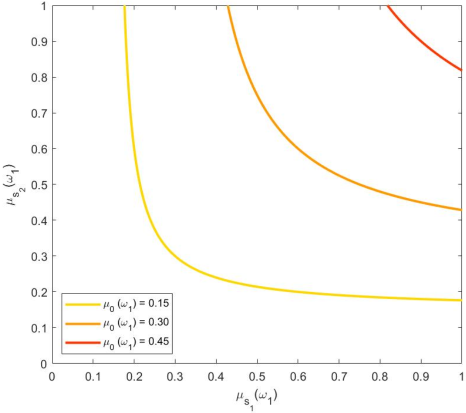
Back

Let $\mu_{0,\varepsilon} = \left(\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon\right)$ and $\mathcal{F}_\varepsilon$ the set of the feasible vectors of posteriors with respect to the prior $\mu_{0,\varepsilon}$

Proposition

For $\varepsilon' < \varepsilon''$, it holds that $\mathcal{F}_{\varepsilon''} \subseteq \mathcal{F}_{\varepsilon'}$

- The closer the receiver's priors are to 50-50, the more she can be manipulated



Appendix: Firehose of Falsehood Back

- These stories are able to shift any receiver that has prior higher than 33%
- How? The conspiracy theory is adopted when the reported majority is not for Z and the just narrative is adopted when the reported majority is in favor of Z

$$\Pr^{m_1}(R) > \Pr^{m_2}(R)$$

- To see this, calculate for which prior p this is the case

$$p\ 99\% + (1-p)\ 1\% \geq p\ 1\% + (1-p)\ 50\%$$

$$p \geq 33\%$$

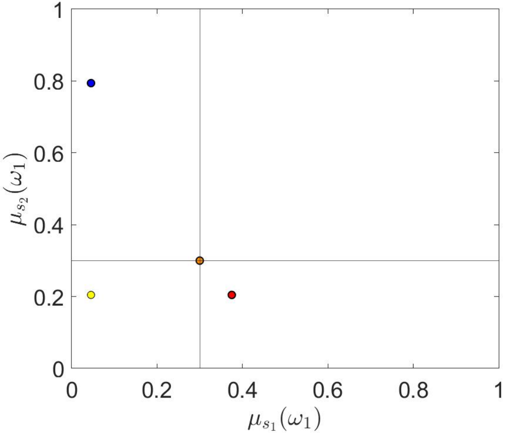
- Depending on how priors are distributed in the population, Z might be able to leverage on this bundle of truth and conspiracy theory to be elected

Conflicting Narratives Back

Binary Case

Conflicting Narratives m, m' if $\pi^m(s_1|\omega_1) > \pi^m(s_1|\omega_2)$ and $\pi^{m'}(s_1|\omega_2) > \pi^{m'}(s_1|\omega_1)$

- m implies that $\mu^m_{s_1}(\omega_1) > \mu_0(\omega_1)$ and $\mu^m_{s_2}(\omega_1) < \mu_0(\omega_1)$
  - South-East quadrant
- m' implies that $\mu^{m'}_{s_1}(\omega_1) < \mu_0(\omega_1)$ and $\mu^{m'}_{s_2}(\omega_1) > \mu_0(\omega_1)$
  - North-West quadrant
- Together induce always a vector of beliefs that is Bayes-inconsistent



Belief Polarization

Back

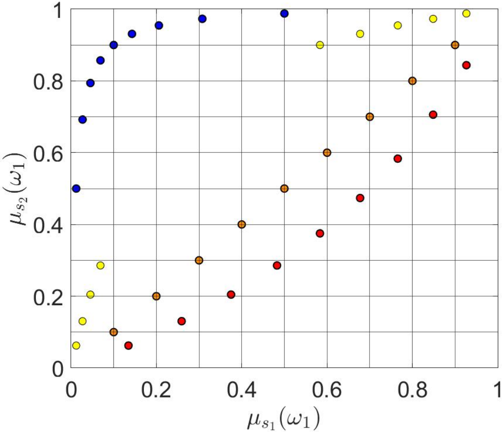
Binary Case

- Whenever two conflicting narratives are communicated, belief polarization occurs
- Depending on the prior, Bayes-consistency is violated differently

Proposition

For each pair of conflicting stories, there exists a threshold in prior p such that for every signal s, it holds that

1. $\mu_s(\omega_1) < \mu_0(\omega_1)$ if $\mu_0(\omega_1) < p$
2. $\mu_s(\omega_1) > \mu_0(\omega_1)$ if $\mu_0(\omega_1) > p$



Solving the Sender's Problem

- All the information the sender needs to learn how to maximize his value through stories is the receiver's prior and the number of models she accepts
- Assume that, if the sender does not communicate any model to the receiver, she does not update her beliefs, discarding the realized signal: for each signal s, $\mu_s = \mu_0$
- Persuasion is beneficial if there exists a feasible vector of posterior beliefs $\mu \in \mathrm{I}$ such that its value is higher than the value of the prior:

$$V(\mu) \geq V(\mu^{\varnothing})$$

- Binary case: graphical solution, plotting the sender's value on $\mathcal{F}$

Applications [Back]

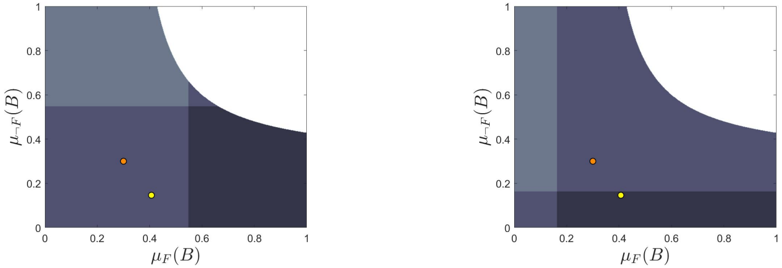
1. Firehose of Falsehood: model of Russian propaganda based on a large number of possibly contradictory and mutually inconsistent messages (Paul & Matthew, 2016)
    - With conflicting narratives, belief polarization occurs: there is a threshold in prior such that voters with prior higher (lower) than the threshold prior would hold extreme high (low) posteriors regardless the election outcome
2. Finance: with misaligned incentives an advisor can effectively manipulate investors to invest in his preferred asset
    - Even without knowing investors' relavant information such as past experience, the advisor communicates ad-hoc stories to maximize his return
3. Nudging: proposing ad-hoc narratives can be seen as a soft intervention to influence in a not coercive manner choices of an agent with the purpose of increase her welfare
    - Confidence manipulation by a paternalistic planner, via distorting the interpretation of signals, is optimal to influence the agent's behavior in a risky task
4. Intra-personal Phenomena: a mechanism through which the individual may distort his beliefs without assuming exogenous parameter of memory loss, inattention, first-impression, etc.
    - In a multi-selves model, an agent has incentives to distort his self-confidence in order to offset his time inconsistent preferences

Applications Back

1. Firehose of Falsehood
2. Finance
3. Nudging: proposing ad-hoc narratives can be seen as a soft intervention to influence in a not coercive manner choices of an agent with the purpose of increase her welfare
   - Proposing ad-hoc narratives can be seen as a soft intervention adopted by a paternalistic planner to influence in a not coercive manner choices of an agent with the purpose of increase her welfare
   - Confidence manipulation by a paternalistic planner, via distorting the interpretation of signals, is optimal to influence the agent's behavior in a risky task
4. Intra-personal phenomena

Nudging Risk Attitude

- States: $\{B, \neg B\}$, where "B" is the event that Arthur is brave enough
- Signals: $\{F, \neg F\}$, where "F" is the event that Ford gives Arthur a positive feedback
- Arthur has to choose whether to go on adventure (high reward if B, low if $\neg B$) or not (medium reward)
- Marvin wants Arthur to undertake the adventure only conditional on a positive feedback
- Marvin believes Ford would bias his advice optimistically:
  low false negative $\pi^t(F|B) = 0.8$ but high false positive $\pi^t(F|\neg B) = 0.5$



(a) Risk adverse (b) Risk seeking

- It is beneficial to distort the agent's beliefs to be moderately overconfident (underconfident) if risk averse (risk seeking)

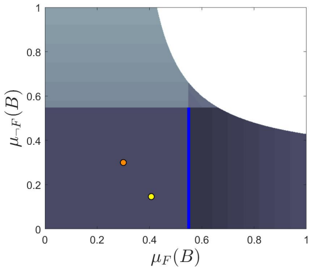
Refinment Back

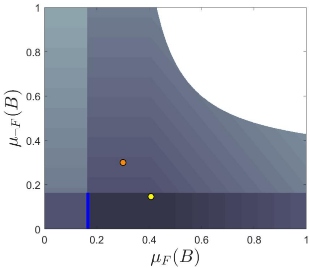
Appendix: Nudging Risk Attitude

- Assume that belief distortion bears some psychological costs for the sender, such as disappointment aversion
- Disappointment = the positive gap between the expected payoff calculated with the induced beliefs and the expected payoff with the true model
- The resulting sender's value function

$$V(\mu) = \sum_{s \in \{F, \neg F\}} \Pr^t(s) \left[\mathbb{E}_{\mu^t}\left[U^S(a(\mu))\right] - k \cdot \underbrace{\max\left\{0, \mathbb{E}_{\mu^t}\left[\pi^R(a(\mu))\right] - \mathbb{E}_{\mu}\left[\pi^R(a(\mu))\right]\right\}}_{disappointment} \right]$$

where k is a sensitivity parameter to disappointment, and $\pi_R : \{a, \neg a\} \to \mathbb{R}$ is the receiver's payoff





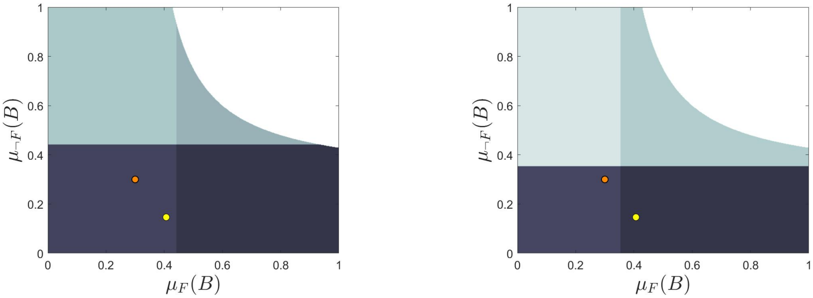
Applications

Back

1. Firehose of Falsehood
2. Finance
3. Nudging
4. Intra-personal phenomena: a mechanism through which the individual may distort his beliefs without assuming exogenous parameter of memory loss, inattention, first-impression, etc.
    - Multiple stories allow motivated reasoning to take root
    - In a multi-selves model, an agent has incentives to distort his self-confidence in order to offset his time inconsistent preferences

Commitment (Bénabou & Tirole, 2002) Model

- Arthur has to decide whether to go on an adventure or not
- Arthur is risk-neutral with quasi-hyperbolic discounting preferences
- Arthur knows that at the moment of the decision the imminent cost of undertaking the adventure c (t = 1) will be more salient than the future reward of a success v if B (t = 2)
- Deep down he believes Ford to be optimistic: $\pi^t(F|B) = 0.8$ and $\pi^t(F|\neg B) = 0.5$



(a) Present Bias (b) No Present Bias

- At date 0, Arthur might have incentives to distort his interpretations of Ford's advice only to overcome his present bias

Back

Appendix: Bénabou & Tirole (2002)

Timing: At $t = 0$, the individual can take an action that potentially affects his information at $t = 1$ with some utility flow. At $t = 1$, he decides whether to take an action with disutility c that, if successful, would yield benefit v at $t = 2$

- Consider a risk-neutral individual with quasi-hyperbolic discounting
- No action a leads to zero utility, hence $U^1(\neg a) = 0$ and $U^0(\neg a) = u_0$
- Utility at $t = 1$ when taking the action conditional on s:

$$U^1(a) = u_1 + \beta\delta \; \mathbb{E}_{\mu_s}[u_2] = -c + \mu_s(success)\; \beta\delta v,$$

  where $\delta \leq 1$ is his discount factor and $\beta > 0$ is his present bias.
- At $t = 1$, the action is optimal if $\mu_s(success) \geq \frac{c}{\beta\delta v}$.
- Utility at date 0 when taking the action:

$$U^0(a) = u_0 + \beta \; \mathbb{E}_{\mu^t}\left[\delta \; u_1 + \delta^2 \; u_2\right] = u_0 + \beta\delta\big(-c + \mu_s^t(success)\; \delta v\big),$$

- At date 0, the action is optimal if $\mu_s^t(success) \geq \frac{c}{\delta v}$, lower if he suffers from present bias $\beta < 1$
- The probability of success, discussed as his self-confidence by the authors, may depend on either new information received or forgotten

Back

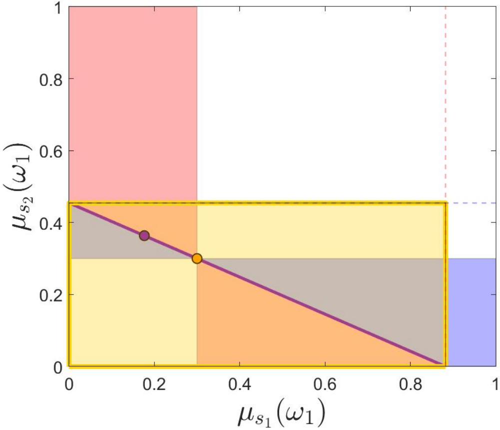
Appendix: Receiver's Default Model Back

- The receiver has endowed with a default model d, known by the sender
- More challenging for the sender to induce posteriors:
  conditional on the signal, the receiver adopts the posterior distribution induced by the sender's model only if the latter has higher fit than the default model

Proposition $|M| \geq |S|$

The set of feasible vectors of posterior beliefs is

$$\mathcal{F}^d = \Big\{ \mu \in [\Delta(\omega)]^S : \forall \omega \in \Omega, s \in S, \delta_s(\omega; \mu_s)\mathrm{Pr}^d(s) \leq 1 \Big\}$$



Appendix: Receiver's Default Model Back

The set of the feasible vectors of posteriors without the default model is the union of all sets of the feasible vectors of posteriors with default model for every default model

Proposition

The union of $\mathcal{F}^d$ for all possible default model is $\mathcal{F}$, i.e.,

$$\bigcup_{d \in \mathcal{M}} \mathcal{F}^d = \mathcal{F}.$$