Online Appendix for "The Status Quo and Belief Polarization of Inattentive Agents: Theory and Experiment"

Vladimír Novák, Andrei Matveenko, and Silvio Ravaioli*

January, 2024

*Novák: Corresponding author. Department of Economic and Monetary Analysis, National Bank of Slovakia, and the Faculty of Business Administration, Prague University of Economics and Business (e-mail: vladimir.novak@nbs.sk); Matveenko: Department of Economics, University of Mannheim (e-mail: matveenko@uni-mannheim.de); Ravaioli: Cornerstone Research (e-mail: silvio.ravaioli@cornerstone.com).
We thank two anonymous referees and the Editor (Leeat Yariv) for useful comments and suggestions. Many thanks to Filip Matějka for invaluable advice and crucial help in setting up the framework employed in this paper, and to Mark Dean for continuous support on this project. We are grateful for comments from Sandro Ambuehl, Niels Boissonnet, Andrew Caplin, Alessandra Casella, Yeon-Koo Che, Navin Kartik, Alexander Nesterov, Pietro Ortoleva, Jacopo Perego, Debraj Ray, Avner Shaked, Jakub Steiner, Michael Woodford, Jan Zápal, and numerous seminar participants for helpful comments and suggestions. This project has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme (grant agreements No. 678081 and No. 740369) and we acknowledge the financial support of grants from the Columbia Experimental Laboratory for Social Sciences. Funding by the German Research Foundation (DFG) through CRC TR 224 (project B03) is gratefully acknowledged. Vladimír Novák was supported by the Charles University in Prague (GA UK projects No. 197216 and No. 616219) and by the H2020-MSCA-RISE project GEMCLIME-2020 GA. no 681228. The authors affirm no financial interest in the outcomes of the experiments detailed herein. All data were collected with the approval of the Columbia University Institutional Review Board (protocol AAAS5801). The views expressed in this paper are those of the authors and should not be attributed to the National Bank of Slovakia or Cornerstone Research.

B Appendix: Belief divergence without polarization

In this appendix, we demonstrate a possibility of belief divergence when these beliefs are updated in the same direction, which, in our opinion, is an interesting feature of our model. Specifically, we explore the influence of the prior beliefs on the magnitude of the change in the mean of beliefs. To study this question, we take advantage of an example with three states and two actions. This problem is a simple benchmark, and its solution exhibits the basic features of solutions to the problems with n states and two actions. The solution we analyze in this section is symbolic. Let us start with the definition of divergence of beliefs updated in the same direction.

Definition 5. We say that two agents $j \in \{A, B\}$, who are characterized by the pair $(R^j, \mathbf{g}^j)$ and are choosing between actions $i = \{1, 2\}$, *diverge in their belief updated in the same direction* when in the state $s^* \in S$ the following two conditions are satisfied

1. $\left|m^A(s^*) - m^B(s^*)\right| > \left|\mu^A - \mu^B\right|$.

2. $\Delta_A(s^*) \cdot \Delta_B(s^*) > 0$.

The parameter values that we use in this appendix are as follows: $v_1 = 0$, $v_2 = 1/2$, $v_3 = 1$, $g_1 = g \in (0, 2/3)$, $g_2 = 1/3$, $g_3 = 2/3 - g$, $R = 3/8$, $\lambda = 1/8$. Note that keeping the prior probability of state 2, g_2, fixed, we can vary the prior probability of state 1, g only between $(0, 2/3)$. Also, $\mathbb{E}v$ can vary only from 1/6 to 5/6. To solve the problem (1)-(4) it is necessary to find the unconditional probabilities $\mathcal{P}(i = 1)$ and $\mathcal{P}(i = 0)$, which we then use for finding the conditional probabilities.

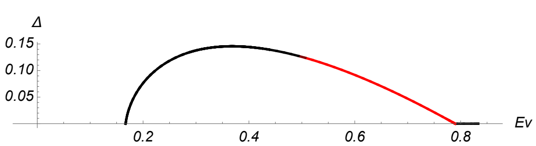


Figure 12: $\Delta(s^* = 2)$ as a function of $\mathbb{E}v$ for $R_1 = 3/8$ and $\lambda = 1/8$. The red area depicts the region of updating in the opposite direction from the realized value.

Figure 12 provides an example in which two agents diverge in their beliefs while they update in the same direction. Since two agents differ only in their prior expectations about the new policy, it is sufficient to look at how a single agent's change in the mean of beliefs $\Delta(s^* = 2)$ depends on $\mathbb{E}v$. We are interested in finding two prior expected beliefs for which there is a divergence of posterior beliefs. To do so, we need to find two points such that Δ for the left point is lower than Δ for the right point. In our example, the red part of the plot is a decreasing function. This means that, in our example, two agents updating in the same direction with the same valuation of the status quo might diverge in their opinions only when they are updating towards the realized value. However, in the black part of the plot, it is easy to find two points at which the agents diverge in their opinions.

C Appendix: Experimental Design and Procedure

C.A Experimental Interface and Payment

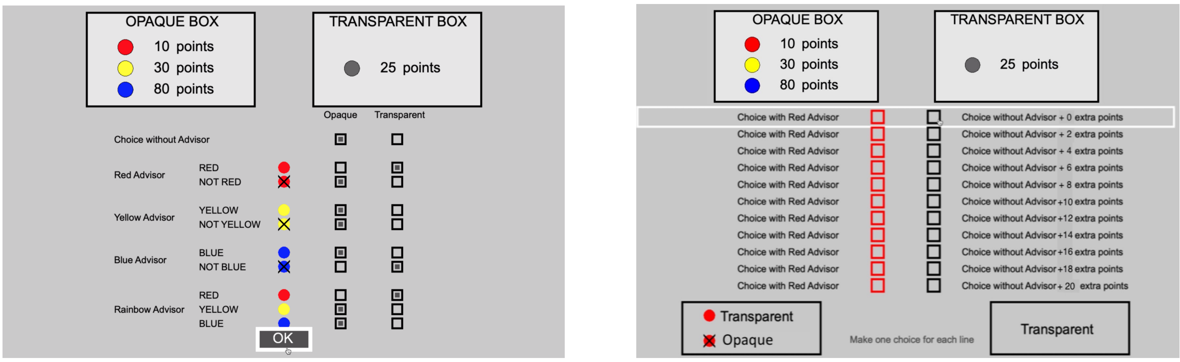
Task 1 - Colorblind advisor game

If one round from this part is selected for the bonus payment, a subject receives the $15 bonus with the percentage probability equal to the number of points that she collected in that round. Since each line counts as a separate decision, one of which might be randomly drawn for payment, truthful revelation is strictly optimal. We constrain subjects to have at most one switching point for every advisor.

Task 2 - Imprecise advisor game If one round from this part is selected for the bonus payment, subjects receive the $15 bonus with the percentage probability equal to the number of points that she collected in that round.

Task 3 - Card color prediction game If one round from this part is selected for the bonus payment, the computer randomly determines the state and realized signal, and subjects receive the $15 bonus with the percentage probability determined by the quadratic loss scoring rule.

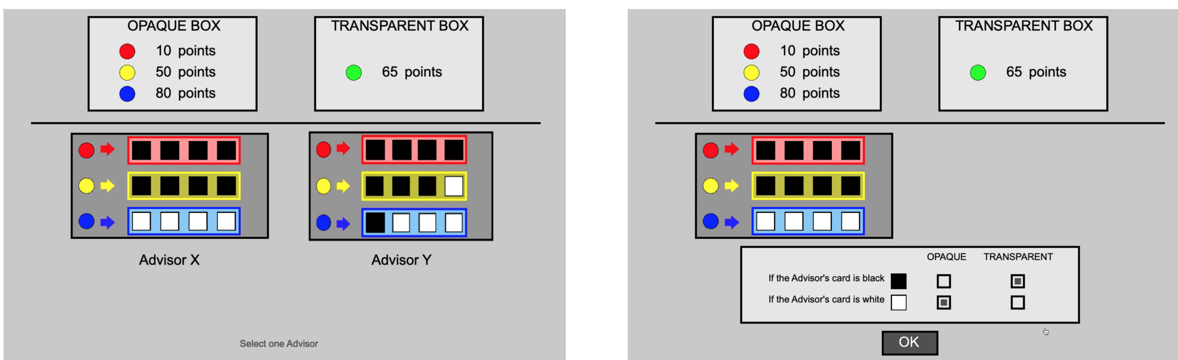
Task 4 - Ball color prediction game If one round from this part is selected for the bonus payment, the computer randomly determines the state and realized signal, and subjects receive the $15 bonus with the percentage probability determined by the quadratic loss scoring rule.



Task 1, Screen 1: Action choice

Task 2, Screen 2: WTA for each advisor

Figure 13: Task 1: Colorblind advisor game. Left: Subjects choose an action (box) contingent on the advisor and signal received. The possible values of each action are indicated on the top of the screen. Each state (ball color) is equally likely to occur. Right: Subjects indicate for each advisor the willingness to accept renunciation of its signal in a series of binary choices (BDM method). At most, one switch is allowed. Action choices selected in the previous stage are reported on the bottom of the screen.



Task 1, Screen 1: Advisor choice

Task 2, Screen 2: Action choice

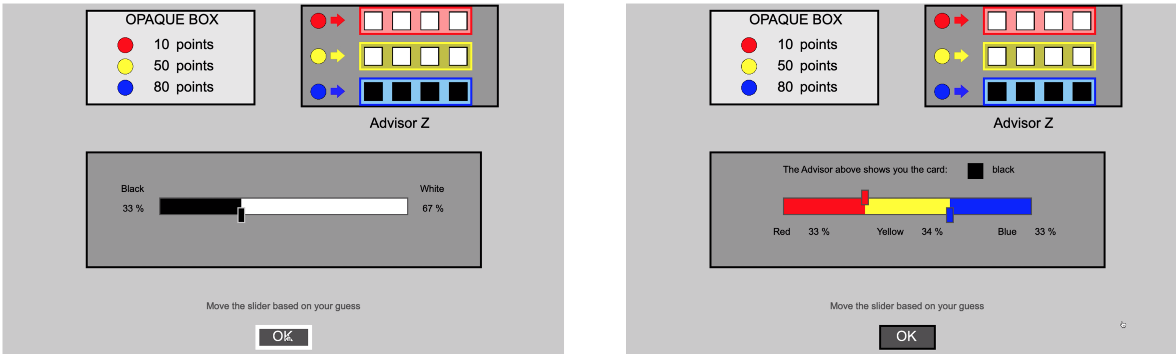
Figure 14: Task 2: Imprecise advisor game. Left: Subjects choose one signal structure (advisor) between the two options available. Each advisor is a triplet of state-contingent signal probabilities. Right: Subjects indicate the signal-contingent action for each signal (strategy method).

C.B Randomization

In all the tasks we randomize the order of the trials. For task 2 only, the first 3 trials are randomly drawn from the subset of trials where both the advisors provide certainty (in order to facilitate the transition from task 1 to task 2).

In task 1 we randomize the order of the four hiring screens within each trial.

In task 2 we randomize the positions of the two advisors on the screen.



Task 3: Beliefs over signal likelihood

Task 4: Beliefs over state likelihood

Figure 15: Left: Task 3 (Card color prediction game). Subjects indicate the likelihood of observing each signal (card color) for the given advisor. Right: Task 4 (Ball color prediction game). Subjects indicate the likelihood of each state (ball color) given an advisor and signal. In both tasks subjects move the slider(s) and receive a number of probability points according to the quadratic loss scoring rule described in the instructions.

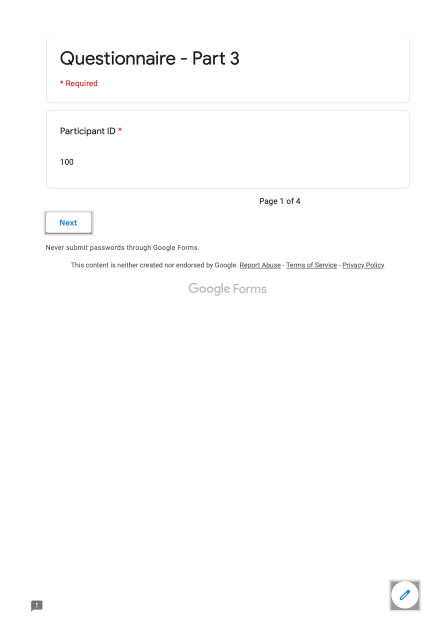
In tasks 2, 3, and 4 we randomize the advisors' card colors (black and white). This means that the signal-contingent choice in the second part of the round requires the subjects to analyze every advisor separately, since the colors do not convey any intrinsic message, and this procedure reduces the concern regarding inertia in the evaluation of the advisor and in actions.

C.C Subject understanding

Instructions were provided on both the computer screen, as slides that can be browsed by each subject at the desired pace, and as a paper printout. The two versions of the instructions contained the same information verbatim. Before proceeding with every section of the experiment, subjects were required to correctly answer all the multiple-choice questions of the comprehension test to check understanding of the instructions. The number of questions ranged from two to four for every section, and subjects received a one-minute timeout before having a new attempt. Subjects were initially informed about the payment structure, the no-deception policy of the laboratory, and that choices in one section of the experiment did not affect any other section, or the questionnaire. A small number of subjects were recruited for each laboratory session (6 on average) in order to facilitate clarification of questions during the experiment.

D Appendix: Questionnaire

At the end of the four tasks, we have an additional section with a Holt and Laury test of risk aversion (Part 1), Raven matrices test of fluid intelligence (Part 2, five matrices of different difficulty), and a series of questions (Part 3), that we show here as they were presented to subjects.



Questionnaire - Part 3

* Required

Participant ID *

100

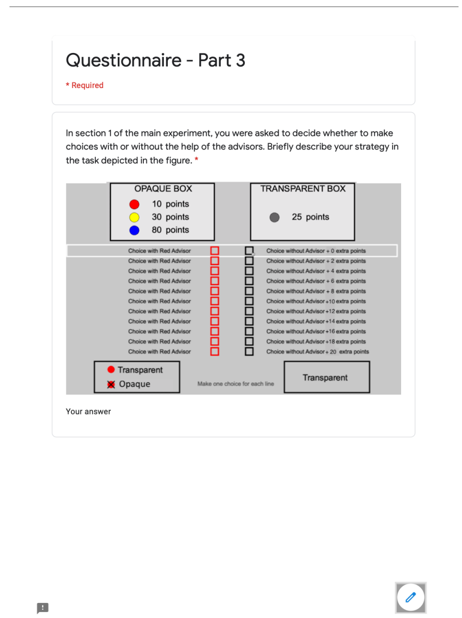
Page 1 of 4

Next

Never submit passwords through Google Forms.

This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Privacy Policy

Google Forms

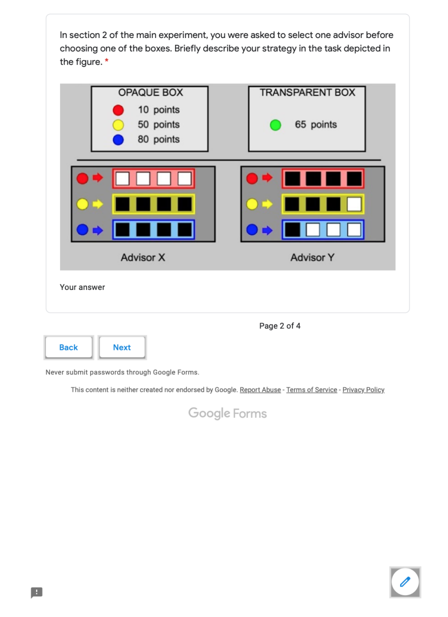


Questionnaire - Part 3

* Required

In section 1 of the main experiment, you were asked to decide whether to make choices with or without the help of the advisors. Briefly describe your strategy in the task depicted in the figure. *

Your answer



In section 2 of the main experiment, you were asked to select one advisor before choosing one of the boxes. Briefly describe your strategy in the task depicted in the figure. *

Your answer

Page 2 of 4

Back

Next

Never submit passwords through Google Forms.

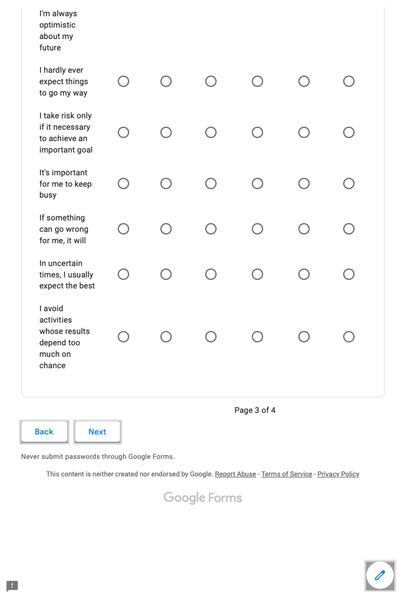
This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Privacy Policy

Google Forms

How much do you agree with the following statements? *

Try not to let your response to one statement influence your responses to other statements. There are no "correct" or "incorrect" answers. Answer according to your own feelings, rather than how you think "most people" would answer.

	I agree a lot	I agree a little	I neither agree nor disagree	I disagree a little	I disagree a lot	Prefer not to say
I enjoy my friends a lot	○	○	○	○	○	○
I own an object that I feel brings me good luck during exams	○	○	○	○	○	○
I often take risk just for fun	○	○	○	○	○	○
I rarely count on good things happening to me	○	○	○	○	○	○
I am uncomfortable if I am assigned seat number 13 in a theatre	○	○	○	○	○	○
It's easy for me to relax	○	○	○	○	○	○
I refrain from taking risks when I feel I am having an unlucky day	○	○	○	○	○	○
I don't get upset too easily	○	○	○	○	○	○
Overall, I expect more good things to happen to me than bad	○	○	○	○	○	○
	○	○	○	○	○	○


I'm always optimistic about my future
I hardly ever expect things to go my way
I take risk only if it necessary to achieve an important goal
It's important for me to keep busy
If something can go wrong for me, it will
In uncertain times, I usually expect the best
I avoid activities whose results depend too much on chance

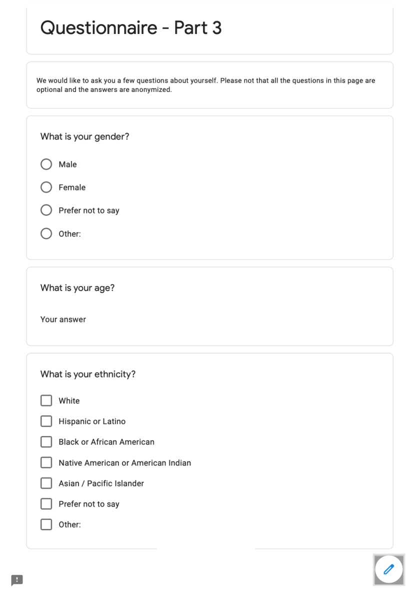
Page 3 of 4

Back Next

Never submit passwords through Google Forms.

This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Privacy Policy

Google Forms


Questionnaire - Part 3

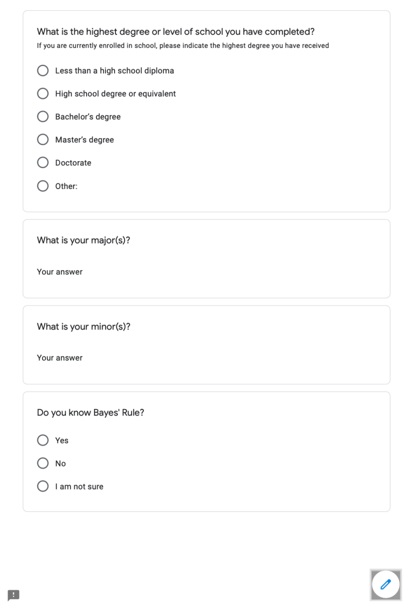
We would like to ask you a few questions about yourself. Please not that all the questions in this page are optional and the answers are anonymized.

What is your gender?
- Male
- Female
- Prefer not to say
- Other:

What is your age?

Your answer

What is your ethnicity?
- White
- Hispanic or Latino
- Black or African American
- Native American or American Indian
- Asian / Pacific Islander
- Prefer not to say
- Other:


What is the highest degree or level of school you have completed?
If you are currently enrolled in school, please indicate the highest degree you have received
- Less than a high school diploma
- High school degree or equivalent
- Bachelor's degree
- Master's degree
- Doctorate
- Other:

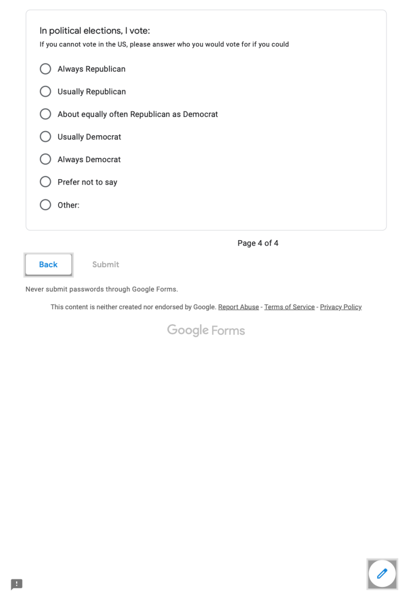
What is your major(s)?

Your answer

What is your minor(s)?

Your answer

Do you know Bayes' Rule?
- Yes
- No
- I am not sure


In political elections, I vote:
If you cannot vote in the US, please answer who you would vote for if you could
- Always Republican
- Usually Republican
- About equally often Republican as Democrat
- Usually Democrat
- Always Democrat
- Prefer not to say
- Other:

Page 4 of 4

Back Submit

Never submit passwords through Google Forms.

This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Privacy Policy

Google Forms

E Appendix: Advisor Pairs in the Main Task

In task 2, subjects play 40 rounds with different pairs of advisors and values for the ball in the transparent box. The ball in the transparent box can take two values: 30 points (low status quo) and 65 points (high status quo). The values for the balls in the opaque box are unchanged during the task (10, 50, and 80 points, with uniform probability of being drawn). The rounds are designed as a combination of 20 advisor pairs and two values for the ball in the transparent box.

The table 4 shows the pairs of advisors, here labeled X and Y. Advisor names, positions on the screen, and signal colors were randomized at the subject level. Each advisor is presented as a triplet of conditional signal probabilities, conditional on the realized state (value of the risky action). For each pair of advisors, the table indicates which of them has the highest instrumental value under each status quo value (low or high), with $\sim$ to denote ties.

	Advisor X			Advisor Y			Best Low R	Best High R
Pair	x_{bad}	x_{med}	x_{good}	y_{bad}	y_{med}	y_{good}	$(R = 30)$	$(R = 65)$
1	0	1	1	0	0	1	X	Y
2	0.25	1	1	0	0	1	X	Y
3	0	1	0.75	0	0	1	X	Y
4	0	0.75	1	0	0	1	X	Y
5	0	1	1	0	0	0.5	X	Y
6	0	1	1	0	0.25	0.75	X	Y
7	0.25	1	1	0	0	0.75	X	Y
8	0.5	1	1	0	0	0.75	X	Y
9	0.25	1	1	0	0.25	0.75	X	Y
10	0.25	0.75	1	0	0	0.75	X	Y
11	0.25	0.75	1	0	0.25	0.75	X	Y
12	0	0.5	1	0.25	0.5	0.75	X	X
13	0	1	1	0	1	0	X	$\sim$
14	0.25	0.75	1	0.75	0.25	1	X	$\sim$
15	0	1	1	0.5	1	1	X	$\sim$
16	0	0	1	0	1	0	$\sim$	X
17	0	0	0.5	0.25	0.5	0.75	$\sim$	X
18	0	0.5	0.5	0	1	0	$\sim$	$\sim$
19	0.5	0	0.5	0.5	0.5	1	$\sim$	$\sim$
20	0	0.25	0.5	0	0.5	0.75	$\sim$	$\sim$

Table 4: Pairs of advisors used in Task 2. Each pair contains two different advisors (X and Y), with the triplet of signal probabilities $p_i = Pr(\sigma = 1|s = i)$. The last two columns show the theoretical predictions for a Bayesian decision-maker. Each pair of advisors is presented with two different status quo values (low R and high R). For each of these values, we indicate which of the two advisors has the highest instrumental value, with $\sim$ in case of a tie.

The advisor pairs are selected in order to examine preference over sources of information and formulate predictions about the effect of the safe option on information collection and posterior beliefs. Based on the predicted behavior of a Bayesian agent, we can classify the pairs into the following groups:

- Pairs 1-11: pick different advisors by changing the safe option (strict preference);
- Pairs 12-16: pick different advisors by changing the safe option (weak preference);
- Pair 17: always pick advisor X regardless of the safe option (Blackwell ordered signals);
- Pairs 8-20: indifference between the advisors for both safe options.

The table 5 shows the pairs of advisors as in the table 4 extended for corresponding measures of simplicity. In particular, we include dummy for certainty, discrete complexity and the continuous measure of complexity. The discrete complexity measure is defined by equation 6 and the continues measure of complexity for state s and a signal σ is defined by following equation

$$c_C = \sum_s \sum_\sigma \sqrt{\mathcal{P}(s|\sigma_S)} \tag{9}$$

F Appendix: Further Analysis on Advisor Choice

Certain advisors provide an answer to a question of the kind: "Is the state red(/yellow/blue)?" and allow the subject to learn with certainty if a particular state is realized (with probaility one) or not realized (with probability zero).

Figure 18 shows advisor choice in the trials in which both advisors provide certainty. We display separately the trials with different status quo values. When the subjects have to choose between advisors that provide certainty and are also state poolers, that is, between an advisor providing information whether the state is blue and another advisor providing information whether the state is red (first couple of bars for $R = 30$ and $R = 65$), they significantly select the former for the high value of the status quo and latter for the low value of the status quo. This switch between advisors confirms our theoretically predicted state pooling effect. In particular, for a status quo value R the subject wants to learn whether the state-dependent payoff of the new policy is greater or lower than R. When subjects face

	Advisor X			Value		Simplicity measures		
Advisor	x_{bad}	x_{med}	x_{good}	$(R = 30)$	$(R = 65)$	Certainty	Discrete complexity	Continuous complexity
1	0	0	1	46.667	70	1	1	2.4142
2	0	1	0	46.667	65	1	1	2.4142
3	0	1	1	53.333	65	1	1	2.4142
4	0	0	0.5	46.667	67.5	0	2	2.7121
5	0	0	0.75	46.667	68.75	0	2	2.6667
6	0	0.5	1	50	67.5	0	2	2.7877
7	0	0.75	1	51.667	66.25	0	2	2.7522
8	0	1	0.75	49.167	65	0	2	2.7522
9	0.25	1	1	51.667	65	0	2	2.6667
10	0.5	1	1	50	65	0	2	2.7121
11	0	0.25	0.5	46.667	66.25	0	3	3.1093
12	0	0.25	0.75	46.667	67.5	0	3	3.0391
13	0	0.5	0.5	46.667	65	0	3	3.1213
14	0	0.5	0.75	46.667	66.25	0	3	3.0755
15	0.25	0.75	1	50	65	0	3	3.0391
16	0.5	0	0.5	46.667	65	0	3	3.1213
17	0.5	0.5	1	46.667	65	0	3	3.1213
18	0.75	0.25	1	46.667	65	0	3	3.0391
19	0.25	0.5	0.75	46.667	65	0	4	3.3854
20	0.25	0.25	0.25	46.667	65	0	4	3.4641

Table 5: List of advisors used in Tasks 2, 3, and 4, and their certainty and complexity scores. Each advisor is represented by a triplet of signal probabilities $p_i = Pr(\sigma = 1|s = i)$. The next two columns show the theoretical predictions for a Bayesian decision-maker: the expected value conditional on choosing the advisor based on the different status quo values (low R and high R). The last three columns consist of measures of simplicity: dummy for certainty, discrete complexity score from 1 to 4, and continuous complexity score, respectively.

a choice between a certainty state pooler[46] and certainty advisors, they select on average the certainty state pooler in 74% of the trials. This result appears at odds with the Experimental Result 7 (higher WTP for information on high-value states, in task 1). This suggests that, in case of conflict between state pooling and high-value preferences, the discrete choice task favors the first effect over the second one.

In the two scenarios of choice between the red-advisor and blue-advisor, we see that each option is chosen by more than one quarter of the participants, and this is at odds

[46] A certainty state pooler advisor is certain (can fully reveal one state, as previously defined) and the revealed state is the singleton one from the state pooling.

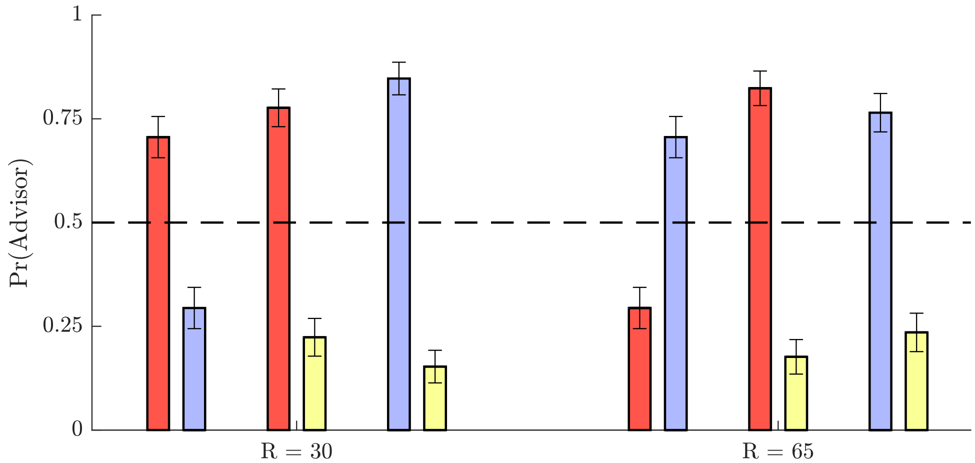


Figure 18: Comparison of advisors providing the ideal state pooling question: "Is the state red/yellow/blue ?" for two different status quo values. The color of the bar shows which state the question is about. The figure demonstrates the state pooling behavior, and also that participants do switch between advisors when it is valuable to do so.

with the model's prediction, especially in a trial with a relatively simple problem. This can be interpreted as a general signal of noise in the participants' actions, or as a systematic preference towards information about low or high states. Figure 19 suggests that the latter interpretation can partially explain the pattern. Starting from all the trials, we calculate for every subject the probability of choosing the advisor that is best under low or high status quo value.

If choices are just noisy, we should observe most of the subjects to be clustered around the coordinates 0.5-0.5, which would be consistent both with optimal behavior (always pick the best advisor), and with completely erratic choices (pick randomly). If participants have non-instrumental preferences over skewed sources of information (as shown by Masatlioglu, Orhun and Raymond (2017)), and such preferences are heterogeneous, we would expect a distribution of subjects that systematically deviate towards 1-0 (reveal information about the low state) and 0-1 (about the high state). In fact, we observe that participants deviate in both directions, and some of them also deviate towards lower probabilities in both dimensions - this can be the case when the chosen advisor is the worse choice under both status quo scenarios.

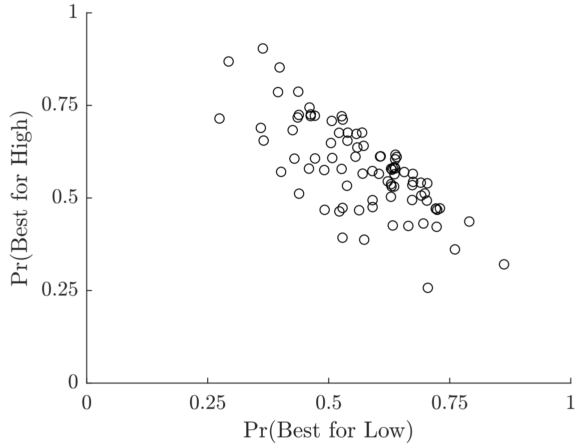


Figure 19: Distribution of participants' advisor choices: probability of choosing the advisor that provides more information about the low or high state in different types of trials.

G Appendix: Further Analysis on Risk Attitude

In Section III.B we discussed how risk preferences could represent an explanation for the subjects' deviation from the model's predictions. We provide here further details about how subjects' actions are, on average, not characterized by a significant deviation from risk neutrality.

Figure 20a shows the realized probability of selecting the risky option as a function of the difference in the EVs between the actions. Trials are grouped based on the x-axis value for visualization purposes. The optimal agent would have a sharp jump in probability from 0 (when the value difference is negative) to 1. We observe a smoother transition in our data, suggesting that action probability is modulated by the cost of mistakes, similarly to our discussion in Figure 5b about the choice between advisors. Such a sigmoid curve is normally found in experiments involving choice under risk (Mosteller and Nogee, 1951; Khaw, Li and Woodford, 2019).The indifference point appears close to the trials in which both actions have the same values, suggesting that the participants are overall close to risk neutrality.

We replicate the analysis for the choices made in task 1. An advantage of this dataset is that we observe two types of action choice scenario.

First, when the advisor confirms the color of the hidden ball, the decision maker faces a choice between two degenerate lotteries with different values, for example 80 points for sure

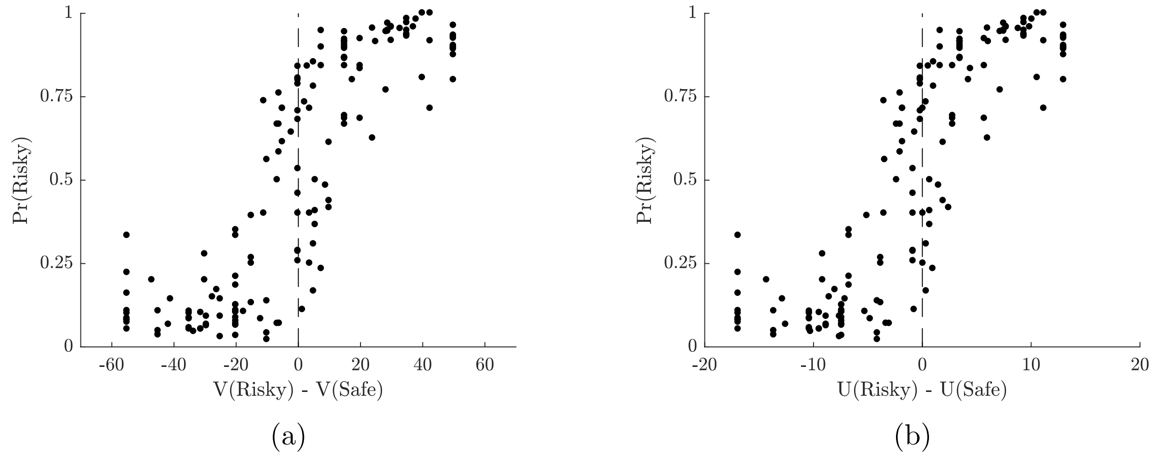


Figure 20: Task 2: Action selection probability. Left: Action choice under risk neutrality. Observed probability of choosing the risky action in task 2. 6,800 observations unequally divided across 160 cases (2 cases per trial, conditional on the advisor choice). Right: Action choice under risk aversion (best fit). The expected values for each action is replaced with the expected utility, with CRRA utility and the MLE coefficients $\hat{\alpha} = 0.34$ estimated from the dataset.

(risky action if you know the color is blue) or 60 points for sure (safe action). In these cases, participants pick the best option 90% of the times, confirming the small amount of noise in the action implementation in these simple choices.

Second, all the participants encounter choices with full uncertainty about the color (decision without any hint) or with a hint about two possible colors (e.g. red or yellow with equal chance, but not blue). The participants pick the option with the highest expected value 84% of the times, and we encounter again the sigmoid curve discussed above. The MLE of the risk aversion parameter under CRRA utility is $\hat{\alpha} = 0.52$.

H Appendix: Further Analysis of Beliefs

In Section III.B we discussed how subjective beliefs could represent an explanation for the subjects' deviation from the model's predictions. We provide here further details about how subjects' subjective beliefs elicited in tasks 3 and 4 display, on average, only mild evidence of conservatism.

In both tasks we observe accurate probability estimates, close to the predictions of an

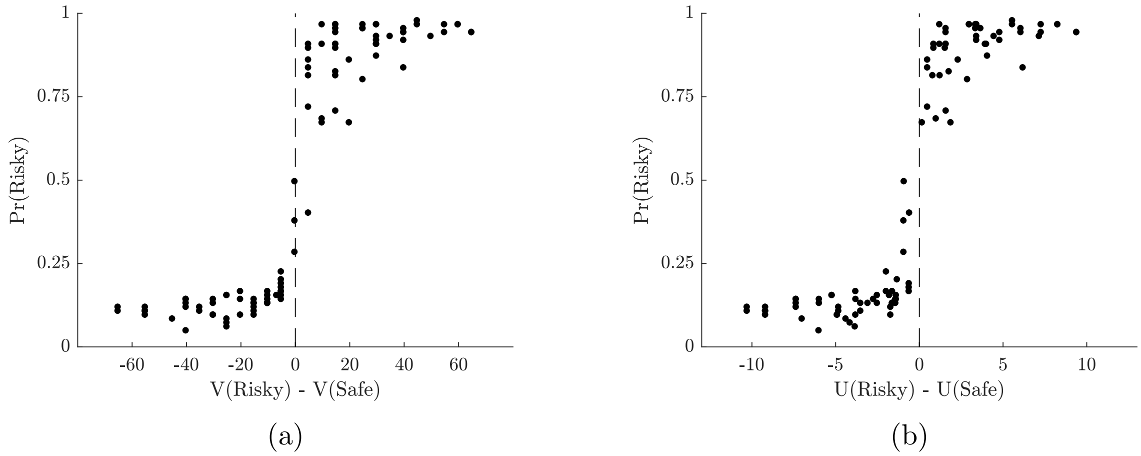


Figure 21: Task 1: Action selection probability. Left: Action choice under risk neutrality. Observed probability of choosing the risky action in task 1. 100 questions (10 questions for each of the 10 trials), 85 observations per question. Right: Action choice under risk aversion (best fit). The expected values for each action is replaced with the expected utility, with CRRA utility and the MLE coefficient $\hat{\alpha} = 0.52$ estimated from the dataset.

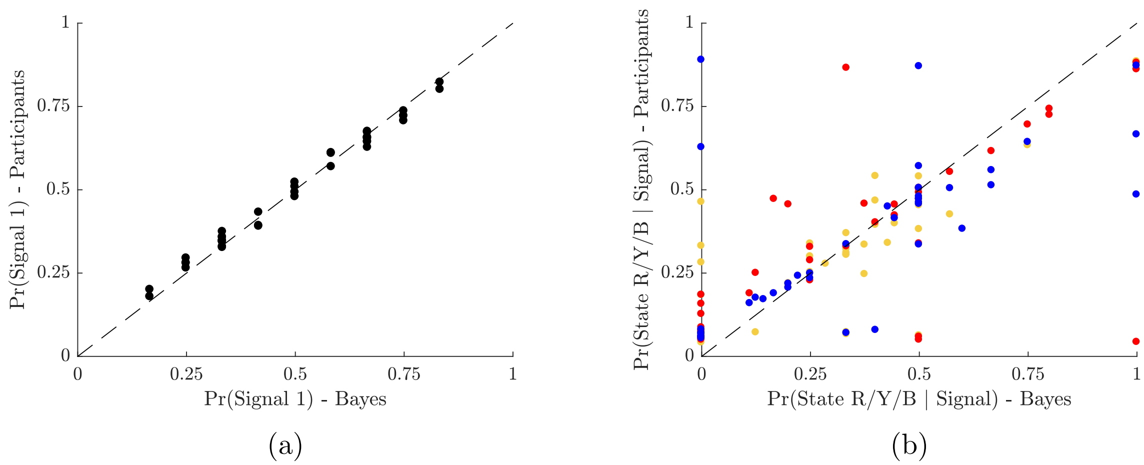


Figure 22: Average subjective beliefs. Left: Estimated probability of receiving a signal realization in Task 3. The plot compares the average of the subjective estimates collected with the optimal estimates of a Bayesian decision maker. $1,700$ observations across 20 trials (85 observations per point). Right: Estimated posterior probability of each state in task 4 conditional on the realized signal. Colors indicate which state was estimated (red, yellow, blue). The plot compares the average of the subjective estimates collected with the optimal estimates of a Bayesian decision maker. $20,400$ observations across 40 trials (6 observations per trial, 85 observations per point in the plot).

optimal Bayesian agent. Figure 30a shows the subjective estimate of a signal realization (y-axis, averaged across participants) compared to the optimal estimates (x-axis). Similarly, Figure 30b shows the subjective estimate of each of the three possible states in the posterior compared to the unbiased posterior, with different colors in the figure matching the state. In both plots, the 45 degree lines represent our theoretical benchmark and we can see that 1) participants are on average accurate in the estimate of probabilities, 2) we do not observe a systematic difference between estimates involving different states (i.e., we do not have evidence of motivated beliefs, Bénabou (2015)), and 3) both tasks show mild evidence of conservatism (central tendency of judgement), as vastly reported in experiments with subjective estimates (Hollingworth, 1910; Anobile, Cicchini and Burr, 2012).

For the signal probability (task 3) a linear fit of the subjective estimates $\hat{p}$ over the true probabilities p returns the coefficients $\hat{p} = 0.041 + 0.918 \cdot p$ with $R^2 = 0.991$. For the state probability (task 4) the linear fit for the whole dataset returns $\hat{p} = 0.058 + 0.825 \cdot p$ with $R^2 = 0.993$. The slopes are not significantly different across the three types of states: $\beta_{red} = 0.831$, $\beta_{yellow} = 0.805$, $\beta_{blue} = 0.827$.

I Appendix: Further Analysis of Willingness to Accept

In this section we add further results from the analysis of the willingness to accept renunciation of an advisor in task 1. We reported in Figure 9 and Table 3 that WTA is characterized by compression, a conservatism in the evaluation of the instrumental value of an advisor that leads to overpayment for the advisor with little or no informative value.

This result is robust across subjects, as displayed in Figure 23. For every subject, we estimate the sensitivity to the instrumental value by using a simple OLS regression of the subjective evaluation $V^j(i)$ over the instrumental value $V^{\text{Bayes}}(I)$. The graph shows the cumulative distribution of the fitted slopes, where 0 indicates no response to the true value and 1 indicates full alignment between the two variables. 82% of the participants show values between 0 and 1.

Another effect discussed in the paper is the excessive value assigned to blue advisors (that reveal the high-payoff state) and rainbow advisors (that provide full disclosure). Using Equation 5 we can easily recognize that the value of the rainbow advisor is equal to the highest value among the other three advisors. Participants do not seem to follow this rule,

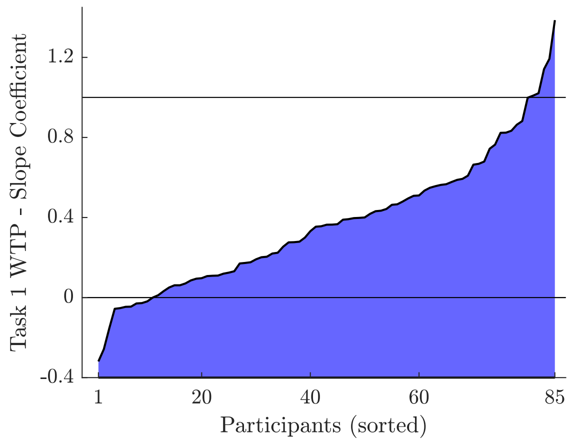


Figure 23: Distribution of participants' responses to the instrumental value of the advisors in Task 1.

as they tend to pay much less for the rainbow advisor. Figure 24 shows the distribution of the differences, within each trial, between the WTA for the rainbow advisor and the maximum of the other three WTA. Participants are willing to pay strictly less 39% of the times, and they are willing to pay strictly more only 17% of the times.

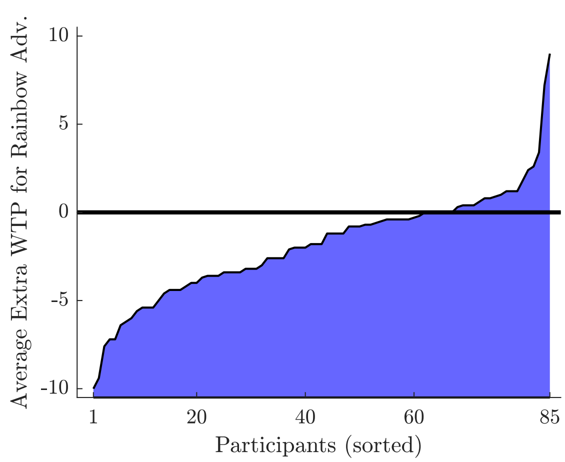


Figure 24: Distribution of participants' average excessive WTP for the rainbow advisor with respect to the highest WTP among the three simple advisors.

Finally, we want to show that the extra WTA for blue and rainbow advisors is due to a fixed premium that participants are willing to add, and not because of different elasticity to the instrumental values. In Table 3 we assumed that advisors' values can have different intercepts but share the same slope. We relax the assumption in a series of regressions displayed in Table 6. Compared to the benchmark model (column 1, same slope and intercepts for all), we notice a major improvement in the fit when we add the advisor-specific intercepts, which is not as much as for advisor-specific slopes.

Method: OLS, Dependent variable: $V^i(I)$

	(1)	(2)	(3)	(4)
Constant	6.66***	5.62***	6.80***	5.65***
	(0.166)	(0.231)	(0.167)	(0.257)
Red - constant		-0.193		-0.496
		(0.353)		(0.447)
Blue - constant		3.41***		3.42***
		(0.349)		(0.422)
Rainbow - constant		2.74***		2.96***
		(0.373)		(0.505)
$V^{\text{Bayes}}(I)$ - slope	0.372***	0.269***	-0.123	0.222
	(0.027)	(0.030)	(0.164)	(0.172)
Red - slope			0.252	0.099
			(0.163)	(0.181)
Blue - slope			0.631***	0.038
			(0.164)	(0.180)
Rainbow - slope			0.550***	0.009
			(0.162)	(0.181)
Trials	All	All	All	All
Observations	2,520	2,520	2,520	2,520

Table 6: Aggregate valuations of information structures in task 1.
Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

J Appendix: Further Analysis of Questionnaire

We combine demographic information with two additional tasks (Holt-Laury test of risk attitude and Raven matrices as a measure of cognitive ability) and a final questionnaire with questions about the field of study, mathematical literacy, and other tests (Revised Life Orientation Test, LOT-R, as a measure of optimism, questions on superstition, and questions on risk attitude). Table 7 shows that neither of our measures of mathematical aptitude and cognitive style is significantly associated with our measure of within-subject polarization. Risk attitude, measured by the Holt and Laury test, shows a negative coefficient (a high risk seeking score is associated with a low polarization score), but the effect disappears once we introduce the demographic controls. Tables 8 - 12 show analogous analyzes for accuracy of advisor choice in task 2, probability of selecting the simple advisor in task 2, probability of selecting the risky action in task 1, accuracy in beliefs elicitation (tasks 3 and 4) and WTP slope in task 1, respectively.

Method: OLS, Dependent variable: Polarization score

	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			✓	✓

Table 7: Polarization score. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Method: OLS, Dependent variable: Accuracy of advisor choice in task 2				
	Baseline (1)	Full (2)	Baseline (3)	Full (4)
Risk attitude (Holt and Laury)	-0.36^{***}	-0.36^{***}	-0.28^{*}	-0.28^{*}
	(0.09)	(0.10)	(0.15)	(0.16)
Fluid intelligence (Raven test)	0.19^{**}	0.16	0.22^{***}	0.13
	(0.07)	(0.10)	(0.08)	(0.10)
Familiar with Bayes rule	0.09^{*}	0.11^{*}	0.17^{***}	0.20^{***}
	(0.05)	(0.06)	(0.06)	(0.06)
Analytical studies	−0.03	−0.03	−0.06	−0.06
	(0.05)	(0.05)	(0.05)	(0.06)
LOT-R scale		−0.01		−0.04
		(0.03)		(0.03)
SUPERSTITION scale		0.02		0.03
		(0.03)		(0.03)
RISK scale		−0.00		−0.03
		(0.03)		(0.03)
Observations	63	63	63	63
Demographic Controls			✓	✓

Table 8: Accuracy of advisor choice in task 2. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Method: OLS, Dependent variable: Probability of selecting the simple advisor in task 2				
	Baseline (1)	Full (2)	Baseline (3)	Full (4)
Risk attitude (Holt and Laury)	−0.07	−0.08	−0.17	−0.18
	(0.11)	(0.11)	(0.15)	(0.15)
Fluid intelligence (Raven test)	−0.06	−0.08	−0.08	−0.11
	(0.08)	(0.10)	(0.09)	(0.13)
Familiar with Bayes rule	−0.02	−0.01	−0.04	−0.03
	(0.04)	(0.05)	(0.05)	(0.06)
Analytical studies	−0.01	−0.01	0.00	−0.00
	(0.04)	(0.04)	(0.04)	(0.04)
LOT-R scale		0.01		0.01
		(0.02)		(0.03)
SUPERSTITION scale		0.02		0.04
		(0.02)		(0.03)
RISK scale		−0.01		−0.00
		(0.03)		(0.03)
Observations	63	63	63	63
Demographic Controls			✓	✓

Table 9: Probability of selecting the simple advisor in task 2. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Method: OLS, Dependent variable: Probability of selecting the risky action in task 1				
	Baseline (1)	Full (2)	Baseline (3)	Full (4)
Risk attitude (Holt and Laury)	0.03	0.03	−0.00	0.00
	(0.08)	(0.08)	(0.10)	(0.10)
Fluid intelligence (Raven test)	0.05	0.06	0.03	0.05
	(0.05)	(0.07)	(0.06)	(0.07)
Familiar with Bayes rule	−0.00	−0.02	−0.05	−0.06
	(0.04)	(0.04)	(0.04)	(0.04)
Analytical studies	−0.01	0.00	0.00	0.01
	(0.03)	(0.03)	(0.03)	(0.03)
LOT-R scale		0.01		0.02
		(0.02)		(0.02)
SUPERSTITION scale		−0.02		−0.02
		(0.02)		(0.02)
RISK scale		−0.01		−0.01
		(0.02)		(0.02)
Observations	63	63	63	63
Demographic Controls			✓	✓

Table 10: Probability of selecting the risky action in task 1. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Method: OLS, Dependent variable: Accuracy in beliefs elicitation (tasks 3 and 4)				
	Baseline (1)	Full (2)	Baseline (3)	Full (4)
Risk attitude (Holt and Laury)	−0.59	−0.60	−0.40	−0.51
	(0.68)	(0.63)	(0.50)	(0.53)
Fluid intelligence (Raven test)	0.45	0.45	0.34	0.48
	(0.39)	(0.41)	(0.35)	(0.35)
Familiar with Bayes rule	0.26	0.35*	0.37	0.42
	(0.20)	(0.20)	(0.26)	(0.30)
Analytical studies	−0.15	−0.21	−0.30	−0.30
	(0.19)	(0.19)	(0.20)	(0.20)
LOT-R scale		−0.03		0.07
		(0.13)		(0.12)
SUPERSTITION scale		0.13		0.11
		(0.14)		(0.15)
RISK scale		0.10		0.14
		(0.09)		(0.11)
Observations	63	63	63	63
Demographic Controls			✓	✓

Table 11: Accuracy in beliefs elicitation (tasks 3 and 4). Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Method: OLS, Dependent variable: WTP slope in task 1	Baseline (1)	Full (2)	Baseline (3)	Full (4)
Risk attitude (Holt and Laury)	-0.42^{**}	-0.39^{*}	-0.12	-0.15
	(0.20)	(0.21)	(0.27)	(0.28)
Fluid intelligence (Raven test)	0.27	0.31	0.31^{*}	0.33
	(0.18)	(0.24)	(0.16)	(0.21)
Familiar with Bayes rule	0.20^{*}	0.17	0.14	0.15
	(0.10)	(0.11)	(0.12)	(0.13)
Analytical studies	-0.20^{**}	-0.16	-0.20^{**}	-0.16^{*}
	(0.10)	(0.11)	(0.09)	(0.10)
LOT-R scale		0.02		0.03
		(0.07)		(0.06)
SUPERSTITION scale		-0.06		-0.03
		(0.05)		(0.06)
RISK scale		-0.02		-0.01
		(0.05)		(0.06)
Observations	63	63	63	63
Demographic Controls			✓	✓

Table 12: WTP slope in task 1. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

K Appendix: Further Analysis of Heterogeneity Across Subjects

We extend Figure 6a by also displaying the minimal and maximal realized polarization among all the subjects (see Figure 25). We can observe that for each trial and state pair, there is a subject whose realized polarized polarization is zero or very close to zero (min=0). The subject with maximal realized polarization in each trial and state pair is close to the 45-degree line and several times above it, polarizing even more than predicted.

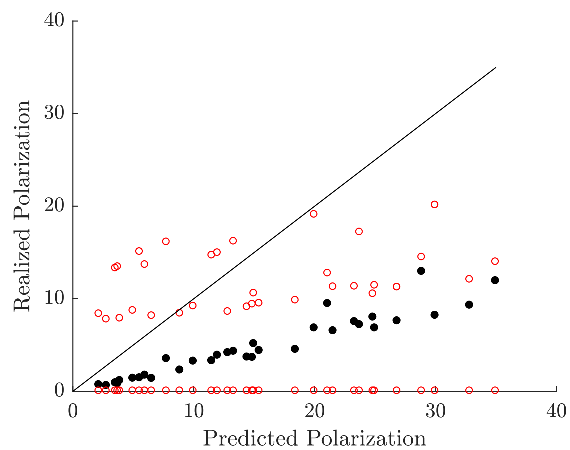


Figure 25: Polarization. Predicted polarization (for the Bayesian decision maker) and realized polarization (based on subjects' responses, n=85) in the 11 pairs of trials with predicted advisor switches (3 states per pair of trials). Average (black circles) and the minimal and maximal realized polarization in the corresponding trial and state across subjects (red circles).

Further we conducted a cluster analysis of the participants' advisor choices using two approaches. In the first approach, we use simple clustering based on two dimensions depicted in Figure 11, i.e., the probability that the best advisor is selected and the probability of selecting the simplest advisor.

As we describe in the main text (Section IV: Heterogeneity across subjects), the participants can be categorized into three broad groups based on these two dimensions. A cluster of accurate participants that display little or no bias on the right side - Green cluster, a group of simplicity-driven participants consistently selecting the advisor with lower complexity on the top - Red cluster, and a smaller group of participants whose advisor choices are close to

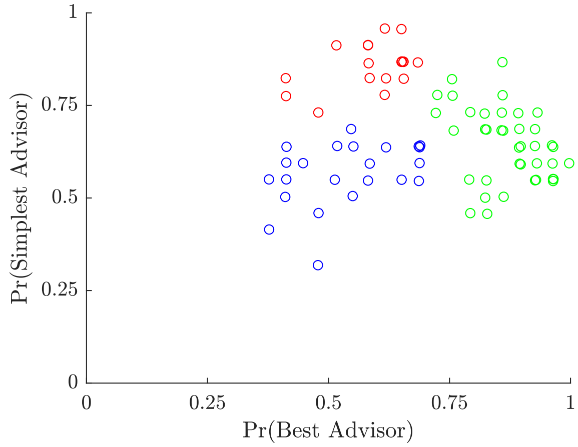


Figure 26: Distribution of participants' advisor choices: the probability of choosing the best advisor (based on instrumental value) and simplest advisor (based on the complexity score). Clustered based on the approach 1.

	Population	Cluster		
		Red	Blue	Green
N	85	20	23	42
% best advisor	0.721	0.593	0.54	0.881
% simple advisor	0.671	0.868	0.563	0.636
% polarization (with Subjective beliefs)	0.531	0.302	0.366	0.731
% muted polarization (with Bayesian beliefs)	0.712	0.576	0.596	0.841
Avg beliefs slope in task 4 (1=Bayesian)	0.872	0.801	0.821	0.933
Avg Raven score	0.424	0.33	0.383	0.49

Table 13: Summary statistics for each cluster based on the approach 1.

random - Blue cluster.

We further explore to what extent they show signs of updating beliefs in a Bayesian fashion and at the same time how much the polarization is mitigated for the participants corresponding to these three clusters.

Figure 27 and table 13 indicate that participants corresponding to all three groups are updating beliefs close to the Bayesian fashion, but those from the blue and the red clusters deviate slightly more from Bayesian updating than the participants from the green cluster. This together with advisor choices of each cluster provides an explanation for the degree of polarization mitigation. In particular, the green cluster's polarization is closest to the predicted one and the red cluster's polarization is the most mitigated. This is due to the fact that the green cluster is best in selecting the best advisor and preference for the simplest

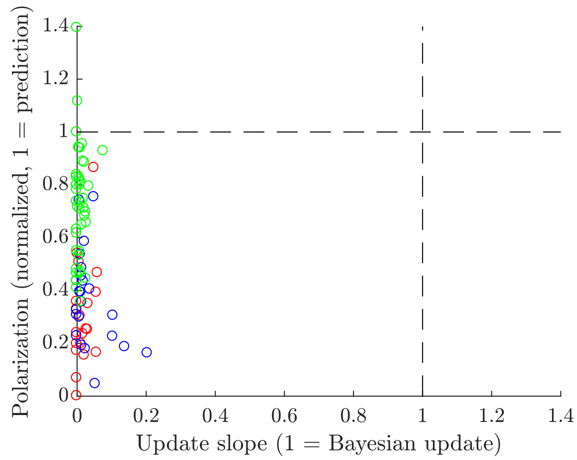


Figure 27: Distribution of participants' polarization and belief updating. Clustered according to participants' advisor choices - approach 1

advisor is a little bit above one-half. The blue cluster on one side has almost no preference for the simplest advisor but identifies the best advisor almost randomly. In addition, deviates non-negligibly from Bayesian updating. Thus, the blue cluster ranks second in the degree of mitigated polarization. The red cluster has the most mitigated polarization, as it demonstrates the strongest preference for the simplest advisor while being only slightly above the coin flip from identifying the best advisor and being non-Bayesian on a similar level as the blue cluster. In table 13 we report how much the mitigated polarization changes when subjective beliefs are replaced by Bayesian beliefs. Importantly, the green cluster accounts for almost half of all participants. These participants also show a marginally higher average Raven score.

	Population	Cluster		
		Red	Blue	Green
N	85	23	18	44
% best advisor	0.721	0.654	0.477	0.856
% simple advisor	0.671	0.848	0.576	0.618
% polarization (with Subjective beliefs)	0.531	0.341	0.337	0.71
% muted polarization (with Bayesian beliefs)	0.712	0.616	0.58	0.817
Avg beliefs slope in task 4 (1=Bayesian)	0.872	0.868	0.739	0.928
Avg Raven score	0.424	0.4	0.378	0.455

Table 14: Summary statistics for each cluster based on the approach 2.

In the second approach, we do clustering based on the vector of all the choices for each subject. Each subject's vector of all the choices consists of 40 choices, where each is 0 or 1

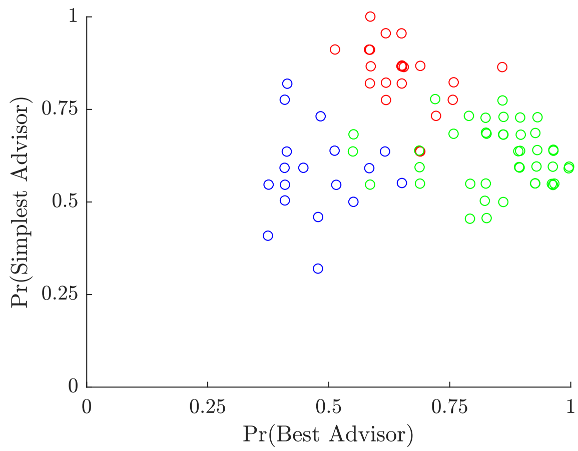


Figure 28: Distribution of participants' advisor choices: the probability of choosing the best advisor (based on instrumental value) and simplest advisor (based on the complexity score). Clustered based on the approach 2.

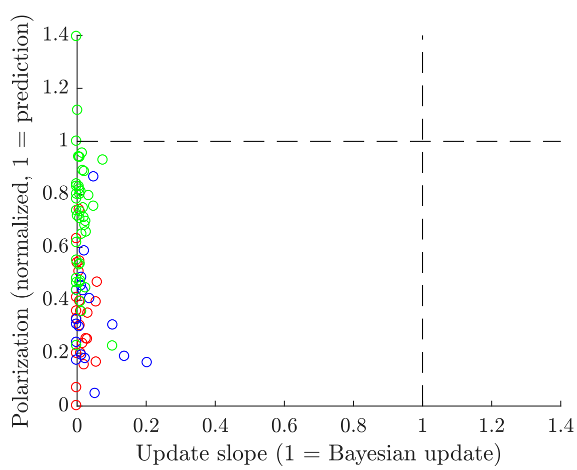
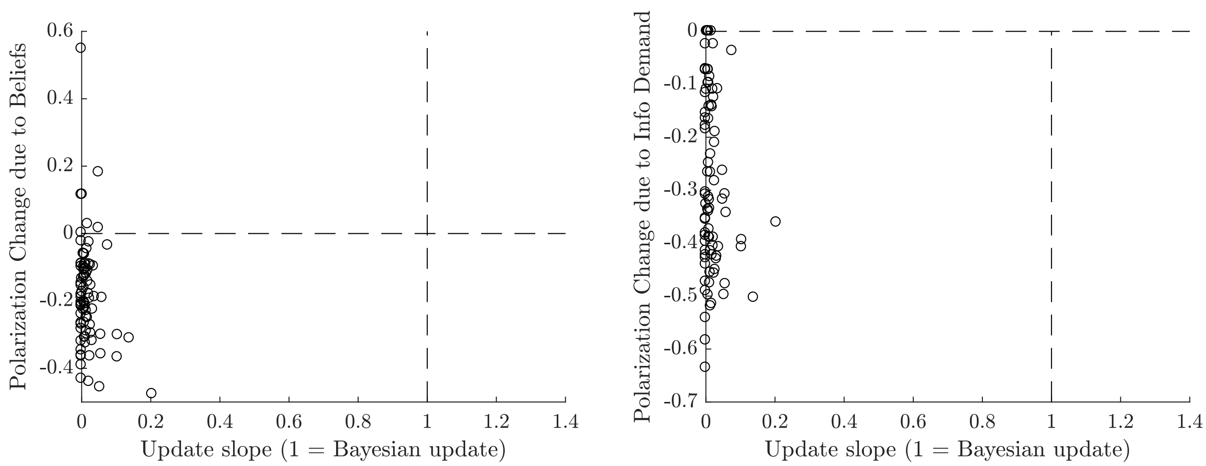


Figure 29: Distribution of participants' polarization and belief updating. Clustered according to participants' advisor choices - approach 2

(referring to the advisor selected).

We report the same figures as in the previous case (Figure 28-29) and summary statistics in table 14. The main difference between approaches 1 and 2 is in the probability of selecting the best advisor for clusters and that the red cluster - with a high preference for the simplest advisor is more numerous.

In the rest of this appendix, we explore participants' over-reaction to the evidence as we documented by polarization exceeding the predicted values (above 1 on the normalized score) in Figure 10a. Specifically, if non-Bayesian agents over-react to the evidence received with the signal, they can reach more extreme posterior beliefs than the ones of a Bayesian agent. In this case, we would observe an actual magnitude of polarization that is larger than the predicted one. From figure 27, we observe that most of the participants show some conservatism in their response, and over-polarization occurs in participants with conservative beliefs. The two channels that can affect (reduce/increase) polarization are 1) non-Bayesian update (yet here it typically reduces polarization due to conservatism), and 2) demand for information. Figures 30 below show the decomposition of these two effects.



(a) (b)

Figure 30: Decomposition of polarization according to two channels. Left: Change in polarization due to beliefs. The plot shows the difference between polarization with subjective beliefs and polarization with Bayesian beliefs, both normalized. Right: Change in polarization due to information demand. The plot shows the difference between polarization with Bayesian beliefs and model predictions, both normalized.

Both in the aggregate and at the subject level, non-Bayesian beliefs reduce the magnitude of polarization (with few exceptions, associated with conservatism). This is still possible because, despite the beliefs being on average conservative, there is dispersion around the regression curve and few participants over-react only to some evidence, but not in a systematic way.

Changes coming from the change due to information demand always reduce the magnitude of polarization, by construction (these are the trials in which the model always predicts the agents should switch, so there is no way for this channel to augment polarization).

L Appendix: Prediction of the status quo type

We now look at our framework from the perspective of a platform that wants to infer the type (status quo value) of the decision maker and has access to a dataset with some observable activities. We can divide these activities into two groups: final actions, like *voting* or the choice between the status quo and a new policy, and information acquisition, like *reading newspapers* or selecting an advisor in our design. For a more concrete example, imagine a social media platform like Facebook or Twitter, that has access to a dataset of actions performed by its users. These actions include publicly observable actions (likes, list of friends or followers), but also a series of additional actions (clicks, searches) that involve the process of information acquisition. Are these search activities helpful in improving the prediction of the type of the user, on top of the observable actions?

	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%

Table 15: Inference of the agent's status quo: predicted and realized accuracy (pairs of trials with expected advisor switch only). The table indicates the accuracy of the prediction of the type (status quo) of the decision maker based on the data available. The model's predictions are based on rational and unbiased agents. The accuracy realized refers to the dataset collected in the laboratory experiment.

We consider separate scenarios in which the platform has access to choices only (opaque or transparent box), searches only (advisor X or advisor Y), or both, under the assumptions of our model (rational decision makers) and in the dataset collected in the laboratory experiment. Table 15 shows the results of this exercise: having access to the search data guarantees a much higher accuracy with respect to the action data, with minor improvements when both datasets are available. When we consider the trials in which we expect to observe an advisor switch (column 2), the type prediction accuracy with advisor choices is 68%, but it is only 62.6% when we observe only the final actions. When both datasets are available, the accuracy increases marginally to 68.4%, with a further improvement if the signal realization (that occurs between search and choice) is also observed.

We can conclude that, in this simple setup, the data about the choice over sources of information is more valuable than the final action from the perspective of an observer who wants to infer the type (status quo value) of the decision maker.

M Appendix: Timeline of the Problem

Our experimental design allows us to estimate how agents evaluate an informative signal structure (advisor), and measure how the subjective evaluation depends on the properties of the signal structure, including instrumental value (expected improvement in the choice process) and non-instrumental properties (ease of interpretation).

The timing of the problem (as in task 2) can be summarized as follows:

1. The agent is informed of the prior $\mathcal{P}(s) = \frac{1}{3}\ \forall s$ and the state-contingent returns $\{v_s\}_S,\ R$.
2. One state is realized, but the agent is unaware of it.
3. The agent is offered two sources of information (advisors) I_1 and I_2.
4. The agent chooses one advisor and discards the other.
5. The selected advisor observes the realized state (ball in the opaque box).
6. The selected advisor returns a binary signal, whose likelihood depends on the realized state.
7. The agent observes the realized signal.
8. The agent chooses one action (opaque or transparent box) and receives the payoff π.
9. The agent plays a lottery and receives the final prize k with probability $\frac{\pi}{100}$.

The problem presented in task 1 is similar up to a change in steps 3 and 4:

3'. The agent is offered one single source of information (advisor) I

4'. The agent indicates how much she is willing to accept renunciation of the advisor.

In the *Colorblind advisor game* (task 1), we elicit the probability w_I such that the agent is indifferent between making a choice after observing the realization of a known signal structure I, and choosing without additional signals but receiving additional w_I tickets to win the prize. In the *Imprecise advisor game* (task 2), we offer pairs of signal structures, and collect binary choices between advisors. If the valuation and choices differ from those a Bayesian expected utility maximizer would display, we would like to pinpoint the source of the deviation. For this reason, we add two control tasks to elicit a subjective signal of beliefs' realization (*Card color prediction game*, task 3) and subjective posterior beliefs (*Ball color prediction game*, task 4). We collect posteriors only after eliciting preference over advisors, so we do not nudge the subjects towards thinking about information valuation in a specific fashion.