Persuasion with Correlation Neglect

Gilat Levy
LSE

Inés Moreno de Barreda
University of Oxford

Ronny Razin
LSE

Introduction

The amount of information we are exposed to and its complexity opens up opportunities to manipulate the beliefs of voters and consumers.

- Brexit campaign: The UK information Commissioner's Office (ICO) is investigating whether Vote Leave (the official pro-Brexit campaign during the EU referendum), BeLeave and others shared data to identify audiences and target adds. In parallel, the Electoral Commission in the UK is investigating allegations of unlawful coordination between Vote Leave and BeLeave.
- Different thinktanks are often called to provide their expertise on issues in the public debate. However, if sharing the same funding donors, those thinktanks might be coordinating their reports.
- Cagé, Hervé and Viaud (2017) show that most online content is not original and repackaged and repeated without references.

Introduction

- While campaigns and information may be correlated, voters or consumers may be unaware that this is the case.
- The literature has recently documented many environments where "correlation neglect" arises. (Ortoleva and Snowberg (2015), Eyster and Weizsker (2011), Kallir and Sonsino (2009) and Enke and Zimmermann (2018))
- correlation neglect: Individuals treat information sources as if they are conditionally independent.

Introduction

We analyze a model of persuasion when the receiver has correlation neglect.

Research Question:

- What is the scope of manipulation when the receiver has correlation neglect?
- How can a strategic sender use correlation strategically to manipulate a receiver who believes that information sources are independent?

Information design problem:

- The sender "designs" m signals to influence the receiver.
- The receiver understands the marginal information structure of each signal separately,
- but believes that the joint information structure over all signals satisfies conditional independence.

Introduction

General Results:

- The sender can manipulate over and above what she could get from a rational receiver.
  - In particular, the sender can move the expectation of the posterior of the receiver in any direction.
- For a fixed number of signals, the sender's ability to manipulate is bounded.
  - For example, the probability with which the sender can convince the receiver that the state *is* ω is at most the prior probability of the state.
- As the number of signals increases, the sender can approximate her First Best.
  - In particular, she can approximate any posterior beliefs, where those posterior may differ across states.

Introduction

Optimal Information Structure:

For the case of state-independent preferences:

- Concavification provides an upper bound on the optimal value.
  - However, such upper bound might not be achievable: concavification might point to correlation structures that are not feasible.
- We analyse a modified problem in which the sender uses full correlation:
  - The solution to the modified problem converges to the optimal solution as the number of signals is large.
  - Under supermodular utilisites full correlation is optimal even with finite signals.
  - Full correlation is not always optimal. For submodular utilities, some negative correlation is optimal.

The Model

- There are two players, a sender (she) and a receiver (also she).
- There is a finite set of states of the world, $\Omega = \{\omega_1, ..., \omega_n\}$, with commonly known interior prior $p \in \Delta(\Omega)$
- The sender designs an information structure, which consists of m distinct signals.
- The receiver observes the realisation of the signals $\{s_1, ..., s_m\}$ and chooses an action $a \in A$ where A is compact.
- Given action a and state ω, the receiver gets utility $u(a, \omega)$ and the sender gets $v(a, \omega)$. We assume that:
  - There exists an optimal action for the receiver
  - v evaluated at the optimal action for the receiver is upper hemi-continuous in the receiver's information.

Signal Structures

- An information structure with m signals is defined by $\{\mathbf{S}, \{q\cdot \mid \omega)\}_{\omega\in\Omega}\}$ where,
  - $\mathbf{S} = \prod_{i=1}^{m} S_i$ with S_i finite, and
  - $q(\cdot \mid \omega) \in \Delta(\mathbf{S})$ is a joint distribution conditional on ω.
- We denote by $\{S_i, \{q_i(\cdot \mid \omega)\}_{\omega\in\Omega}\}$ the marginal information structure for signal i.

Signal Structures

As in the Bayesian Persuasion literature it will be convenient to interpret a signal as a distribution over posteriors:

- A realisation s_i of signal i generates a posterior

$$\mu_{s_i}(\omega) = \frac{p(\omega) q_i(s_i \mid \omega)}{\sum_{\nu \in \Omega} p(\nu) q_i(s_i \mid \nu)} \quad \forall \omega \in \Omega$$

- Hence signal i generates a family of conditional distributions over posteriors $\{\tau_i(\cdot \mid \omega)\}_{\omega \in \Omega} \subset \Delta(\Delta(\Omega))$ where

$$\tau_i(\mu_{s_i} \mid \omega) = q_i(s_i \mid \omega)$$

(and uncoditional distribution $\tau_i(\mu) = \sum_{\upsilon \in \Omega} p(\upsilon) \tau_i(\mu \mid \upsilon)$)

Signal Structures

As the sender might want to correlate signals depending on the realisation of the state, it will be important to characterise the possible conditional distributions that might be generated by a signal:

Lemma: *The family of conditional distributions over posteriors* $\{\tau_i(\cdot \mid \omega)\}_{\omega \in \Omega}$ *is inducible by a signal given a Bayesian updater if and only if for any* $\mu \in \Delta(\Omega)$ *with* $\tau_i(\mu) > 0$,

$$\mu(\omega) = \frac{p(\omega)\tau_i(\mu \mid \omega)}{\sum_{\upsilon} p(\upsilon)\tau_i(\mu \mid \upsilon)} \tag{1}$$

We will therefore abstract from the explicit distributions of the signals and work directly with conditional distributions over posteriors.

Correlation Neglect

We assume that the receiver

- perfectly understands the marginal distribution of each signal,
- treats all the signals as conditionally independent.

Lemma: *Given m signals, consider $\boldsymbol{\mu} = (\mu_1, ..., \mu_m)$ a vector of posteriors. Then, the posterior of a receiver with correlation neglect is given by:*

$$\mu^{CN}(\omega \mid \boldsymbol{\mu}) = \frac{\frac{\prod_{i=1}^{m} \mu_i(\omega)}{p(\omega)^{m-1}}}{\sum_{\upsilon \in \Omega} \frac{\prod_{i=1}^{m} \mu_i(\upsilon)}{p(\upsilon)^{m-1}}}$$

- Note that if $m = 1$, then $\mu^{CN}(\mu) = \mu$ and the receiver acts as a completely rational agent.

Receiver's choice

- Once the receiver updates her beliefs, she chooses her optimal action given those beliefs.
- Given a posterior $\mu \in \Delta(\Omega)$, a_μ is the receiver's optimal choice:

$$a_\mu \in \arg\max_{a \in A} \sum_{\omega \in \Omega} \mu(\omega) u(a, \omega).$$

- In case of multiple solutions we assume she chooses the sender-preferred solution in order to guarantee existence of the sender's problem.

The Sender's Problem

The sender's problem is to choose $\{\tau(\cdot \mid \omega)\}_{\omega\in\Omega} \subset \Delta(\Delta(\Omega)^m)$ to solve:

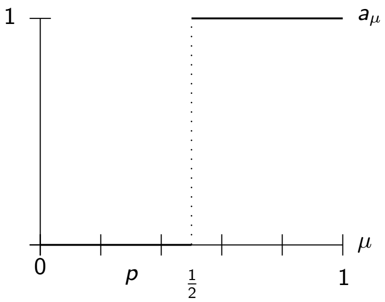
$$\max_\tau \hat{v}(\tau) = \sum_{\omega\in\Omega} p(\omega) \sum_{\boldsymbol{\mu}\in\Delta(\Omega)^m} \tau(\boldsymbol{\mu} \mid \omega) v(a_{\mu^{CN}(\boldsymbol{\mu})}, \omega)$$

$$s.t. \quad \mu^{CN}(\omega \mid \boldsymbol{\mu}) = \frac{\frac{\prod_{i=1}^m \mu_i(\omega)}{p(\omega)^{m-1}}}{\sum_{\upsilon\in\Omega} \frac{\prod_{i=1}^m \mu_i(\upsilon)}{p(\upsilon)^{m-1}}}$$

$$and \quad \mu_i(\omega) = \frac{p(\omega)\tau_i(\mu_i|\omega)}{\sum_{\upsilon\in\Omega} p_\upsilon \tau_i(\mu^i|\upsilon)} \quad \forall \mu_i \text{ s.t. } \tau_i(\mu_i) > 0, \quad \forall i \in \{1, ..., m\}$$

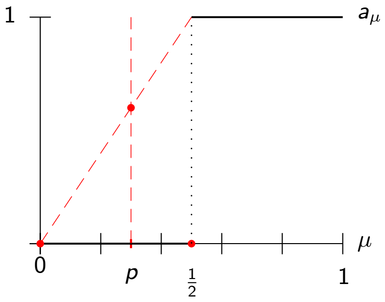
Binary Example

- $\Omega = \{0, 1\}$, with prior $p = prob(\omega = 1) < \frac{1}{2}$.
- Reciever's optimal action is $a = 1$ if $\mu^{CN}(1 \mid \boldsymbol{\mu}) \geq \frac{1}{2}$ and $a = 0$ otherwise.
- The sender has state independent preferences and prefers $a = 1$ to $a = 0$.



Binary Example

Consider first the case of a single signal:



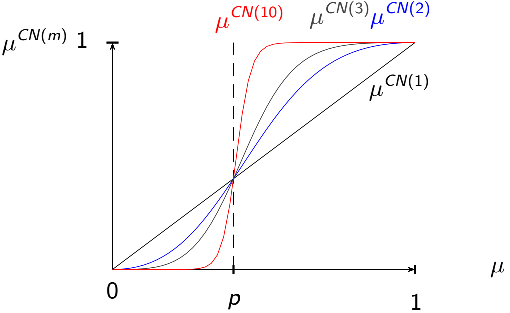
The optimal conditional distribution is:

$$\begin{aligned}\tau(1/2 \mid 1) &= 1 \\ \tau(1/2 \mid 0) &= \frac{p}{1-p},\ \tau(0 \mid 0) = \frac{1-2p}{1-p}.\end{aligned}$$

Binary Example

Suppose that the sender sends m fully correlated signals. The posterior of a receiver with correlation neglect becomes:

$$\mu^{CN}(\mu, ..., \mu) = \frac{\frac{\mu^m}{p^{m-1}}}{\frac{\mu^m}{p^{m-1}} + \frac{(1-\mu)^m}{(1-p)^{m-1}}}$$



There is an amplification effect.

Binary Example

For example for $m = 2$, we can induce $\mu^{CN}(1 \mid \boldsymbol{\mu}) = \frac{1}{2}$ by generating marginal posteriors

$$\mu = \frac{p}{p + (1-p)\left(\frac{p}{1-p}\right)^{\frac{1}{2}}} < \frac{1}{2}$$

But since $\mu < \frac{1}{2}$, it can be induced with higher probability in state 0:

$$\tau(\mu|1) = 1, \quad \tau(\mu|0) = \left(\frac{p}{1-p}\right)^{\frac{1}{2}}, \quad \tau(0|0) = 1 - \left(\frac{p}{1-p}\right)^{\frac{1}{2}}$$

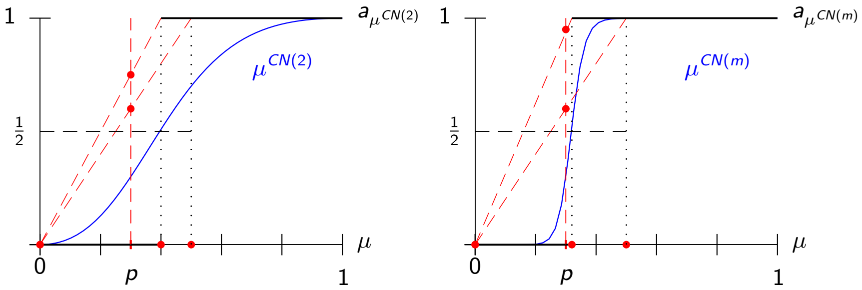
This then leads to a higher expected posterior for the receiver, as now

$$E_\tau(\mu^{CN}(\boldsymbol{\mu})) > p$$

- The sender is able to manipulate the expected posterior of the receiver.

Binary Example

In particular, this will lead to a higher expected utility for the sender:

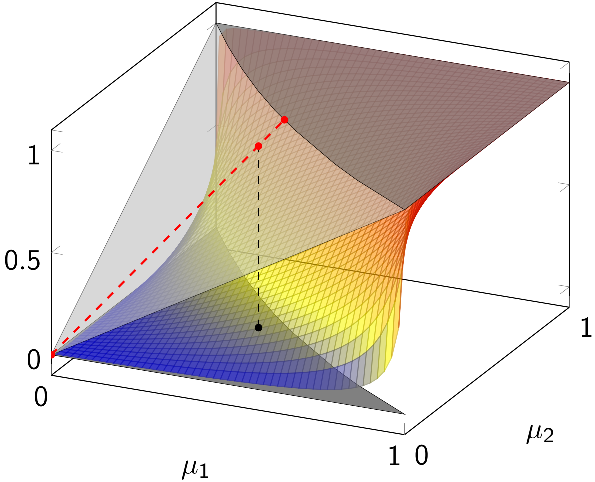


- For fixed m, the sender benefits from correlation, but cannot fully manipulate the receiver.
- As the number of signal increases the sender achieves her first best.

Binary Example

What is the optimal signal structure?

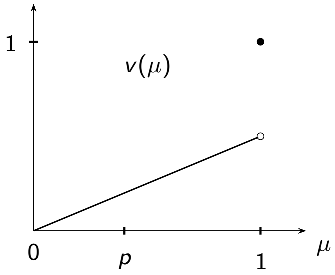
For $m = 2$ signals, the utility of the sender is:



- For the binary example, full correlation is optimal.

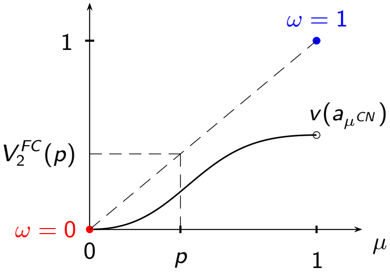
Example 2: Suboptimality of Full Correlation

- $\Omega = \{0, 1\}$, $p(0) = p(1) = \frac{1}{2}$.
- The receiver's optimal action is her expected state: $a_\mu = \mu$
- The sender has state independent preferences, increasing in a continuous for $a < 1$ and with a discontinuity at $a = 1$.



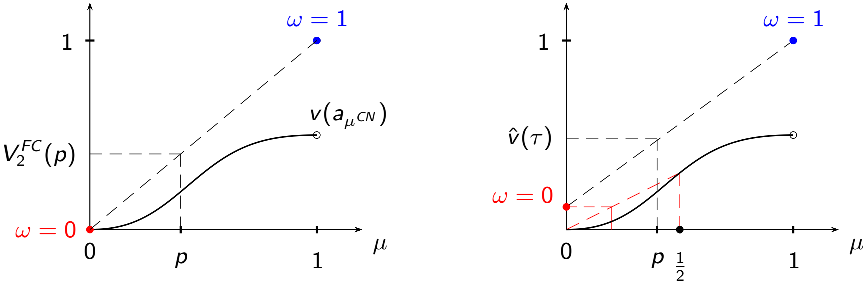
Example 2: Suboptimality of Full Correlation

With full correlation the sender would perfectly reveal the state:



Example 2: Suboptimality of Full Correlation

With full correlation the sender would perfectly reveal the state:



However, for example, in the case of 2 signals, the following information structure is an improvement over full correlation:

$$\tau((1, \tfrac{1}{2})|\omega = 1) = \tau((\tfrac{1}{2}, 1)|\omega = 1) = \tfrac{1}{2}$$
$$\tau((0,0)|\omega = 0) = \tfrac{1}{2} \quad \tau((\tfrac{1}{2}, \tfrac{1}{2})|\omega = 0) = \tfrac{2}{3}$$

General Results

Changing the Expected beliefs:

Proposition: *For any distribution* $q \in \Delta(\Omega)$*, there exists* $\epsilon > 0$ *and a signal structure* τ *with two fully correlated signals, such:*

$$E_\tau(\mu^{CN}) = (1-\epsilon)p + \epsilon q$$

General Results

Bound on Manipulation:

Consider the set

$$M_\omega(\delta) = \{\boldsymbol{\mu} \in \Delta(\Delta(\Omega)^m) \mid \mu^{CN}(\omega \mid \boldsymbol{\mu}) > 1 - \delta\}$$

Proposition: *For any $\omega, \upsilon \in \Omega$ with $\omega \neq \upsilon$,*

$$\tau(M_\omega(\delta) \mid \upsilon) \leq m \frac{1}{p(\upsilon)} \left(\frac{\delta}{1-\delta}\right)^{\frac{1}{m}} \left(\frac{p(\upsilon)}{p(\omega)}\right)^{\frac{m-1}{m}}$$

In particular,

$$\lim_{\delta \to 0} \boldsymbol{\tau}(M_\omega(\delta)) \leq p_\omega$$

General Results

Full manipulation in the limit:

Proposition: *When $v(a, \omega)$ is continuous, the sender can achieve her first best in the limit. That is, given $\{q^\omega\}_{\omega \in \Omega}$, there exists a sequence of signal structures $\{\tau_m\}_{m \in \mathbb{N}}$ indexed by the number of signals, such that for any $\omega \in \Omega$ and any $\epsilon > 0$,*

$$\lim_{m \to \infty} \boldsymbol{\tau}_m(\{\boldsymbol{\mu} \mid |\mu^{CN}(\boldsymbol{\mu}) - q^\omega| < \epsilon\} \mid \omega) = 1$$

Optimal Information Structure

Homogeneous signals:

Proposition: *For any joint information structure τ, and its associated expected utility for the sender $\hat{v}(\tau)$, there exists a joint information structure τ' with homogenous marginals, i.e., $\tau'_i = \tau'_j$, for all $i, j \in \{1, ..m\}$, such that $\hat{v}(\tau') = \hat{v}(\tau)$*

Sketch of Proof: Define $\tau^* = \frac{1}{m!}\sum_\sigma \tau^\sigma$, where τ^σ are permutations of τ:

- τ^* has homogeneous marginals
- $\hat{v}(\tau^*) = \hat{v}(\tau)$
- τ^* satisfies condition (1) and hence can be generated by a signal.

Optimal Information Structure

State-independent Preferences:

- The expected utility of the sender only depends on the receiver's posteriors:

$$\begin{aligned}\hat{v}(\tau) &= \sum_{\boldsymbol{\mu}\in\Delta(\Omega)^m} \tau(\boldsymbol{\mu}) \sum_{\omega\in\Omega} \mu(\boldsymbol{\mu})(\omega) v(a_{\mu^{CN}(\boldsymbol{\mu})}) \\ &= \sum_{\boldsymbol{\mu}\in\Delta(\Omega)^m} \tau(\boldsymbol{\mu}) v(a_{\mu^{CN}(\boldsymbol{\mu})})\end{aligned}$$

- Denote $\tilde{v}(\boldsymbol{\mu}) = v(a_{\mu^{CN}(\boldsymbol{\mu})})$, and $V(\cdot)$ the concavification of $\tilde{v}(\cdot)$.

Proposition: *Given state-independent preferences, $V(\mathbf{p})$ is an upper bound for the expected utility of the sender*

The upper bound might not be reached because it may yield correlation patterns that are not compatible with (1).

Optimal Information Structure

Full Correlation: The Modified Problem

Denote by $\tilde{v}_m^{FC}(\mu) = v(a_{\mu^{CN}(\mu,\ldots,\mu)}$. Suppose that the sender chooses one signal $\tau \in \Delta(\Delta(\Omega))$ to maximise $\sum_{\mu \in \Delta(\Omega)} \tau(\mu)\tilde{v}_m^{FC}(\mu)$, subject to τ being a Bayes-plaussible distribution of posteriors.

Proposition: *Let $v(\cdot)$ a continuous state-independent utility function. Then the solution of the modified problem converges to her first best as the number of signals increases. That is, for any $\epsilon > 0$,*

$$\lim_{m \to \infty} |V_m^{FC}(p) - v(\mu^{FB})| < \epsilon$$

Optimal Information Structure

Full Correlation:

Proposition: *Given $\Omega = \{0, 1\}$ and state-independent preferences, if $v(a_{\mu^{CN}(\boldsymbol{\mu})})$ is supermodular in $\boldsymbol{\mu}$, the optimal solution to the sender's problem is achieved by fully correlating a set of homogeneous signals.*

This result follows from Lorentz (1953) and the result on homogeneity of signals.

In that case we can solve for the optimal signal structure by solving the modified problem.

Optimal Information Structure

Sub-optimality of Full Correlation:

Proposition 7: *Given state-independent preferences, and $m = 2$, if $\tilde{v}(\mu_1, \mu_2) = v(a_{\mu^{CN}(\mu_1,\mu_2)})$ is sub-modular in (μ_1, μ_2), then the optimal information structure consists of homogeneous signals with a joint cumulative distribution function which is the lower Fréchet bound, and hence negative correlation arises.*

This result follows from Muller and Stoyan (2002) and the result on homogeneity of signals.

(The lower Fréchet bound of a distribution Γ is given by $\max\{0, 2\Gamma - 1\}$)

Conclusion

We consider the general problem of strategic information design when the receiver has correlation neglect.

General Manipulation:

- We show that the sender can manipulate the receivers expected posterior in any direction.
- Given a fixed number of signals manipulation is bounded.
- But as the number of signals increases the sender approaches her (state-dependent) first best.

Conclusion

Optimal Information Structure:

For state-independent utilities:

- Concavification provides an upper bound for the expected utility of the sender.
- Full correlation is optimal in the limit.
- If $\tilde{v}(\boldsymbol{\mu})$ is supermodular in $\boldsymbol{\mu}$ then full correlation is optimal.
- Full correlation is not always optimal. (submodularity, example 2)