The Status Quo and Belief Polarization of Inattentive Agents: Theory and Experiment

Vladimír Novák
(National Bank of Slovakia)

joint work with

Andrei Matveenko
(University of Mannheim)

Silvio Ravaioli
(Cornerstone Research)

EEA-ESEM Congress 2023

August 31, 2023

Societies polarized

- in their beliefs about future policies
- significant disagreement in their evaluations of the implemented status quo policies
  (e.g., Alesina, Miano, Stantcheva 2020)

Heterogeneity in the evaluation of the status quo

- leads to differences in perceived gains and losses associated with the adoption of a new policy
- significantly influences the demand for information and, consequently, essential economic decisions
- example: policies that aim to achieve climate neutrality (e.g., carbon tax) and many other applications

This Paper

How do valuations of the status quo influence belief polarization, and what important environmental factors determine the demand for information?

Model

- Rationally inattentive decision maker
- Mechanism (**state pooling**) by which endogenous information leads to **polarization ex-ante** conditional on a state

Lab Experiment

- **Generates polarization ex-ante** through state pooling, the magnitude is mitigated
- Demand for **simple signal structures** (fewer possible outcomes) and **preference for certainty** (degenerate posteriors)

Illustrative example

Setting

- State of the world $v \sim U[0, 1]$
- Two risk-neutral agents A and B facing a binary action $a \in \{0, 1\}$ representing preservation of the status quo and adoption of a new policy
- Agent A prefers $a = 1$ if $v \geq R_A$ and agent B prefers $a = 1$ if $v \geq R_B$, where $R_i \in (0, 1) \forall i$
- For simplicity assume $R_B < R_A$ and that both have the same uninformative prior

Illustrative example

Information acquisition

If information acquisition is costly, agents will demand the most instrumental signal structure

- agent A will ask whether $v \geq R_A$
- agent B will ask whether $v \geq R_B$
- none of them cares about the exact value of v
- **state pooling** - agents do not distinguish some states of the world, and pool states associated with the same action together

Illustrative example

Polarization

When the true state of the world $v \in (R_A, R_B)$

- the agents receive **opposite signals** whether they should adopt a new policy (given the assumption that signal is noiseless and truthful)
- agents' posterior expected values from the new policy would get polarized
  - move in the opposite direction
  - further apart as they were

The full-fledged model in the paper is much more general and shows **polarization ex-ante** - polarization of expected posterior conditional on a true state over all possible signal realizations from the selected information structure.

Literature: Suen (2004); Nimark and Sundaresan (2019); Bloedel and Seagal (2021); Hu, Li and Segal (2022), ...

Illustrative example

Polarization

When the true state of the world $v \in (R_A, R_B)$

- the agents receive **opposite signals** whether they should adopt a new policy (given the assumption that signal is noiseless and truthful)
- agents' posterior expected values from the new policy would get polarized
  - move in the opposite direction
  - further apart as they were

The full-fledged model in the paper is much more general and shows **polarization ex-ante** - polarization of expected posterior conditional on a true state over all possible signal realizations from the selected information structure.

Literature: Suen (2004); Nimark and Sundaresan (2019); Bloedel and Seagal (2021); Hu, Li and Segal (2022), ...

Illustrative example

Polarization

When the true state of the world $v \in (R_A, R_B)$

- the agents receive **opposite signals** whether they should adopt a new policy (given the assumption that signal is noiseless and truthful)
- agents' posterior expected values from the new policy would get polarized
  - move in the opposite direction
  - further apart as they were

The full-fledged model in the paper is much more general and shows **polarization ex-ante** - polarization of expected posterior conditional on a true state over all possible signal realizations from the selected information structure.

Literature: Suen (2004); Nimark and Sundaresan (2019); Bloedel and Seagal (2021); Hu, Li and Segal (2022), ...

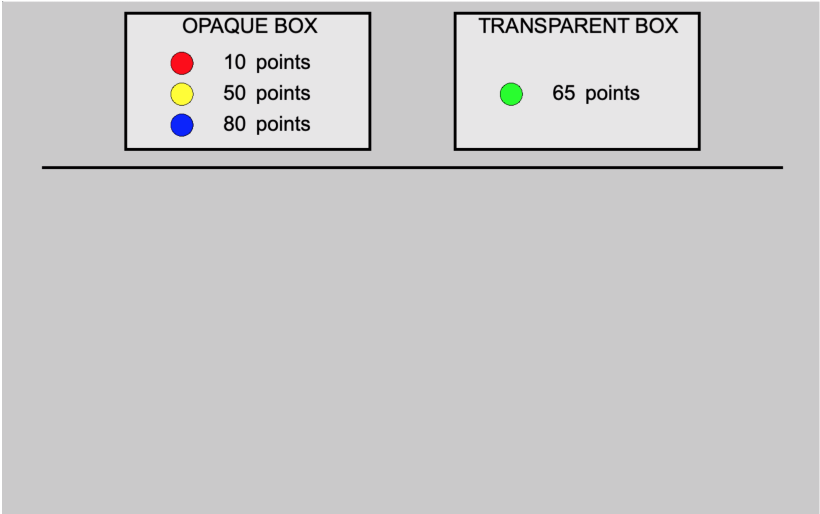
Laboratory Experiment - Procedure

- 85 participants (Columbia CELSS Lab)
- Avg time 80 minutes, Avg payment $\sim$\$25
- Payoffs expressed in probability points

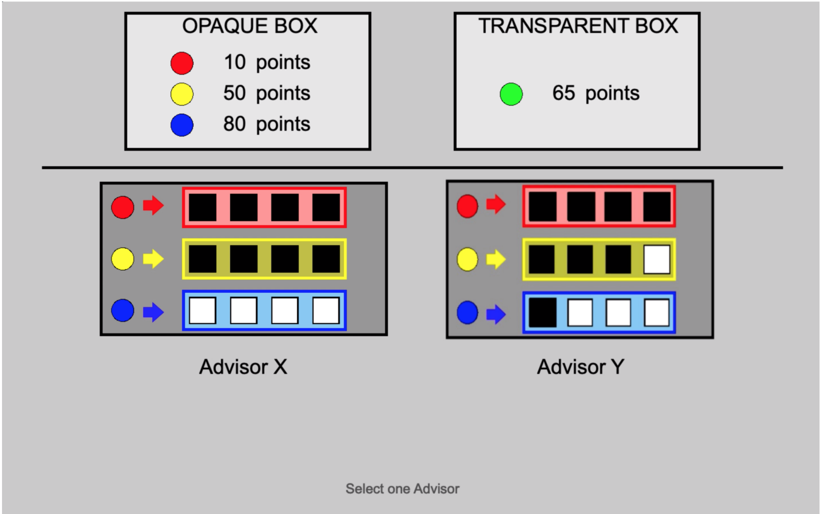
- **Main task**: choose advisor, make a choice (safe/risky)
  - Same pair of advisors, vary the status quo
  - Different pairs of advisors (value, complexity)

- **Extra tasks**: subjective beliefs elicitation
  - State probability (posterior)
  - Signal probability
- Additional data: Risk (Holt&Laury), Cognitive (Raven), Demographics, Questionnaire (optimism, superstition)

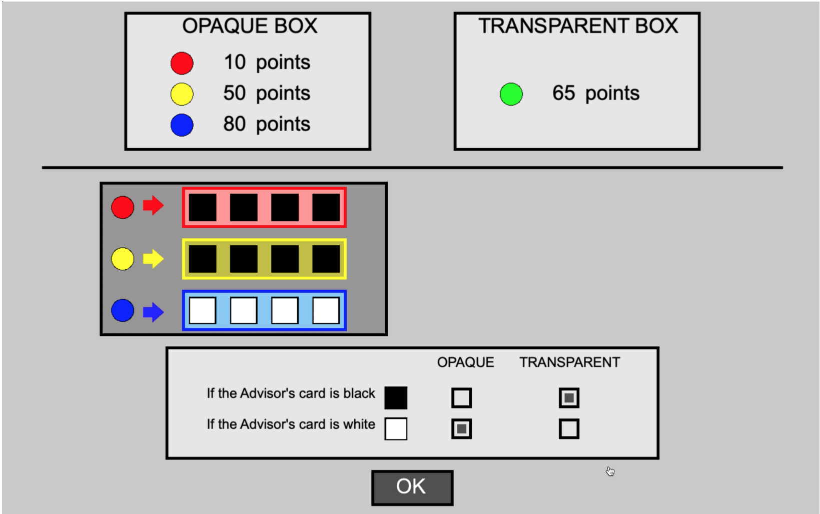
Main Task


OPAQUE BOX
10 points
50 points
80 points
TRANSPARENT BOX
65 points

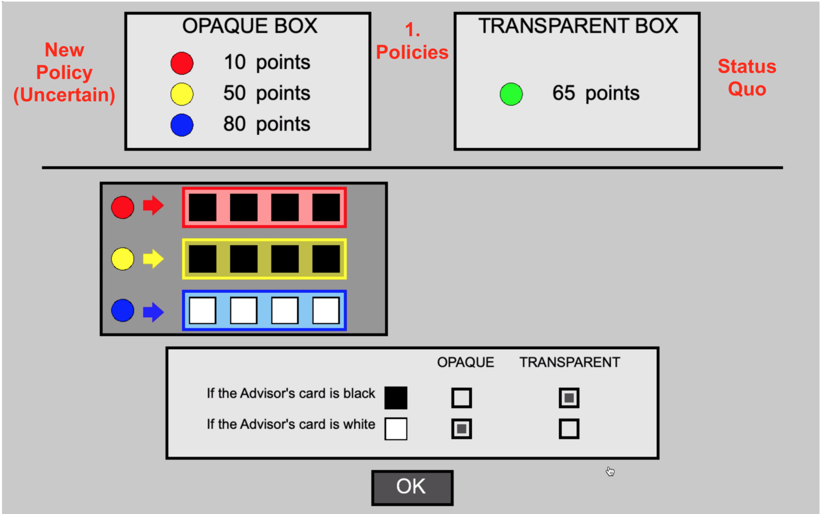
Main Task - Hiring screen



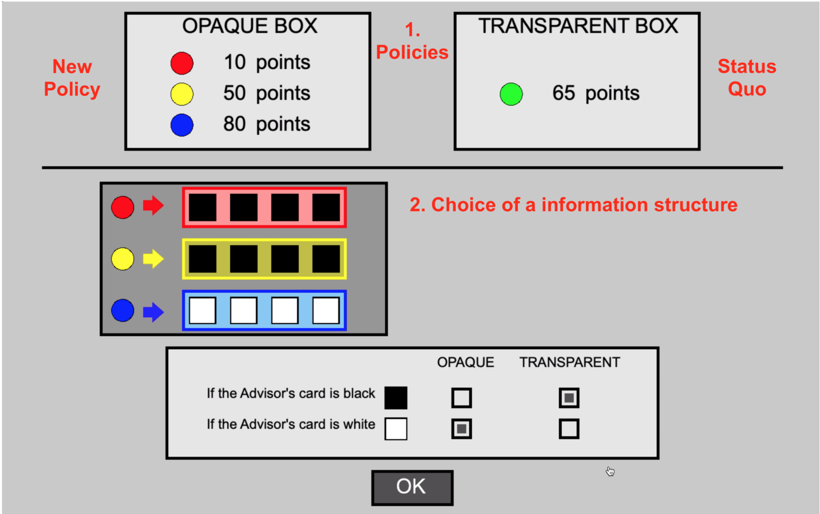
Main Task - Choice screen



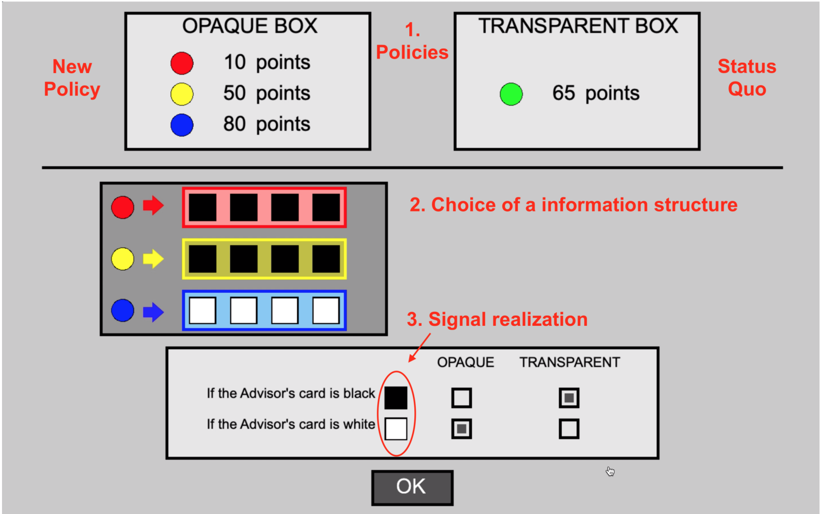
Main Task - Choice screen



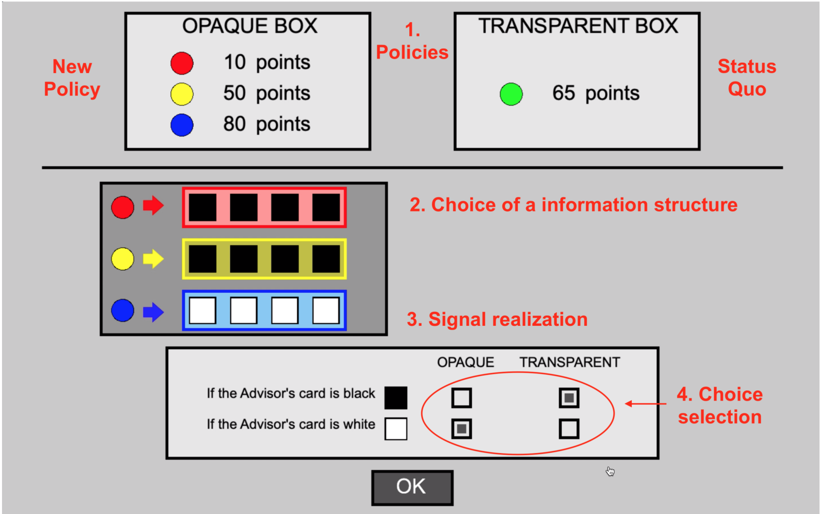
Main Task - Choice screen



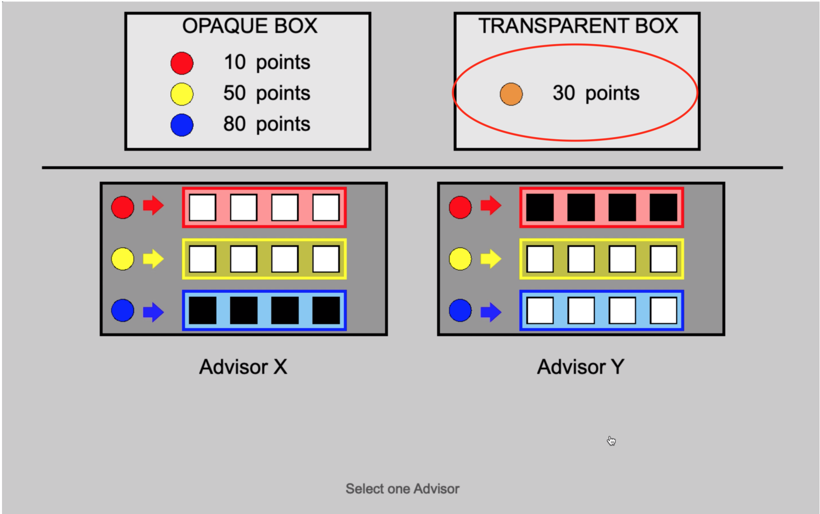
Main Task - Choice screen



Main Task - Choice screen

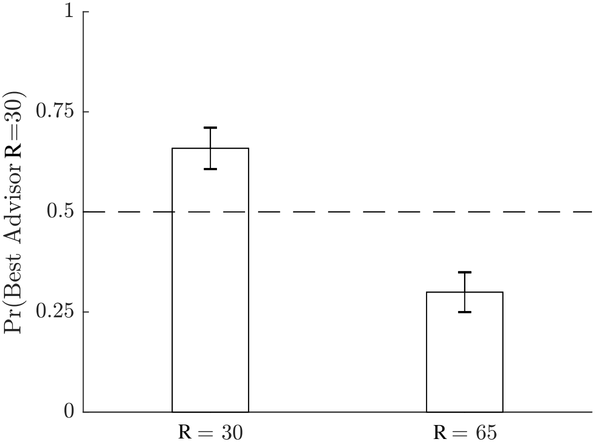


Main Task - Status Quo Manipulation



1. **Do Participants Switch Advisor?**
2. Do Participants have Unbiased Beliefs?
3. Do Participants get Polarized?
4. What is Mitigating Polarization?

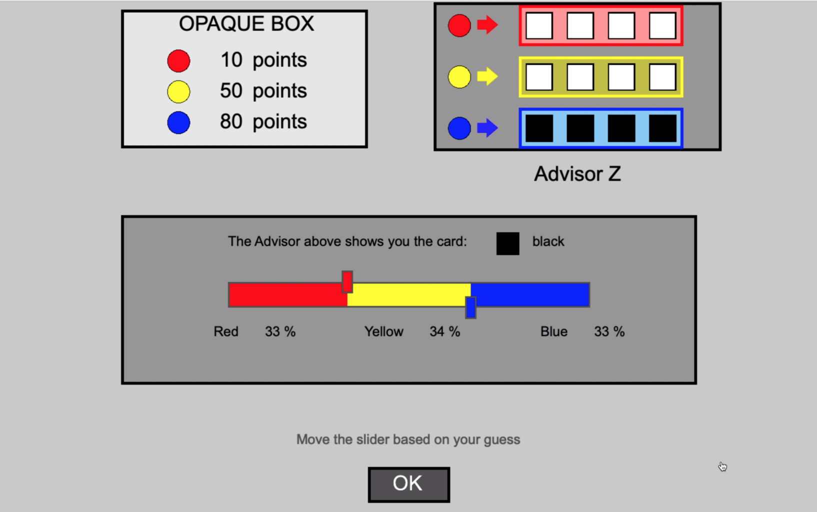
Most participants "switch" advisor



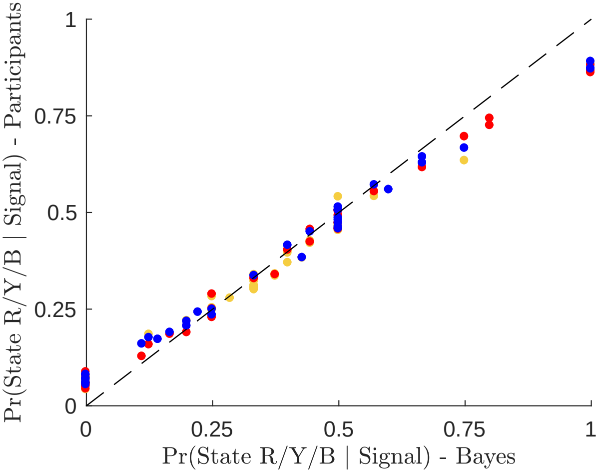
Advisor selection probability, all participants (n=85).
Each bar: 11/40 trials (935 observations).

1. Do Participants Switch Advisor?
2. **Do Participants have Unbiased Beliefs?**
3. Do Participants get Polarized?
4. What is Mitigating Polarization?

Beliefs elicitation - Posterior beliefs



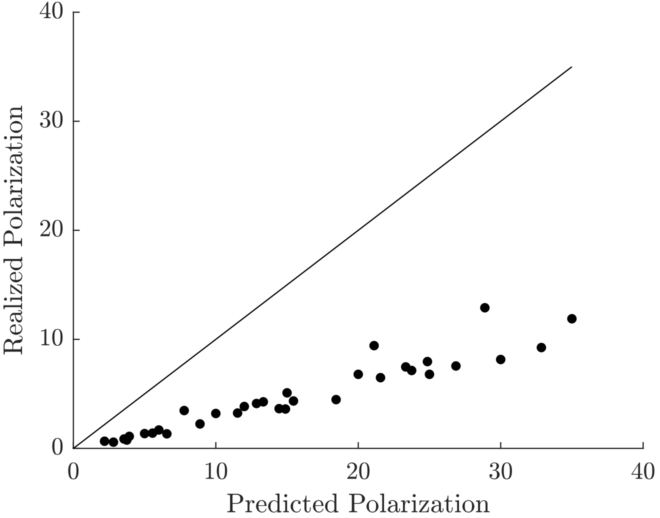
Beliefs elicitation - Predictions vs Behavior



Estimated probability: optimal and average subjective estimates. Posterior beliefs, colors indicate the state (40 trials, 85 observations per trial).

1. Do Participants Switch Advisor?
2. Do Participants have Unbiased Beliefs?
3. **Do Participants get Polarized?**
4. What is Mitigating Polarization?

Average Polarization



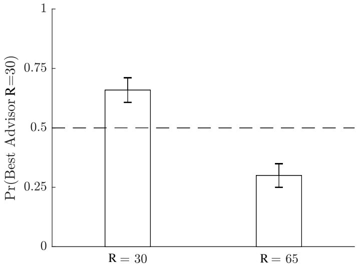
Predicted and realized within-subject polarization (n=85),
in the 11 pairs of trials with predicted advisor switch (3 states per trial).

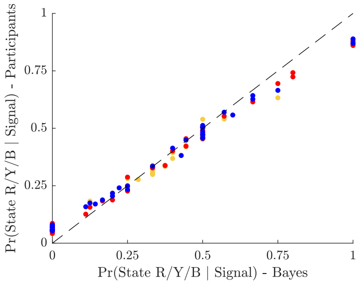
1. Do Participants Switch Advisor?
2. Do Participants have Unbiased Beliefs?
3. Do Participants get Polarized?
4. **What is Mitigating Polarization?**

What is Mitigating Polarization?

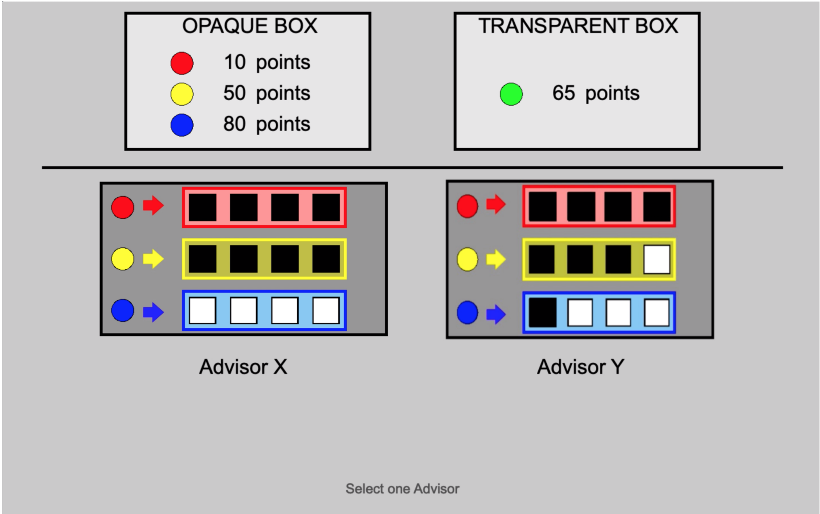
The usual suspects...

- **Advisor choice** Not respond to the manipulation
- **Beliefs** Not update enough after signal

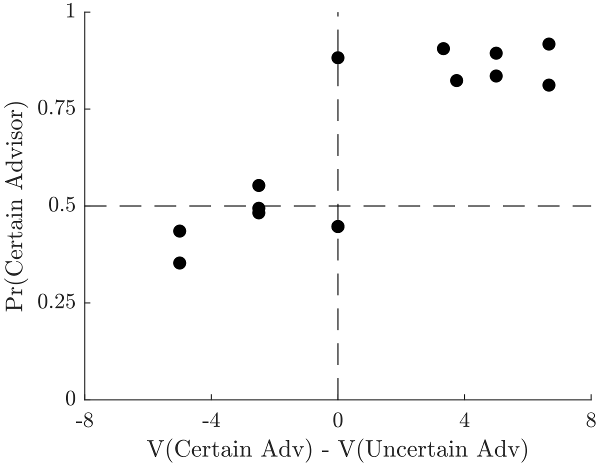




Difference between Advisors' complexity

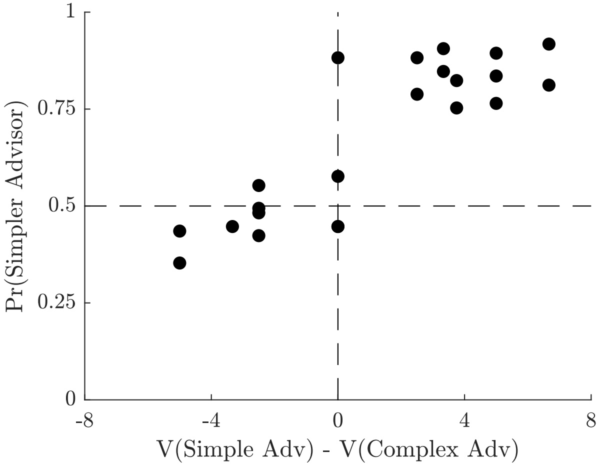


Certain vs Uncertain Advisor



Probability of choosing the certain advisor, in the trials that have a certain advisor
and an uncertain advisor (14/40 trials). 85 observations per trial.

Simple vs Complex Advisor



Probability of choosing the simplest advisor, in the trials that have different complexity scores (21/40 trials). 85 observations per trial.

Complexity score $c_I = \Sigma_\sigma(\Sigma_s \mathbb{1}(p(s|\sigma) > 0) - 1)$.

Advisor Choice - Logit Regressions

	(1)	(2)	(3)	(4)
Value w_I^{Bayes}	**0.246*****	**0.217*****	**0.235*****	**0.232*****
	(0.018)	(0.011)	(0.011)	(0.019)
Best Advisor	**-0.084**			**-0.007**
	(0.096)			(0.102)
Complexity c_I		**-0.359*****		**-0.074*****
		(0.037)		(0.076)
Certainty			**0.511*****	**0.428*****
			(0.069)	(0.110)
State Pooling			**0.404*****	**0.330****
			(0.069)	(0.102)
Trials	All	All	All	All
Observations	3,400	3,400	3,400	3,400

Advisor choice. Notation: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ (H0: $\beta = 0$)

State Pooling advisors under status quo value R can provide a signal σ that generate posterior beliefs either $Pr(\pi_s > R|\sigma) = 0$ or $Pr(\pi_s > R|\sigma) = 1$.

Conclusions

Model:

- **Rational and endogenous** belief polarization
- Role of the status quo for information acquisition
- Key mechanism: state pooling

Lab experiment:

- A change in the safe option generates **“advisor switches”**
- and creates (mitigated) **belief polarization**
- Causes of mitigation: instrumental + non-inst. features

Implications:

- Interventions to reduce polarization
- Infer the agent’s type (status quo) based on action and info

Vladimír Novák
(National Bank of Slovakia)

joint work with

Andrei Matveenko (University of Mannheim)
Silvio Ravaioli (Cornerstone Research)

EEA-ESEM Congress 2023

August 31, 2023

Previous literature

Rational Inattention

- Discrete choice - Sims (1998, 2003), Matějka & McKay (2015), Steiner et al. (2017)
- Posterior based approach - Caplin & Dean (2015)

Polarization - persistent/exogenous biases

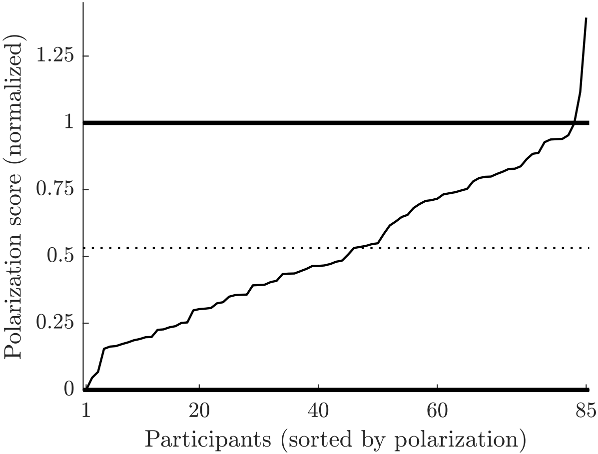
- De-polarization - Savage (1954), Blackwell & Dubins (1962)
- Exogenous bias - Rabin & Schrag (1999), Dixit & Weibull (2007), Ortoleva & Snowberg (2015)
- Biased search or signal interpretation - Rabin & Schrag (1999), Klayman & Ha (1987), Ortoleva & Snowberg (2015)
- Inattentiveness - Nimark & Sundaresan (2019)

Laboratory experiments

- Preference for skewed info - Masatlioglu et al. (2017)
- Demand for information - Ambuehl & Li (2018)
- Choice over biased info - Charness, Oprea & Yuksel (2020)

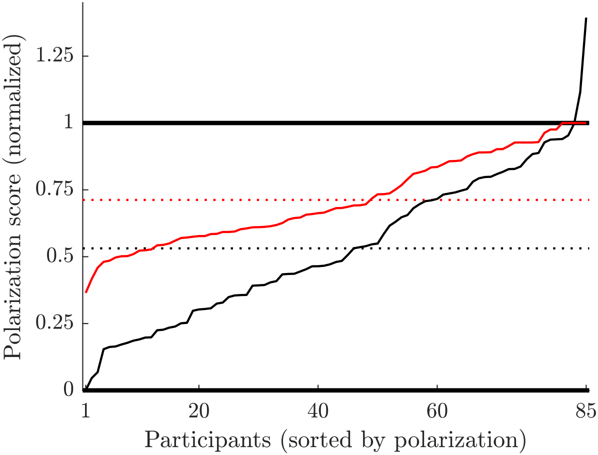
Extra - Experiment

Polarization - Subject Level Analysis



Estimated polarization coefficient $\hat{p}_i$ by subject. Distribution of coefficients, subjects ordered by $\hat{p}_i$.

Polarization - Subject Level Analysis



Estimated polarization coefficient $\hat{p}_i$ by subject (black) and by controlling for beliefs (red).

Polarization - Subject Level Analysis

	Baseline (1)	Full (2)	Baseline (3)	Full (4)
Risk attitude (Holt and Laury)	-0.52^{***}	-0.50^{***}	-0.27	-0.26
	(0.16)	(0.16)	(0.24)	(0.25)
Fluid intelligence (Raven test)	0.13	0.10	0.20	0.07
	(0.11)	(0.14)	(0.12)	(0.15)
Familiar with Bayes rule	0.03	0.02	0.10	0.12
	(0.10)	(0.10)	(0.11)	(0.09)
Analytical studies	0.09	0.10	0.06	0.07
	(0.09)	(0.10)	(0.10)	(0.11)
LOT-R scale		-0.03		-0.06
		(0.04)		(0.05)
SUPERSTITION scale		-0.03		-0.01
		(0.04)		(0.05)
RISK scale		-0.02		-0.07^{*}
		(0.04)		(0.04)
Observations	63	63	63	63
Demographic Controls			✓	✓

Polarization score. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Predict Type from Observables

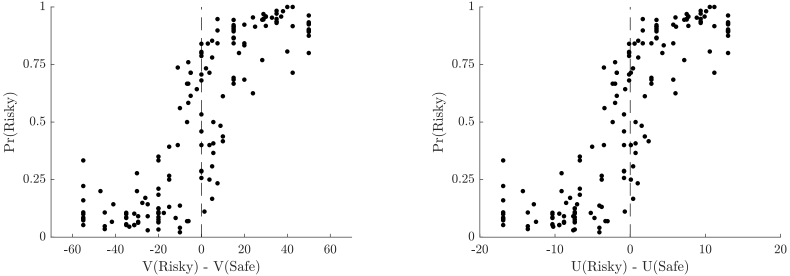
How accurately can we predict the type (status quo value) from observable behavior?

	Prediction	Data
No information	50.0%	50.0%
Choice only	69.7%	62.6%
Search only	100.0%	68.0%
Search+Choice	100.0%	68.4%
Search+Signal+Choice	100.0%	72.9%

Inference of the agent's status quo: predicted and realized accuracy.

Imagine a social media platform like Facebook has access to a dataset of actions performed by an user: publicly observable ones (likes, list of friends) and search actions (clicks, searches).

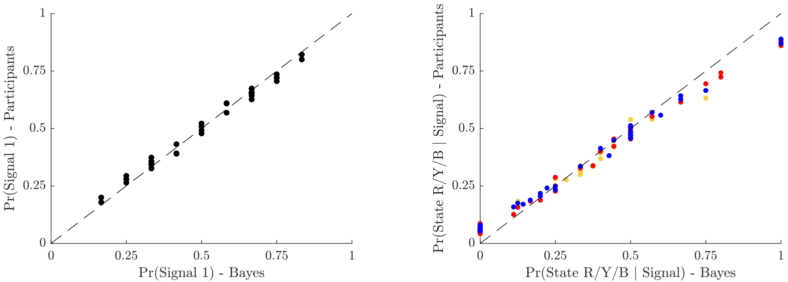
Risk Attitude



Action selection probability: EV (L) and calibrated CRRA EU (R).

MLE for risk aversion coefficient (CRRA): $\hat{\alpha} = 0.34$.
Reject the null hypothesis $\alpha = 0$ (p<0.001).
Pseudo-R^2: from $R^2_{risk.neutral} = 0.382$ to $R^2_{risk.averse} = 0.422$.

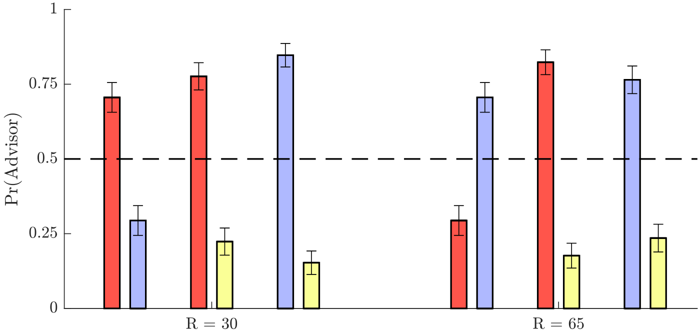
Subjective Beliefs



Average subjective beliefs: Task 3 (L) and 4 (R), 85 observations per point.

L: Signal probability $\hat{p} = 0.041 + 0.918 \cdot p$ 7.9cm with $R^2 = 0.991$
R: Posterior probability $\hat{p} = 0.058 + 0.825 \cdot p$ 7.9cm with $R^2 = 0.993$

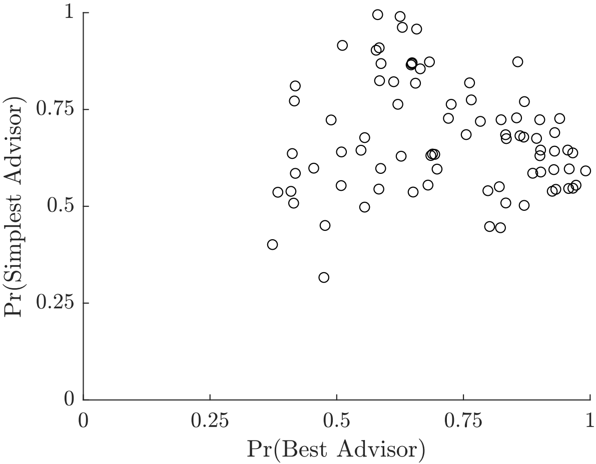
Certainty Advisors



Advisor choice under yes/no questions (main task).

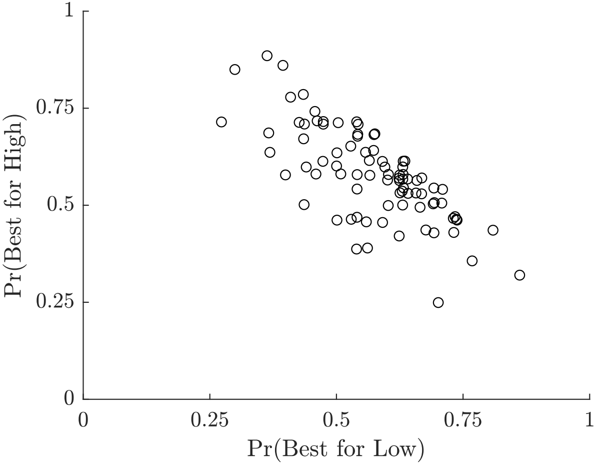
When subjects face a choice between certainty state pooler and certainty advisors, they select on average the certainty state pooler in 74% of the trials.

Subject Heterogeneity /1



Probability of choosing the best advisor (based on instrumental value) and simplest advisor (based on the complexity score).

Subject Heterogeneity /2



Probability of choosing the advisor that provides more information about the low or the high state in different types of trials.

Extra - Model

Agent's problem

$$\max_{\text{Information strategy}} \{\mathbb{E}(U) - \text{cost of information}\}$$

$$\max_{\{\mathcal{P}(i|s)|i=1,2;\ s\in S\}} \left\{ \sum_{s=1}^{n} \left(v_s \cdot \mathcal{P}\left(i=1|s\right) + R \cdot \mathcal{P}(i=2|s)\right) g_s - \lambda\kappa \right\},$$

subject to

$$\forall i: \ \mathcal{P}(i|s) \geq 0 \qquad \forall s \in S\,,$$

$$\sum_{i=1}^{2} \mathcal{P}(i|s) = 1 \qquad \forall s \in S\,,$$

$$\kappa = \underbrace{-\sum_{i=1}^{2} \mathcal{P}(i) \log \mathcal{P}(i)}_{\text{prior uncertainty}} - \sum_{s=1}^{n} \left(\underbrace{-\left(\sum_{i=1}^{2} \mathcal{P}(i|s) \log \mathcal{P}(i|s)\right)}_{\text{posterior uncertainty in state } s} g_s \right).$$

Lemma 1: Solution

Conditional on the realized state of the world s^*

$$\mathcal{P}(\text{new policy } | s^*) = \mathcal{P}(i = 1 | s^*) = \frac{\mathcal{P}(i = 1) e^{\frac{v_{s^*}}{\lambda}}}{\mathcal{P}(i = 1) e^{\frac{v_{s^*}}{\lambda}} + (1 - \mathcal{P}(i = 1)) e^{\frac{R}{\lambda}}}$$

$$\mathcal{P}(\text{status quo } | s^*) = \mathcal{P}(i = 2 | s^*) = \frac{(1 - \mathcal{P}(i = 1)) e^{\frac{R}{\lambda}}}{\mathcal{P}(i = 1) e^{\frac{v_{s^*}}{\lambda}} + (1 - \mathcal{P}(i = 1)) e^{\frac{R}{\lambda}}}$$

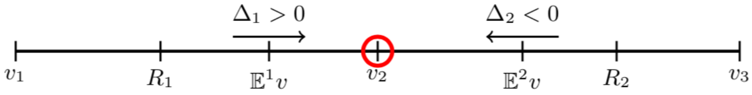
$\mathcal{P}(i = 1)$ - unconditional probability of choosing a new unknown policy

$\lambda = 0$ chooses the option with the highest value with probability one

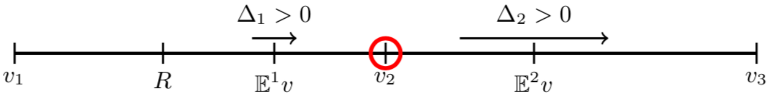
Convergence

Theorem

Let us assume that there are two agents $j = 1, 2$ that are characterized by the pair $(R^j, \mathbb{E}^j v)$.
If in state of the world $s^* \in S$ the conditions $(\mathbb{E}^1 v - \mathbb{E}^2 v)(v_{s^*} - R^1) < 0$ and $(\mathbb{E}^1 v - \mathbb{E}^2 v)(v_{s^*} - R^2) > 0$ hold, then the two agents converge in their beliefs in this state of the world.



Divergence updating in the same direction



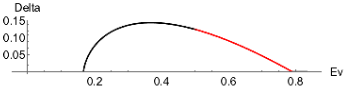


Figure 1: $\Delta(s^* = 2)$ as a function of $\mathbb{E}v$ for R_1 and λ_2. The red area depicts the region of wrong updating.

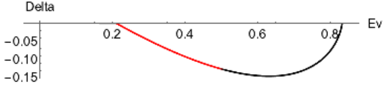


Figure 2: $\Delta(s^* = 2)$ as a function of $\mathbb{E}v$ for R_2 and λ_2. The red area depicts the region of wrong updating.

Comparative statics

Cheaper information ($\lambda_2 < \lambda_1$) might lead to higher polarization

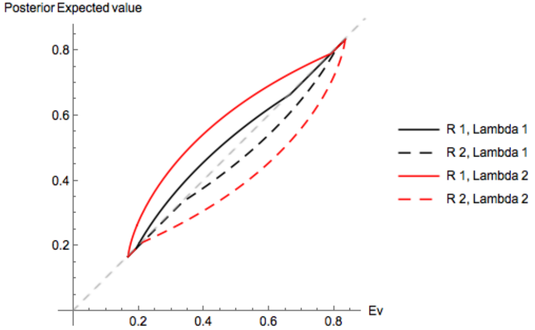


Figure 3: $\mathbb{E}_i[\mathbb{E}(v|i)|s^*]$ as a function of $\mathbb{E}v$ for different levels of R and λ. The solid lines are the case with R_1 and dashed with R_2. Black corresponds to cases with λ_1 and red is used for λ_2.