‘Identify the Expert’: an Experimental Study in Economic Advice

Theodore Alysandratos, Aris Boukouras, Sotiris Georganas and Zacharias Maniadis*

November 14, 2023

Abstract

We develop the novel ‘identify the Expert’ task and run preregistered online experiments on a representative sample. Participants receive recommended answers to an economics questionnaire by two computerized advisors. One advisor is of high-accuracy (‘the Expert’) and recommends the answers produced by academic consensus. The other advisor is of low accuracy (‘the Populist’), and recommends the modal answers of lay participants from a pilot study. Participants do not know who the Expert is, and have to judge this from the recommendations. We examine which advisor participants identify as the Expert via revealed preference, i.e. participants select an advisor to answer the questionnaire on their behalf. Participants overwhelmingly choose the Populist, even when fully informed about the advisors’ modus operandi. Bayesian models fail to explain these choices, even in the degenerate case where participants should be able to identify the Expert with 100% accuracy. Overconfidence in one’s ability and ego-involvement do not fully account for participants’ difficulty in identifying the expert, although they make the task harder. These results are relevant for a wide range of everyday expert selection tasks.

Keywords: Democracy, Economic Literacy, Expert Advice, Populism.

JEL Codes: C91, A11

*Alysandratos: Department of Economics, University of Heidelberg, Germany. Boukouras: Department of Economics, Finance and Accounting, University of Leicester, UK (Corresponding Author). Email: Aris.Boukouras@leicester.ac.uk. Georganas: Department of Economics, City, University of London, UK. Maniadis: Department of Economics and SInnoPSis, University of Cyprus, and Department of Economics, School of Social Sciences, University of Southampton. This work was ethically approved in July 2021 as: ‘ 31171.A1 - A follow-up study of knowledge of economics and economic advice. (Amendment 1)’ by the University of Southampton.

1 Introduction

τότ' ἔφη τὰς πόλεις ἀπόλλυσθαι, ὅταν μὴ δύνωνται τοὺς φαύλους ἀπὸ τῶν σπουδαίων διακρίνειν
Cities, said he, fail when they cannot distinguish fools from great men.

Antisthenes

Democracies are faced with a constant trade-off between involving expert technocrats in shaping public policy and giving a meaningful voice to laypeople (Caplan, 2011; Chakraborty et al., 2020). This tension has not faded in the last decades. On the contrary, scholars suggest that the advent of social media has amplified the gap between experts and laypeople, as public debate is increasingly shifting from traditional media to online platforms (Gillespie, 2018; Allcott et al., 2020). We unpack a specific aspect of this tension, by asking whether laypeople are able to distinguish the opinions of experts from those of other laypeople. In a set of preregistered online experiments we show that a populist non-expert whose advice agrees with the public's priors but has no concern for the truth, outperforms a true expert systematically.

We focus on economic expertise, which is presumably of high importance in modern democracies, and present participants with no other information except for the advisors' opinions. In our treatments we find that most participants robustly choose the non-expert as advisor, mainly because they fail to discover a simple but powerful heuristic: in the presence of populists, choosing advisors you agree with is only a good idea if you are knowledgeable enough already. The implication is that in technical and counterintuitive subjects (such as economics) most people, not being knowledgeable, should be choosing advisors *they actually disagree with*!

To be more precise, we employ an economics questionnaire (validated by experts) in a two-stage experiment with participants from the general population in England and Wales. In Stage 1, participants first provide their own answer and then see the suggested answers of two computerized 'advisors' on the same questions. One advisor is the 'Expert', who is designed to give the answer deemed correct by the academic consensus.[1] The other advisor is the 'Populist' who, for each

[1] We distributed an expanded version of our questionnaire to academics in economics departments in Europe and the USA. Questions which received more than 70% agreement were validated and these answers became the choices of the Expert. The online Appendix describes the details of our validation exercise.

question, proposes the most popular answer from a pilot study. Participants do not know which advisor is the Expert, they only see in words the recommendations that the two advisors make. Furthermore, they receive no other information that could operate as a cue, such as credentials, visual characteristics or exhibited confidence. Subsequently, in Stage 2, after observing a summary of the recommendations by both advisors, participants are asked to pick one advisor to answer all of the questions on their behalf. They are financially incentivized to answer correctly in Stage 1 and to select the Expert in Stage 2.

We deploy three experimental treatments to examine a number of pre-registered hypotheses. Treatments are identical with respect to Stage 1, but vary the information presented to participants in Stage 2. In the baseline treatment (Treatment 1), participants review the answers they provided to each question and the recommendation of the two advisors from Stage 1, but do not know how accurate their own responses to the questionnaire have been. In Treatment 2, they learn how many questions they answered correctly in Stage 1, in addition to the information of the baseline treatment. Finally, Stage 2 of Treatment 3 provides similar information as Stage 2 of Treatment 2, but with respect to another participant. That is, in Stage 2 of the third treatment, each participant sees the summary table of answers and the two advisors' recommendations referring to a third participant from a prior experimental session. They also see how many correct answers that third participant had in Stage 1. For future reference, we shall call the person whose choices a participant observes in Stage 2 (the participant themselves in Treatments 1 and 2 or the third person in Treatment 3) as the Decision-Maker.

The baseline treatment tells us whether people can distinguish the Expert from the Populist. Treatment 2 examines the extent to which overconfidence in their ability to answer the questionnaire drives performance in selecting the Expert. Treatment 3 tests whether presenting the problem from a third-person perspective helps reduce ego-involvement and improve performance (choosing the Expert) in Stage 2. Our results indicate that in economic matters people have a strong tendency to follow advisors who suggest similar answers to their own priors, which often leads to a mistaken choice of advisor. Indeed, in all three treatments, the percentage of participants who chose the Expert in Stage 2 is significantly below 50%. On average, *a participant would do better by selecting an advisor at random rather than using their intuition.* In our experimental environment, participants are fully informed of how the two advisors choose their answers. Sophisticated individuals lacking expertise in economics should infer that the Decision-Maker is more likely to have more common answers with the Populist than with the Expert. This rational inference requires no feedback or experience. Yet, the large majority of participants fail to apply it correctly in this experiment.

In Treatments 2 and 3 in particular, participants know that the Expert is always correct and they also know the Decision-Maker's number of correct answers. Sophisticated participants should infer that the Expert is the advisor who has as many common answers with the Decision-Maker as the latter's number of correct answers. Nonetheless, participants' performance improves only marginally in comparison to Treatment 1, and it does not exceed 50% in either treatment. Hence, the task is intrinsically hard for most subjects. This finding, however, does not mean that overconfidence and ego-involvement play no role at all. When participants should pick the advisor with the least number of common answers with the Decision-Maker, both treatments have a significant positive effect on identifying the Expert. Thus, these two biases complicate further an already hard task for many participants.

There are several possible factors that can influence laypeople to ignore the advice of the experts. Our experimental design shuts down many of them and focuses on the counter-intuitive nature of following someone you disagree with. Reputation concerns and conflict of interest are absent in our setting, while we have designed treatments to counter confirmation bias and overconfidence. In addition, complex mathematical computations are not required for providing a correct answer in treatments 2 and 3, and several of our findings would not emerge if participants were not paying attention to the task. The most probable explanation is that laypeople find it hard to select an advisor they disagree with because the reasoning behind the task (i.e. having few common opinions with an expert if you are not knowledgeable on a subject yourself) is alien to them.[2]

Moreover, these preregistered results were obtained in a concrete applied setting, using a relatively large sample of the general population and focusing on a single aspect: the content of advice. In addition, the experiments were conducted online. There are two important reasons for these choices. Firstly, we want to abstract away from the multitude of reasoning and psychological factors which underpin evaluations of credibility (see Bonaccio and Dalal, 2006 for a review), so as to build a valid benchmark for subsequent research. Secondly, social media platforms, which are often suspected as the main culprit for the public's mistrust of experts (Allcott et al., 2019; Pennycook et al., 2020), naturally lend themselves to these features. Therein laypeople exchange opinions and views on all sorts of issues, with no physical presence, and with scant cues on one's knowledgeability other than one's own opinion on the issue at hand. Therefore, it is sensible to ask whether laypeople can distinguish expertise in such an information setting. Thus, our experimental

[2]A full discussion of the various channels behind our results can be found in section 3.6.

design likely contributes to the ecological validity and generalizability of our results. If indeed these results generalize, it appears likely that a shrewd populist, who promotes the most popular view on any issue, would have a good chance of gaining public support in a contemporary western democracy.

To our knowledge, this is the first experimental study of the 'identify the Expert' problem: one where participants try to distinguish who the real expert, i.e. without knowing which degree of accuracy corresponds to which individual advisor. Prior studies of expertise typically identify the source of information (and its accuracy) and examine the effect on participants' beliefs or choices regarding the subject of interest (Algan et al., 2021; Bailey et al., 2022; Amaral-Garcia et al., 2022). Our theoretical and empirical results regarding Treatment 1 (in particular, the popularity of the populist) are consistent with the model by Gentzkow and Shapiro (2006) where, in the absence of any feedback, people simply infer that sources of information closer to their priors are the correct ones. However, in our two further treatments the additional information provided is sufficient to identify the Expert regardless of one's priors. Our participants' systematic failure to do so indicates the possible existence of a deeper behavioral mechanism, which we discuss in section 3.

Charness et al. (2021) is the study closest to ours. They present participants with two distinct information structures and ask them to select the one with the highest overall accuracy. In their study, subjects frequently select the sub-optimal information structure and stick close to their priors, similarly to our findings. However, there is a crucial difference between their decision problem and ours. In their paper participants are informed exactly about the statistical properties of each structure in the task description, while in our paper participants do not know which statistical properties belong to which advisor. They need to infer them indirectly by comparing the similarity of their responses to the advisors' responses. Indeed, theoretically, the two decision tasks are distinct since they give out different optimal responses. In the case of *bias by commission* (i.e. when the information is biased in one direction), which is the only case we study in all treatments, a decision maker should always pick the advisor closest to their priors in the task of Charness et al. (2021). In contrast, this is not always optimal in our task, where the optimal heurestic depends on the quality of the subjects' priors. Furthermore, we argue that our setting is more relevant for certain applications of interest. For example, when evaluating political contestants citizens rarely have access to the full track record of the candidates. Instead, they need to evaluate who they deem more competent based on the similarity of the candidates' political positions with their own. Finally, our findings provide an important insight for the broader literature in the area. In Charness et al. (2021), when information structures are Blackwell-ranked (i.e. one less informative than the other), an overwhelming fraction of participants identify correctly the optimal one, whereas in our treatments 2 and 3 most subjects fail to answer correctly even though a Bayesian decision maker can achieve full accuracy. Thus, the problem of advisor selection is a cognitive hard task even for statistically degenerate distributions, exactly because the correspondence between advisor-identity and advisor-accuracy is not known.

Our findings also contrast to that of Meloso et al. (2023), where the failure of advisees to correctly infer the type (informative or not) of the unique advisor is rooted in their miscalibrated beliefs regarding the strategic behavior of advisors. Our advisors are computerized, their modus operandi is known to the participants and in two of our treatments this knowledge is redundant with regards to identifying the Expert.

Populism as a subject area has received significant attention by researchers lately. Examples of this include Mudde and Kaltwasser (2017), Fetzer (2019), Autor et al. (2020), and Funke et al. (2020). Guriev and Papaioannou (2022) provide an extensive literature review and pose some open questions for new research. For instance, why do populists exhibit heterogeneous policy platforms, covering both extremes of the political spectrum? We examine this question from the demand side and provide a partial answer as follows. Since the median voter is more likely to deem a policy that panders to her priors as a good one, a populist selecting the most popular opinion on every policy dimension improves their election chances considerably. This strategy implies substantial variation in populists' policies, as the median voter's preferences change across regions and times.

Our work builds naturally on an established literature, which documents differences in beliefs and perceptions between professional economists and laypeople. Sapienza and Zingales (2013) show that the median American has different views from the median American economist. Andre et al. (2022) document the diversity of opinions of laypeople in terms of the implicit models they use for the macro-economy and how they diverge from economists' models. Within this literature, the dispersion of beliefs both within laypeople and academics is well documented (Blendon et al., 1997; Gordon and Dahl, 2013; Angeletos et al., 2021). We acknowledge this divergence and through our validation exercises, we document it as well. But our focus is the policy implication it generates when populists exploit the opinion gap between laypeople and experts in order to direct attention to themselves.

We also examine the problem of expert choice in a concrete applied context with clear theoretical predictions, while the economics literature has hitherto focused on abstract or hypothetical domains. Chakraborty et al. (2020) study theoretically the role of experts in electoral competition and contrast it to a populist alternative. Unlike their setting, the interests of experts and participants are aligned in our experiment and there is no potential benefit from picking the non-expert. Ronayne and Sgroi (2018) and Schotter (2003) examine how individuals respond to advice, while we are interested in a different question, employing a design where advisors effectively compete for attention. Krishna and Morgan (2001) and Gentzkow and Shapiro (2006) study competition for information provision when information sources are biased and care for their reputation. Unlike these papers, our participants are not necessarily sophisticated, and we are interested in the demand side of the problem (i.e. how advisees chose whom to heed) rather the supply side (i.e., optimal strategies for advisors). Indeed, in our experiments, participants are informed about the exact modus operandi of the two advisors and there is no conflict of interest. Recent empirical studies, such as Aksoy et al. (2020) and Algan et al. (2021), examine how laypeople come to trust expert advice and what factors may underline such decisions. However, in these papers, competition between different parties for unsolicited advice-giving is not examined.

Our study is also related to the burgeoning literature on competing media sources and polarisation. Several papers assume or investigate people's preference for bias (Oliveros and Várdy, 2015; Chopra et al., 2022; Bursztyn et al., 2022; Thaler, 2021). In our paper we shut down the preference channel using Treatment 3, since participant choice of advisor is disconnected from their own views on the questions and participants are explicitly rewarded for choosing an "objectively" correct source. Our results indicate that at least part of the problem lies in people's inability to identify good sources of information, not merely in motivated beliefs.

The remainder of the paper is organized as follows. Section 2 presents the theoretical framework with our main predictions and describes the conducted experiments in detail. Section 3 presents the empirical results, while section 4 concludes.

2 Theory and Experimental Design

The basic design is a simple two-stage computerized and incentivized experiment with the following structure. Participants are exposed to an economics questionnaire of ten multiple-choice questions, with two options each.[3] In Stage 1, they are asked to answer each of these ten questions without any feedback, and to record their confidence in their answer on a scale from 0 to 100. After answering

[3]See section 2.3 for more details about the questionnaire, which we first introduced in Alysandratos et al. (2020). The full set of questions are available in our online Appendix.

each question, they observe the proposed answer by two 'advisors'. They are informed that one of them is a high-accuracy advisor, who answers all questions correctly. The other advisor is of low-accuracy, for whom participants know that he answers only four out of the ten questions correctly (i.e. according to academic consensus).[4]

Stage 2 follows after participants answer all questions in Stage 1. They then view a summary of information on the selections they made (in Treatments 1 and 2) or another participant made (in Treatment 3) and the recommendations of the two advisors in Stage 1 (see the online Appendix for details). On the basis of this information, participants are asked to select an advisor, who will answer the *same* questionnaire on their behalf, allowing them to earn money for every correct answer of the selected advisor. Stage 1 is common in all three treatments, while experimental manipulations are introduced in Stage 2 as follows:

- **Treatment 1 (Baseline.)** In Stage 2 of the baseline participants receive a table, which shows their Stage-1 answers to each question and the corresponding recommendations of the two advisors. Participants *do not* receive any feedback on how many correct answers they gave in Stage 1.

- **Treatment 2 (Addressing Overconfidence.)** In Stage 2 participants view the same type of table as in the baseline. In addition, they are shown the number of questions they answered correctly in Stage 1.

- **Treatment 3 (Addressing Ego-involvement.)** In Stage 2 participants are given the same information as in Treatment 2, but in relation to another person. That is, each participant is shown the number of correct answers, the answers to the questionnaire and the corresponding recommendations of the two advisors referring to another participant who went through Stage 1 previously. The participant's task is still to choose one of the advisors pertaining to the given table to answer the questionnaire on the participant's (not the Decision-Maker's) behalf and the choice is incentivized.[5]

[4] In all three treatments, participants are fully informed about the process by which the low-accuracy advisor choose their answers and of the percentage of participants that gave the most popular answer (for each question) in the pilot study that determined the Populist's answers.

[5] To avoid confusion, participants are explicitly told that the order of questions in the table is different from their own from Stage 1. So, when participants choose advisor at Stage 2 of Treatment 3, they do not know the advisors' responses to individual questions. This treatment is designed to counter ego-involvement, since someone may be reluctant to select an advisor with different answers from her own, because that would contradict her original choices from Stage 1.

2.1 Theoretical Predictions

In this section we present theoretical predictions and testable hypotheses for our experimental treatments. A simple Bayesian model captures the decision process of a rational participant, and, as we shall see, its predictions are greatly simplified for Treatments 2 and 3. Suppose that there are three agents: the participant P, advisor A, and advisor B.[6] $\mathcal{G} = \{P, A, B\}$ denotes the set of Players. There is a set $\mathcal{Q}$ of questions, with cardinality Q. Each question q has two candidate answers and only one is correct. Let μ_q be the participant's prior of their own answer on question $q \in \mathcal{Q}$ being correct.

By construction, the Expert always provides the correct answer to each question. The Populist gives the same answer as the participant in question q with probability π_q, which is the fraction of laypeople who gave the most popular answer to question q in the pilot study with an identical sample frame. There are two states of the world. In state $s = 1$, A is the Expert and B is the Populist. In state $s = 2$ the reverse happens, namely A is the Populist and B is the Expert. For simplicity, both states are assumed to have an equal prior probability: $prob\ _0(s = 1) = prob\ _0(s = 2) = 1/2$, which is a natural assumption in the absence of other information.

The participant observes recommendations by the two advisors on each question. Formally, P observes $a = \{a_A, a_B\}$, with $a_A = \{a_{A1}, a_{A2}, ..., a_{AQ}\}$ and $a_B = \{a_{B1}, a_{B2}, ..., a_{BQ}\}$, where a_{iq} denotes the choice of player $i \in \mathcal{G}$ on question $q \in \mathcal{Q}$. Thus, by comparing the answers of the two advisors to her own, P constructs the set X of questions for which she has common answers with B and the set Y of questions with common answers with A. In notation, $X = \{q \in \mathcal{Q} | a_{Pq} = a_{Bq}\}$ and $Y = \{q \in \mathcal{Q} | a_{Pq} = a_{Aq}\}$. X^C and Y^C are the complements of X and Y, namely the sets of questions with non-common answers with B and A, respectively. The question is how a rational participant evaluates the posterior probabilities of states 1 and 2 given the sets X and Y and the priors $prob\ _0(s)$ and μ_q. The solution to this simple Bayesian problem is given by equations (1) and (2) below:

$$Prob(s = 1|X, Y) = \frac{prob\ _0(s = 1)}{prob\ _0(s = 1) + prob\ _0(s = 2) \times OR} \tag{1}$$

Where

[6] In the experiment we used the labels 'J' and 'M' to avoid priming the participants with ordering effects.

$$OR \equiv \frac{prob(X,Y|s=2)}{prob(X,Y|s=1)} = \frac{\prod_{q\in X}\mu_q \prod_{q\in X^C}(1-\mu_q)\prod_{q\in Y}\pi_q \prod_{q\in Y^C}(1-\pi_q)}{\prod_{q\in Y}\mu_q \prod_{q\in Y^C}(1-\mu_q)\prod_{q\in X}\pi_q \prod_{q\in X^C}(1-\pi_q)} \quad (2)$$

Expressions (1) and (2) can be precisely estimated with the use of our experimental data. In particular, participants' subjective prior beliefs μ_q on the correctness of each answer are elicited after participants answer each question. The probability π_q of having a common answer with the Populist in question q is derived from the pilot study: it is the fraction of participants in the pilot giving the most popular (i.e. modal) answer. Note that our participants in the main experimental treatments are informed of this fraction for each question. This allows us to construct a benchmark of rational beliefs regarding the advisor most likely to be the Expert, and whom a money-maximizing participant would select at the end of Stage 2.

Furthermore, note that for experimental Treatments 2 and 3 the computations for Expressions (1) and (2) are greatly simplified. This is because, in these two treatments, participants know the number of correct answers of the Decision-Maker. In addition, they know that the Expert is always correct. These two facts imply that *the Decision-Maker has as many common answers with the Expert as the Decision-Maker's number of correct answers.* In other words, the Expert is always the advisor who exhibits as many common answers with the Decision-Maker, as the latter's number of correct answers. Thus, as long as the number of common answers with the two advisors differs, a rational participant in Treatments 2 and 3 can identify the Expert with 100% accuracy.[7]

2.2 Our Research Hypotheses

We preregistered a series of research hypotheses at the depository of the Open Science Framework, using the OSF template.[8] There we described the preceding model and its predictions. We also specified the research design, hypotheses, sampling plan, variables, and statistical analysis plan. The main preregistered research hypotheses at the aggregate level were the following:

H1: In issues of economic policy, participants systematically select advisors with opinions similar to their own, even if they know that populist (non-expert) advisors strategically express similar opinions to participants. As a result, participants typically select populists. [Directional

[7] When the number of common answers is equal between the advisors, then a Bayesian decision maker still uses (1) and (2) to make an inference, but the posterior can be anywhere between 50% and 100%.

[8] The preregistration can be found in https://osf.io/jr92p.

hypothesis for Treatment 1: the percentage of participants choosing the Expert is lower than the percentage choosing the Populist].

H2: Feedback on the participants' performance in addressing economic policy issues diminishes the tendency to select the Populist (Overconfidence Hypothesis). [Directional hypothesis: the percentage of participants choosing the Expert in Treatment 1 is lower than the percentage choosing the Expert in Treatment 2].

H3: Eliminating the direct connection between advisors' opinion and participants' opinion on the same subject diminishes the tendency to select the Populist ('Ego-Involvement' Hypothesis). [Directional hypothesis: the percentage of participants choosing the Expert in Treatment 3 is higher than the percentage choosing the Expert in Treatments 1 and 2].

H4: Participants who lean on the right politically agree more often with economic experts. [Directional hypotheses: Participants with political preferences above the median in the left-to-right scale have more correct answers and select the Expert more frequently in all treatments].

2.3 Experimental Implementation

We run our experiments online, with the key experimental manipulation pertaining to the information provided in Stage 2 and using the real questionnaire of economic reasoning introduced in Alysandratos et al. (2020).[9] The correct answers (corresponding to the Expert's recommendation) were validated by the consensus of academic economists, as described in Alysandratos et al. (2020) and as reproduced in our online Appendix. All experiments took place in the Fall of 2021. Before the main experimental sessions, we run a pilot study where laypeople provided their answers to the questionnaire without any feedback or any recommendations. A sample of 120 participants, representative of the general population in England and Wales in terms of age and sex, were recruited for this pre-study via *Prolific*. The answers of the low-accuracy advisor (whom we term 'Populist' here) were the modal answers from this pilot, i.e. the answers which were selected by the highest number of participants. Four of these ten modal answers coincide with the correct answers, hence the 'Populist' has four correct answers.

[9] We diverged slightly from that questionnaire, in that we reduced the number of options to two, instead of four. This was done in order to simplify the theoretical model and get sharper predictions.

	Experiment
Sample size	600 for the main experiment 120 for the pre-study
Mode of administraton	Online
Recruitment platform	Prolific Representative of England and Wales in terms of age and sex
Expert's accuracy	100%
Populist's accuracy	40% Always the modal answer from the pre-study
Questionnaire	MCQ with 2 options + belief elicitation
Treatments	Treatment 1: Baseline Treatment 2: Addressing Overconfidence Treatment 3: Addressing Ego-involvement

Table 1: *Summary of the experiment.*

For the main experiment, recruitment was also conducted via Prolific with the same sample frame. Each treatment contained 200 participants, selected to be representative of the general population of England and Wales in terms of age and sex (screenshots with exact instructions can be found in the online Appendix). Furthermore, correct answers were incentivized in both stages. For every correct answer in Stage 1, participants earned £0.07. In Stage 2, if they selected the Expert they received £3.15 and if they selected the Populist they received £1.05. After the main part of each experimental session, participants answered a short questionnaire on demographics and socio-political views. A quick summary of the above information is presented in Table 1.

Table 2 gives a break-down of the three treatments across several important demographic variables. Apart from the average age of participants, the table illustrates the proportions of participants that 1) are UK nationals, 2) have income higher than £ 30,000, 3) are married or in a civil union, 4) are female and 5) have at least undergraduate university education. As we can see, the differences across treatments are relatively minor. This demographic information will be incorporated in the regression analysis (see section 3.4).

	UK National	High Income	Married	Female Sex	Attended Uni	Age
Treatment 1	0.920	0.465	0.390	0.510	0.650	44.3
Treatment 2	0.890	0.530	0.385	0.515	0.655	45.2
Treatment 3	0.875	0.475	0.395	0.505	0.590	45.5

Notes: Although absence of significant differences does not provide conclusive evidence of equivalence, we note that no test yielded statistically significant differences at the conventional levels. Proportion tests show no statistical differences between the participants in the three treatments with respect to nationality (p-value = 0.327), high income (p-value = 0.375), marital status (p-value = 0.979), sex (p-value = 0.980) and educational level (p-value = 0.325). A Kruskal-Wallis test finds no difference with respect to age across the three treatments (p-value = 0.667).

Table 2: *Summary of main demographics across treatments.*

3 Results

Starting with the participants' performance in the questionnaire, they found some questions more difficult than others. Table 3 describes the fraction of participants that answered correctly each of the ten questions at Stage 1. Question 3 was the most difficult one, since less than a third of participants gave the correct answer. On the other hand, Question 6 was the easiest one, with more than 80% of participants giving a correct answer. In general, participants in the three treatments answered similarly in Stage 1 for the ten questions. It also needs to be emphasized that the pilot study accurately predicted the most popular answers for each question. In particular, for every single question the most popular answer from the pilot (hence, the Populist's answer) was the answer provided by the majority at Stage 1 of the main experiment.

In terms of our main results, the experimental evidence supports our pre-registered hypotheses H1 and H2 at the aggregate level. Figure 1 shows the frequency of participants choosing the Expert and the Populist when aggregating participants across all three treatments. As we can see, less than 50% of participants choose the Expert, despite the fact that there are only two options, meaning that random choice would be successful on average half of the time. This result indicates that the Populist is highly successful in presenting themselves as the high-accuracy advisor by emulating the answers of the modal participant.

Figure 2 illustrates the main treatment effects. The main result of Figure 1 also holds true for each treatment individually: participants' selections perform worse (in finding the Expert)

Question	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Average correct answers
Treatment 1	0.415	0.440	0.205	0.780	0.795	0.825	0.310	0.33	0.475	0.670	5.24
Treatment 2	0.435	0.410	0.270	0.785	0.810	0.810	0.420	0.32	0.480	0.705	5.44
Treatment 3	0.375	0.345	0.210	0.840	0.795	0.835	0.435	0.38	0.485	0.780	5.48

Notes: We employ Chi-Square tests for each question to look for differences in the proportion of correct answers across treatments. We find statistically significant differences only for question 7 (p-value = 0.02) and question 10 (p-value = 0.04). Using a Kruskal-Wallis rank sum test for the total number of correct answers across treatments, we find no statistically significant differences (p-value = 0.2681)

Table 3: *Participants' accuracy in the economic questionnaire across the three treatments.*

than random choice. This consistency of the 'success' of the Populist's strategy across the three treatments is important, given that we are examining multiple hypotheses with discrete samples. Hypothesis H1, which concerns the Populist's success, seems to be strongly and consistently borne by the data. The proportion of those choosing the Expert in the baseline treatment is 32.5%, significantly lower than 50% (as per H1) according to Z-test or a Chi-Square test ($p < 0.0001$). The propensity to select the Expert increases in Treatment 2 to 42%, an increase which is significant (as per H2) according to a Chi-Square test ($p = 0.04942$). On the other hand, contrary to H3, Treatment 3 is unsuccessful in further increasing this propensity. However, as we explain is section 3.4, this is due to a composition effect. For those participants who face the easy task of selecting the advisor they agree with the most, countering ego-involvement does not improve the quality of their decision making. For those participants who face the cognitive harder task of selecting the advisor with whom they disagree the most, countering ego-involvement significantly improves their performance. In comparison to other treatments, Treatment 3 contains more participants in the first category and less in the second.

3.1 The role of beliefs

A feature of our design is that we elicit participants' level of confidence when answering the questionnaire in Stage 1. Recall that participants were asked to give a number ranging from 0 to 100 for each question they answered. Evidence on participants' confidence is essential for having a theoretically-valid benchmark of rational decision-making, but it may also yield useful insights on

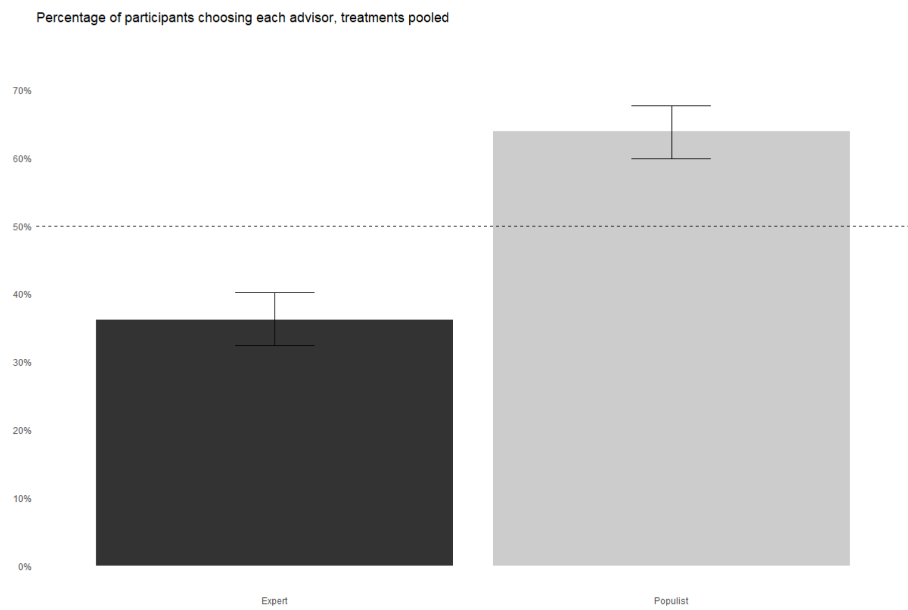

Figure 1: *Percentage of participants choosing the Expert and the Populist over the entire sample of participants. The error bars show the conventional 95% confidence intervals, and the horizontal line at 50% denotes random choice.*

the role of priors in our setting. While the elicitation of participants' confidence was not incentivized, the evidence indicates that responses are non-random as we shall explain below.[10]

Figure 3 shows the average reported confidence in participants' answers for each question across the three treatments, while Figure 4 presents the density distribution for the values of elicited confidence. Three observations emerge from these graphs. First, there is substantial variation in the average confidence across questions. If participants were reporting randomly or without paying attention, we would not expect such variation. Secondly, there is high dispersion of exhibited confidence within each question, meaning that participants display varying degrees of confidence in their answers. Such heterogeneous degrees of confidence indicate that overconfidence may be a valid concern, which we address experimentally in our Treatments 2 and 3. Thirdly, the average elicited confidence in a particular question often diverges significantly from participants' propensity of a correct answer, as reported in Table 3. Hence, participants do not have well-calibrated beliefs on each question. Average confidence is a good proxy for the propensity to answer correctly only for a few questions.

In particular, only in questions 4,5 and 10 is the average confidence comparable to the fraction of participants with correct answers. In questions 1, 2, 3, 7, 8 and 9, average confidence is 70% or

[10]This is in accordance with Trautmann and van de Kuilen (2015), who show that incentives for belief elicitation seem to play little role in improving accuracy.

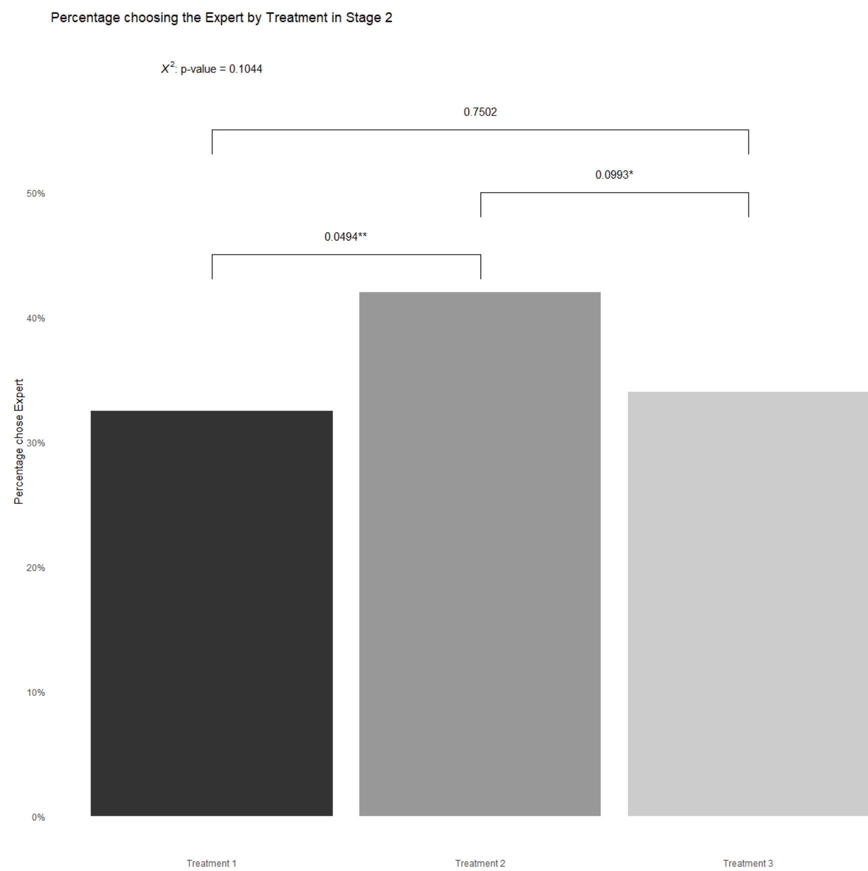

Figure 2: *Main treatment effects: percentage of participants choosing the correct advisor in Stage 2, by treatment. Horizontal lines show the p-values for chi-square tests for the differences between the treatments.*

above, while average accuracy is below 50% in all cases. Question 6 exhibits the opposite pattern, with 80% of participants answering correctly, but with average confidence of 70%. On the other end of the spectrum is question 3, where average confidence is over 80% in all treatments, whereas accuracy is below 30%.

Indeed, digging deeper into the data reveals that the mismatch between true accuracy and beliefs does not come from a minority of individuals, but it is widespread across participants. For each question, we counted the number of people who gave the wrong answer, while reporting confidence 70% or higher. Table 4 shows this count per question and treatment, along with percentages. While in question 6 this categorization captures only 33 participants (5.5% of all the total), in question 3 this increases to 401 participants (66.83% of the total). In questions 1, 2, 7 and 8, over 30% of total participants fall in this categorization, which indicates that overconfidence is a wide-spread occurrence in our sample. Percentages falling in this category are of course higher when calculated as a fraction of participants who provided wrong answers. For example, over 60% of participants

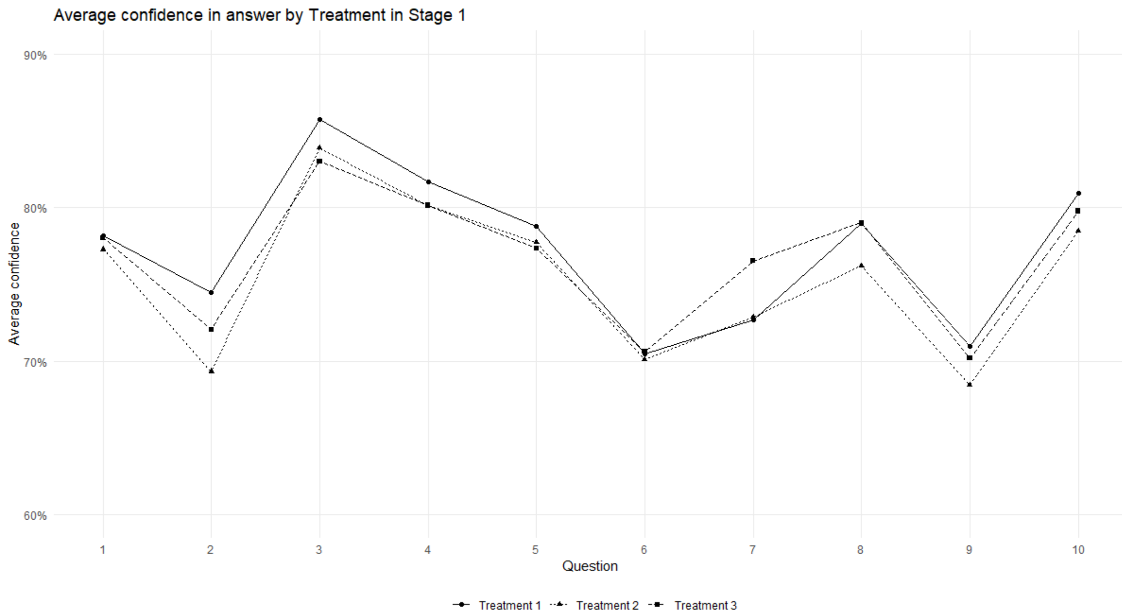

Figure 3: *Average confidence in participant answers per question across treatments.*

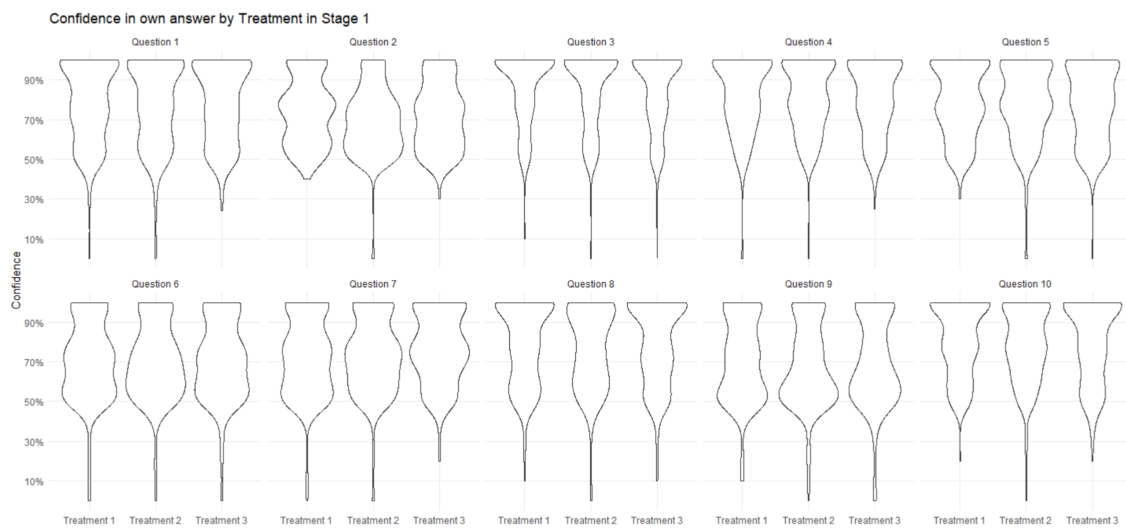

Figure 4: *Distribution of reported confidence per question across treatments.*

who gave the wrong answer in question 8, reported confidence above 70%. The results of Table 4 are consistent across treatments, and highlight systematic differences across questions.

A different potential measure of the disparity between beliefs and actual accuracy is the ratio of the average confidence in questions a participant answered wrongly, over the average confidence in questions where the same participant answered correctly. If this ratio is above (below) one, then the participant exhibits *badly (well) calibrated beliefs*. We can then count the fraction of participants with badly calibrated beliefs across the three treatments. We found that 94 (47.00%), 93 (46.50%) and 90 participants (45.00% of participants) had badly calibrated beliefs in Treatments 1, 2 and 3 respectively. Overall, 46.17% of all participants exhibited higher average confidence for the questions they answered wrongly. It seems that most of our participants are not well calibrated in economic matters.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Count										
Treatment 1	95	68	142	30	24	12	81	85	55	46
Treatment 2	85	59	124	22	17	10	61	83	52	41
Treatment 3	90	77	135	22	21	11	74	84	61	30
Total	270	204	401	74	62	33	216	252	168	117
Percentage of all participants										
Treatment 1	47.50%	34.00%	71.00%	15.00%	12.00%	6.00%	40.50%	42.50%	27.50%	23.00%
Treatment 2	42.50%	29.50%	62.00%	11.00%	8.50%	5.00%	30.50%	41.50%	26.00%	20.50%
Treatment 3	45.00%	38.50%	67.50%	11.00%	10.50%	5.50%	37.00%	42.00%	30.50%	15.00%
Total	45.00%	34.00%	66.83%	12.33%	10.33%	5.50%	36.00%	42.00%	28.00%	19.50%
Percentage of wrong participants										
Treatment 1	81.20%	60.71%	89.31%	68.18%	58.54%	34.29%	58.70%	63.43%	52.38%	69.70%
Treatment 2	75.22%	50.00%	84.93%	51.16%	44.74%	26.32%	52.59%	61.03%	50.00%	69.49%
Treatment 3	72.00%	58.78%	85.44%	68.75%	51.22%	33.33%	65.49%	67.74%	59.22%	68.18%
Total	76.06%	56.51%	86.61%	62.18%	51.67%	31.13%	58.86%	63.96%	53.85%	69.23%

Table 4: *Count and percentage of participants, in a given question, who answered wrongly and reported confidence in their answers of at least 70%.*

3.2 Accounting for Participants' Behavior

We first need to address the empirical performance of our Bayesian model. Participants know that if the state of the world was actually $s = 2$, then Advisor A would be the Populist, whereas Advisor

B would be the Expert. On the other hand, if the state of the world was actually $s = 1$, Advisor B would be the Populist, whereas Advisor A would be the Expert. Without loss of generality, assume that the true state of the world is $s = 1$. So, if a participant believes that the state is $s = 1$, then she infers correctly the identities of the advisors, while if she believes that the state is $s = 2$, then she infers them incorrectly. Using as inputs the participants' actual answers to questions 1-10, their reported confidence in their answers, and the recommendations of the Expert and of the Populist, we can estimate each participant's posterior and the Odds Ratio using Equations 1 and 2.

Since the natural prior is 50%, a simple result of this model is that participants with odds ratios less than one should infer that advisor A is the Expert and they should choose the Expert in Stage 2 if they are maximizing experimental earnings. On the other hand, participants with odds ratios greater than one should infer that advisor B (the actual Populist) is the Expert and should be chosen in Stage 2. By juxtaposing predicted with actual behavior in Treatment 1, we find that 69% of participants choose advisor according to the predictions of the Bayesian model. In Treatments 2 and 3, the model predicts that participants choose the Expert whenever the amounts of common answers of the Decision-Maker with the two advisors differ. By restricting analysis to cases that satisfy this condition, we find that only 45.6% and 38.4% of the relevant participants in Treatments 2 and 3, respectively, choose according to the model's prediction. The model achieves moderate predictive success in Treatment 1 and poor performance in Treatments 2 and 3.

This is an interesting result, because Treatment 1 does not provide feedback on performance and so the Bayesian model is mainly guided by the precision of priors. In our experiment, participants have badly calibrated priors (they are too confident on wrong answers and unsure on correct ones) and, as a result, the model is relatively accurate in predicting participants' choices but very inaccurate in identifying the Expert. Out of 200 observations, the model selects the Expert in only 57 cases, a success ratio of only 28.5%. However, the extra information provided by the feedback in Treatments 2 and 3 is sufficient for the Bayesian model to become 100% accurate whenever the number of common responses with the participant differ across the two advisors. Nonetheless, less than half of participants in these treatments select the Expert even when restricting attention to cases with different numbers of common answers across advisors. As mentioned earlier, only 45.6% of participants in the relevant cases of Treatment 2 and 38.4% of participants in the relevant cases of Treatment 3 answer correctly in Stage 2, whereas the model predicts 100% success in these cases. Overall, the Bayesian model does not provide a good fit for participants' choices.

This begs the question as to how most participants select advisor. An alternative account of behavior can be made using the natural human tendency to consider as an expert the advisor with whom we agree on a given matter. By *'simple heuristic'* we denote the behavioral rule of choosing, in Stage 2, the advisor with whom the Decision-Maker has the most common answers. For each Treatment, we determine which advisor would be chosen by a participant following this simple heuristic, comparing the Decision-Maker's answers with those of each advisor at Stage 2. In our data, excluding ties, about 84.9% of participants behaved according to the simple heuristic in Treatment 1, 62.6% in Treatment 2 and 68% in Treatment 3.[11] This means that an overwhelming majority of participants chose according to this basic rule, but the tendency to do so fell in Treatments 2 and 3, where it was possible to deduct logically that this rule did not result in the optimal choice.

A key behavioral hypothesis is that the simple heuristic constitutes a powerful driving force of behavior, which may often overrule the effect of logical reasoning. For participants who, in Stage 2, observe Decision-Makers with more common answers with the Expert than with the Populist, the optimal Bayesian behavior often coincides (in Treatments 2 and 3 it always coincides) with the prescriptions of the simple heuristic. Accordingly, to examine the simple heuristic, we focus only on the behavior of participants who observe Decision-Makers with strictly more common answers with the Populist than the Expert. Not only do these participants constitute the majority of our observations, but they also face the hard task of having to choose against the prescription of the simple and intuitive heuristic. It is therefore interesting to check how many of them manage to do so.

Figure 5 illustrates the relevant numbers. In Treatments 1, 2 and 3, respectively, there are 122 participants, 113 participants and 153 participants who observed Decision-Makers with more common answers with the Populist than with the Expert in Stage 2. The figure shows that while only 17.2% of such people overcome the tendency to choose according to the simple heuristic in Treatment 1, much higher percentages (38.9% and 33.3%) manage to do so in Treatments 2 and 3. These two percentages are significantly higher than the analogous percentage in Treatment 1 (Chi-Square test with continuity correction, $p < 0.001$ in either case). This can be interpreted as follows: first of all, even though there is a clear optimal choice in Treatments 2 and 3, more than half of participants for whom this optimal choice clashes with the prescription of the simple heuristic choose according to the latter. On the other hand, the fact that it is logically feasible and relatively easy to deduce who is the Expert in Treatments 2 and 3 (as opposed to Treatment

[11] This is even though, in Treatment 3, the 'common answers' were between an advisor and a third person.

1, where prior beliefs play a role) considerably reduced the popularity of the simple heuristic.

Overall, this descriptive evidence indicates that even if the structure of the logical problem is clear, and the modus operandi of the Populist is transparent, the Populist can still gain public support by catering to people's priors. There are many open questions, but this initial evidence on the 'identify the Expert' problem is strong, and there are valid reasons to expect that it is generalizable (see also discussion below).

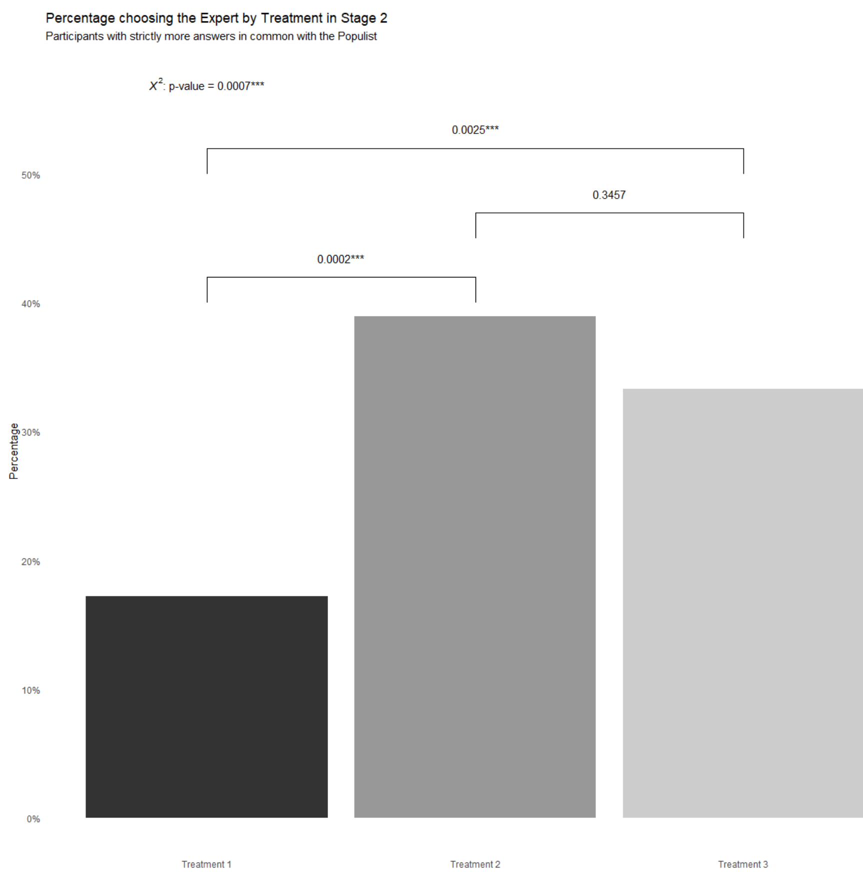

Figure 5: *Choice of Expert against the simple heuristic*

3.3 Regression analysis of preregistered hypotheses

In this section we present regression analysis to delve more deeply in potential drivers of behavior. On the basis of prior evidence, we hypothesized in our preregistration that being male and taking high-level courses in technical disciplines and economics are associated with high performance at Stage 1. The examination of the effects of the remaining variables has an exploratory nature.

We started with a baseline model with only treatment variables and some selected demographics (age, gender, nationality). To this model we added subsequently the remaining demographic variables separately (income, marital status, political leaning, etc), as well as all together in a full model. All regressions use OLS and the standard errors are robust using the HC3 version of the heteroscedasticity-consistent covariance matrix (Long and Ervin, 2000).

Table 5 presents the baseline regression models on the number of questions correctly answered in Stage 1. The treatment variables (i.e. dummy variables corresponding to the treatments of our experiment) are not statistically significant in any of the models in Table 5, indicating that there are no differences in the number of total correct answers among the three treatments. This is reassuring, as it indicates that participants have similar familiarity with economics across the treatments and so our interventions are meaningful. The variable 'Male' is statistically significant in all specifications, suggesting that males achieved an additional half correct answer compared to females. In Table 9 in the appendix we present the full set of control variables. We find no significant effects for the participants' marital status, income level, political leaning, their attention to the experiment (measured as the time spent on answering), the discipline of study, or their occupational sector. Participants with a postgraduate degree give about 0.5 more correct answers. We find no statistically significant differences for other levels of educational attainment. EU nationality has a weakly significant and positive effect in column 7, where we control in addition for the discipline of study of our participants. Finally, those with self-professed low knowledge on matters of economic policy (see column 4 in Table 9 in the Appendix) score about 0.4 fewer correct answers. It is worth noting that only 23.7% of our participants declare themselves not very knowledgeable on these topics.

In Table 6 we present the baseline regression models on the probability of choosing the Expert in Stage 2, which is the main exogenous variable of interest. We observe that Treatment 2 has a rather weak yet positive statistically significant effect (at the 10% confidence level) on correctly identifying the Expert, implying a marginal effect of approximately 1 percentage point. However, this observation misses the true mechanism at play. Some participants are faced with an easy task (selecting the advisor with whom they agree with the most) and perform well across all treatments. Other participants undertake the harder task of selecting the advisor they disagree with. For the latter participants, the treatment has a meaningful effect. Indeed, as we show in section 3.4, when we run the same set of regressions with additional interaction terms for task difficulty and treatment, both treatment dummies become significant at the 1% level. Thus, it is the mix between these distinct groups which reduces the marginal effect of treatment 2. The magnitude of the coefficient is also consistent with the difference between Treatment 1 and 2 of Figure 2.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment: No Overconfidence	0.195	0.197	0.181	0.199	0.192	0.191	0.213	0.192	0.184
	(0.162)	(0.163)	(0.163)	(0.163)	(0.163)	(0.163)	(0.164)	(0.164)	(0.169)
Treatment: No Ego-involvement	0.220	0.221	0.219	0.227	0.223	0.230	0.265^{*}	0.221	0.257
	(0.155)	(0.155)	(0.155)	(0.155)	(0.155)	(0.155)	(0.156)	(0.158)	(0.161)
Age	0.001	0.000	0.002	0.001	0.001	0.002	0.002	0.002	0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
Sex: Male	0.523^{***}	0.525^{***}	0.508^{***}	0.473^{***}	0.524^{***}	0.514^{***}	0.504^{***}	0.515^{***}	0.453^{***}
	(0.132)	(0.132)	(0.133)	(0.135)	(0.132)	(0.132)	(0.138)	(0.140)	(0.151)
Sex: Other	1.633^{*}	1.647^{**}	1.654^{*}	1.446^{*}	1.622^{*}	1.578^{*}	1.550^{*}	1.635^{*}	1.371
	(0.847)	(0.833)	(0.870)	(0.868)	(0.836)	(0.806)	(0.831)	(0.869)	(0.902)
Nationality: EU	0.442	0.449	0.445	0.438	0.426	0.383	0.459^{*}	0.434	0.384
	(0.272)	(0.273)	(0.274)	(0.272)	(0.273)	(0.278)	(0.275)	(0.269)	(0.281)
Nationality: Other	−0.289	−0.292	−0.278	−0.259	−0.306	−0.395	−0.295	−0.329	−0.392
	(0.373)	(0.377)	(0.374)	(0.383)	(0.371)	(0.383)	(0.380)	(0.384)	(0.410)
Constant	4.903^{***}	4.935^{***}	4.772^{***}	5.201^{***}	4.823^{***}	4.598^{***}	4.904^{***}	4.895^{***}	4.897^{***}
	(0.208)	(0.215)	(0.241)	(0.352)	(0.236)	(0.287)	(0.229)	(0.232)	(0.502)
Additional controls	–	Marital status	Income	Political leaning	Attention	Educational level	Discipline studied	Occupational sector	All included
R^2	0.040	0.042	0.043	0.053	0.041	0.051	0.047	0.048	0.077
Adjusted R^2	0.029	0.027	0.027	0.037	0.028	0.033	0.030	0.028	0.030
Observations	600	600	600	600	600	600	600	600	600

Notes: The dependent variable is the total number of correct answers in Stage 1. Robust standard errors in brackets. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Table 5: *Regressions on the number of correct answers in Stage 1*

Correspondingly, Treatment 3 has no statistically significant effect in any specification. Although the coefficient is consistently positive, the implied marginal effect is between 1 and 2 percentage points. Hence, prima facie, we cannot reject the null hypothesis that Treatment 3 has no effect. As mentioned above, this picture changes with the introduction of interaction terms between task difficulty and treatment (see section 3.4 below). In terms of the other control variables, gender does not predict the choice of the Expert in Stage 2. Being an EU citizen is weakly significant in all but one of our specifications. In Table 10 in the appendix we present the full set of results. Low self-professed knowledge on topics of economic policy has a weak, negative effect on the probability of choosing the Expert in the full model. We find no evidence that any of our other controls have an effect in the choice of the Expert. These findings go against our pre-registered hypothesis H4, regarding participants who lean on the right politically.

The online Appendix contains additional robustness checks on these regression results. In our preregistration document we have noted that in expressing their confidence in the answer they have chosen, rational participants should assign probability greater than 50% to their chosen answer. Our additional checks exclude participants who fail to assign probabilities consistent with their

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment: No Overconfidence	0.092*	0.093*	0.093*	0.096*	0.091*	0.092*	0.092*	0.093*	0.094*
	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(0.050)
Treatment: No Ego-involvement	0.009	0.009	0.009	0.013	0.010	0.009	0.013	0.010	0.016
	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)	(0.049)	(0.048)	(0.050)
Age	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	0.000	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Sex: Male	0.023	0.025	0.025	0.020	0.024	0.026	0.023	0.026	0.031
	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)	(0.041)	(0.041)	(0.044)
Sex: Other	0.417	0.417	0.422	0.349	0.414	0.419	0.409	0.405	0.350
	(0.284)	(0.286)	(0.289)	(0.251)	(0.285)	(0.280)	(0.295)	(0.319)	(0.297)
Nationality: EU	0.143*	0.141*	0.138	0.156*	0.139*	0.139*	0.144*	0.149*	0.153*
	(0.083)	(0.084)	(0.084)	(0.083)	(0.083)	(0.084)	(0.084)	(0.084)	(0.088)
Nationality: Other	0.066	0.062	0.063	0.083	0.062	0.067	0.066	0.061	0.077
	(0.118)	(0.119)	(0.121)	(0.120)	(0.118)	(0.120)	(0.121)	(0.125)	(0.136)
Constant	0.333***	0.330***	0.341***	0.523***	0.315***	0.339***	0.362***	0.276***	0.521***
	(0.067)	(0.070)	(0.075)	(0.110)	(0.074)	(0.091)	(0.075)	(0.075)	(0.154)
Additional controls	–	Marital status	Income	Political leaning	Attention	Educational level	Discipline studied	Occupational sector	All included
R^2	0.020	0.021	0.023	0.034	0.021	0.022	0.023	0.025	0.047
Adjusted R^2	0.009	0.006	0.006	0.018	0.008	0.004	0.005	0.005	−0.001
Observations	600	600	600	600	600	600	600	600	600

Notes: The dependent variable is a dummy indicating whether the participant chose the Expert in Stage 2. Robust standard errors in brackets. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Table 6: *Regressions on choosing the Expert in Stage 2*

choices (we chose to exclude those with two or more inconsistencies to allow minimal flexibility for random errors). Another issue with online experiments is possible lack of attention. Participants face a trade-off between better performance, and hence higher payment, from a single study, and participating in multiple studies. We thus also present robustness checks where we drop participants in the lower quartile and the lower half with respect to the variable 'Attention' in our sample. None of our main results are impacted by this sample restriction. In addition, the variables 'Male' and 'Econ Knowledge: Low' retain explanatory power in accounting for the number of accurate answers to the questionnaire in Stage 1. On the other hand, when it comes to choosing the Expert in Stage 2, the 'EU Nationality' variable is no longer significant for many of the specifications of the robustness checks.

3.4 Exploratory regression analysis

As we have mentioned previously, on average, treatments 2 and 3 have weak, if at all any, impact on the ability of participants to discriminate the Expert from the Populist. However, there is a difference in the task difficulty between those who have many common answers with the Expert and those who have many common answers with the Populist. The first group needs to select someone they agree with while the latter someone they disagree with. We conjecture, and in this set of exploratory regressions, we test statistically if this has any impact on the treatment effects. That is, if treatments 2 and 3 differentially impact participants with ‘easy’ tasks vis-á-vis participants with ‘hard’ tasks.

To test this formally, we restrict our dataset to participants who have different number of common answers with the two advisors. Subsequently, we construct the dummy ‘Easy task’, which takes the value one if the Decision-Maker has more common answers with the Expert than the Populist. We include this dummy, along with interaction variables between it and the two treatments in the same set of regressions as those reported in table 6. The summary of the results are reproduced in table 7 while the full set of regressions appears on table 11 in the Appendix. Column 10 includes some additional controls, namely the number of a participant’s correct answers in Stage 1, in linear (‘Total own’) and in quadratic form (‘Total own2’) to account for potential non-linearities, along with the average confidence of a participant in their answers in Stage 1 (‘Average confidence’).

The results are telling. In all ten specifications, the two treatment effects, the task-difficulty dummy, and the interaction terms are all highly statistically significant. Note that the coefficients are positive for treatment and the task-difficulty dummy, but they are negative for the interaction terms. Thus, it seems that, while participants with the easy task perform substantially better than those with the hard task on average, the performance gap vanishes for treatments 2 and 3. In other words, addressing overconfidence and ego-involvement helps subjects with the hard task to perform better but hurts subjects with the easy task. This explains why the average treatment effects in table 6 are not significant.

Moreover, from table 7 we observe that the magnitude of Treatment 2’s effect (approximately 0.2) is not only statistically significant but also economically important. On average, a participant who is faced with the hard task and who receives feedback on their number of correct answers, has about 20 percentage points higher chance to make the right selection of advisor. Treatment 3 has a similar effect but a lower magnitude of approximately 15 percentage points. Hence, our findings are consistent with the view that the inability to admit one’s own mistakes (Eskreis-Winkler and Fishbach, 2019), for example because of self-image or ego-threatening concerns (Falk and Zimmermann, 2017; Eskreis-Winkler and Fishbach, 2022), is one of the obstacles to choosing the Expert. Nevertheless, despite the improvement achieved by Treatments 2 and 3, there is still a large margin for improvement, since most participants did not choose the Expert even in these treatments. This calls for an investigation of other mechanisms that may be at play.

Before proceeding, note that task difficulty is endogenous in our experiments, since the problem

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment: Overconfidence	0.214***	0.213***	0.217***	0.217***	0.215***	0.213***	0.219***	0.209***	0.218***	0.198***
	(0.058)	(0.058)	(0.059)	(0.058)	(0.058)	(0.059)	(0.059)	(0.058)	(0.062)	(0.062)
Treatment: Ego-involvement	0.156***	0.156***	0.156***	0.158***	0.160***	0.156***	0.170***	0.153***	0.170***	0.154***
	(0.052)	(0.052)	(0.052)	(0.052)	(0.052)	(0.052)	(0.053)	(0.052)	(0.054)	(0.056)
Easy task	0.781***	0.783***	0.782***	0.783***	0.784***	0.779***	0.794***	0.781***	0.799***	0.616***
	(0.056)	(0.058)	(0.057)	(0.057)	(0.056)	(0.056)	(0.057)	(0.058)	(0.064)	(0.096)
Treatment: Overconfidence * Easy task	−0.510***	−0.509***	−0.507***	−0.509***	−0.512***	−0.508***	−0.524***	−0.521***	−0.536***	−0.553***
	(0.109)	(0.110)	(0.109)	(0.110)	(0.108)	(0.110)	(0.109)	(0.109)	(0.113)	(0.112)
Treatment: Ego-involvement * Easy task	−0.360***	−0.361***	−0.366***	−0.365***	−0.361***	−0.356***	−0.394***	−0.373***	−0.416***	−0.509***
	(0.123)	(0.125)	(0.126)	(0.122)	(0.122)	(0.125)	(0.122)	(0.125)	(0.128)	(0.136)
Age	−0.001	−0.000	−0.001	−0.000	−0.001	−0.001	−0.001	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Sex: Male	0.001	−0.000	0.004	0.010	0.001	0.002	−0.011	0.007	0.006	0.012
	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.044)	(0.046)	(0.047)
Sex: Other	0.430***	0.428**	0.437***	0.352**	0.423**	0.424***	0.427***	0.460**	0.378**	0.412**
	(0.165)	(0.170)	(0.162)	(0.160)	(0.175)	(0.162)	(0.147)	(0.180)	(0.166)	(0.165)
Total own										−0.145**
										(0.069)
Total own2										0.017**
										(0.007)
Average confidence										−0.003
										(0.002)
Constant	0.191***	0.192**	0.218***	0.350***	0.145*	0.183*	0.209**	0.111	0.270*	0.720***
	(0.072)	(0.077)	(0.078)	(0.115)	(0.081)	(0.096)	(0.082)	(0.078)	(0.162)	(0.252)
Additional controls	Nationality	Marital status	Income	Political leaning	Attentions	Educational level	Discipline studied	Occupational sector	All included	All included
R^2	0.193	0.193	0.197	0.199	0.196	0.193	0.204	0.203	0.222	0.236
Adj. R^2	0.175	0.172	0.174	0.179	0.176	0.170	0.179	0.176	0.170	0.179
Num. obs.	465	465	465	465	465	465	465	465	465	465
RMSE	0.441	0.442	0.442	0.440	0.441	0.443	0.440	0.441	0.443	0.440

Notes: The dependent variable is a dummy indicating whether the participant chose the Expert in Stage 2. Participants who have an equal number of common answers with the two advisers are excluded. Robust standard errors in brackets. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Table 7: *Exploratory regressions on choosing the Expert in Stage 2.*

participants face in Stage 2 is linked to the number of correct answers they provided in Stage 1. This holds true even for Treatment 3. Thus, we refrain from making any causal inferences. The impact of the task-difficulty on subject performance could be due to the difficulty of the task itself or due to unobserved individual characteristics. Even so, our interpretation of the treatment effects on performance (after controlling for task difficulty) stands. Treatments have differential impact on those faced with the easy task and those faced with the hard task, regardless of the underlying reason that causes individuals to select task difficulty. The section below delves deeper on the issue of task endogeneity.

3.5 Heterogeneity in the decision task and robustness

As explained earlier, identifying the Expert is not an easy task, but the task is particularly hard for participants who have few correct answers, hence they are likely to have more answers in common with the Populist. In this sense, there is heterogeneity in the type of problem that participants are called to solve. For participants who answer many questions correctly, finding the Expert is intuitive: if they simply select the advisor they agree with, they are very likely to pick the Expert. For less knowledgeable participants the challenge is far greater. First, they have to realize (Treatment 1) or accept (Treatment 2) that they are not knowledgeable. Second, they have to admit that their lack of knowledge means they cannot evaluate the advisors' answers properly. In fact, they have to choose the one they *disagree* with because this gives them higher chances of selecting the Expert. In Figure 6 we show the percentage of participants that find the Expert per number of correct answers they gave in the questionnaire.

The effect of the number of correct answers is non-linear. For less than 6 correct answers, between 10% and 50% of the participants identify the Expert (the case of one correct answer is the exception, but there are only 5 participants in this group). For 8 correct choices and above, expert identification improves dramatically, reaching 80% to 100%.

Why is the effect so strong in the neighborhood of 6 correct answers? There are two forces at play. First, the task is becoming easier, because participants intuitively prefer the correct advisor. Second, differences in underlying traits may play a role. For example, performing well in the questionnaire may correlate with some measures of intellectual skills. The question is raised whether very skillful participants would do well if they were faced with the decision problems of less skillful individuals. In this case, they would have to detect the heuristic of choosing the advisor they disagree with. Since the number of correct answers is by design endogenous in these experiments, we cannot answer the question using our data and new experiments would be needed.

Indeed, in a follow-up study (Alysandratos et al., 2023) we assign participants to the Stage 2 task of Treatment 3 of the present paper while varying the number of common answers between the layperson and the Expert and the layperson and the Populist randomly. We find that socio-economic demographics, education background and age in particular, have a strong predictive ability for the performance in the task. Thus, the likely mechanism behind our results is that cognitive abilities, as shaped by innate abilities, training and experiences, lead to the correct selection of the appropriate heuristic as the information changes. Note, however, that in the present paper we have consciously opted for the more ecologically relevant environment where individuals endogenously assign themselves to tasks of varying difficulty. In reality, individuals need to select between advisors of different quality without knowing who has better accuracy and using only their own opinions as signals. Thus, the design of this paper both allows us to define 'the identify the Expert' problem in comparison to other information structures and to document the cognitive challenge for most laypeople.

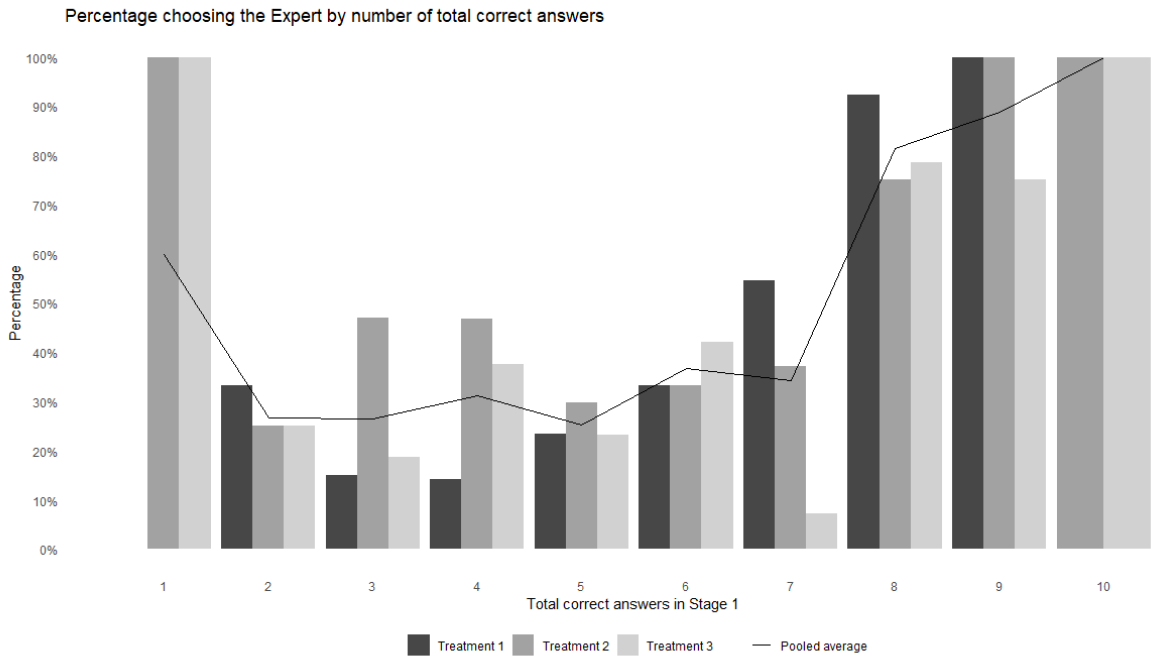

Figure 6: *Percentage of participants choosing the Expert and the Populist by treatment. The coloured bars show the percentage for each treatment and the line shows the average when all three treatments are pooled together.*

3.6 Identifying the mechanism

On the basis of all the above results, there is a wide range of channels that we can preclude as driving them. We summarize and discuss them below.

Reputation and moral hazard. Laypeople may knowingly ignore experts' recommendations because they do not trust their incentives to provide impartial advice. Indeed, the divergent objectives between the two groups is a standard feature of most models. Our design abstracts away from this problem by providing the exact modus operandi of both the Expert and the Populist to participants. In fact, in our setting, money-maximising participants strictly prefer to follow the Expert, since they are rewarded only on the basis of decision accuracy and the Expert has no way of deceiving them.

Imperfect signals of expertise. Laypeople frequently use cues or personal information as indications of ability. Dressing style, use of language, family background, CVs and personal websites, are all used to gauge competence or knowledgeability. Our design abstracts away from this type of complication by restricting attention to the content of advice alone. In addition, the only signal participants can use in Treatments 1 and 2 to reach a conclusion is their own responses to the questionnaire. In Treatment 3, the task is simplified further. They can only use the similarity of the answers between the third participant and the two advisors to infer who the Expert is. Thus, confusing the interpretation of signals is impossible in our experiments.

Overconfidence and beliefs. Another common concern in the literature is that laypeople overestimate their ability to answer correctly and so they may deem themselves as knowledgeable in an area they are not. In this case, people may follow the advice of non-experts simply because they think their own opinions must be closely correlated to experts' opinions. Indeed, we find evidence of this channel for Treatment 1. By measuring their beliefs in their responses, not only do we find that participants overestimate the number of correct answers they have, but they also have badly calibrated beliefs. That is, they are overly confident of the answers they got wrong and feel uncertain about the answers they actually have correct. This is evidence of the strong counter-intuitive nature of many economic findings. Thus, Treatments 2 and 3 are used to shut down this channel by giving participants direct feedback on the exact number of correct answers.

Confirmation bias and motivated reasoning. Participants in our experiments may have a natural tendency to stick to their priors, especially since they answer for themselves the questionnaire in Stage 1. We have acknowledged this possibility from the outset and Treatment 3 was explicitly designed to counter it. Note that the results from the regression analysis in section 3.4 validate the concern for ego-involvement. Indeed, participants who disagree with the Expert perform significantly better in Treatment 3 than in the other treatments. However, a larger fraction of participants fails to identify the Expert. Therefore, confirmation bias is not the major driving force behind our results.

Mathematical complexity and failures of Bayesian reasoning. It is well-established that most laypeople have difficulty dealing with complex mathematical problems and the failures of Bayesian reasoning are well-documented. This could be an issue in Treatment 1, where the correct inference requires complex computations via equations (1) and (2). However, the provision to the participant of the exact number of correct answers of the Decision-Maker in the questionnaire of Stage 1, along with the information that the Expert is always correct, renders the problem mathematically trivial. In fact no computation is required at all to identify the Expert in Treatments 2 and 3. Merely the realization that the Expert must have as many common answers with the participant as the participant has correct answers in the questionnaire should suffice for this task. Yet, most participants fail to perform this simple reasoning. Our conclusion is that consulting a person one disagrees with is a deeply counter-intuitive idea.

Inattention. Finally, it could be that participants failed to pay attention to key information in the experiment and were systematically wrong because of it. Several experimental features and results reassure us that this is highly improbable. First, we provided a good compensation for participation in the experiment. Given the relatively short duration (slightly less than 20 minutes on average) and that participants could earn close to £4, the pro-rate reward for participation stands above the minimum living wage of £10. Second, we made sure that participants could not skim through instructions. After instructions were over, they had to go through a short questionnaire and provide answers to key experimental information, including each advisor's accuracy. Participants were not allowed to move to the main experiment until they provided the correct answers. They also spent approximately 2 minutes in Stage 2 (average across all treatments), which is the most crucial decision stage for us, or 10% of the total experiment time. This is good indication that they put effort in selecting the Expert. The experimental results also strongly indicate that participants paid attention. For example, as mentioned in 3.1, participants' beliefs varied significantly and systematically across treatments, while robustness checks with measures of attention (see section 3.3) yielded the same main effects. These patterns would not emerge if participants were providing random responses.

Our main conclusion from the above is that we can discard many different channels for explaining our results. Given the findings of our follow-up study, the most reasonable explanation for the ones we report here is that the heuristic of selecting the advisor one disagrees with is deeply counter-intuitive. Most participants cannot imagine settings where such a heuristic would yield useful guidance and so they opt for the more natural heuristic of following the advisor they agree with. We claim that this feature lies in the root of the 'identify the Expert' problem.

4 Conclusions

We have conducted an experimental examination of the novel 'identify the Expert' task. With people's everyday work, and knowledge in general, being ever more specialized, seeking expert advice is becoming an ubiquitous need. From choosing a politician to finding medical advice, to even selecting a technician, even people with the highest human capital constantly need experts to heed or hire, on unknown (to them) topics. Our main message is that a Populist influencer who promotes advice that panders to people's prior beliefs is more likely than not to be identified as the Expert – falsely. This holds in the relatively technical and counterintuitive domain of economics, but we conjecture that results will be similar in other domains where people are not particularly knowledgeable.[12] The inability of detecting experts is robust to attempts to reduce participants' overconfidence or ego-involvement. The existence of a clear, relatively straightforward Bayesian optimum, is also to not much avail (see Treatments 2 and 3).

The online environment of our experiments matches the natural setting in which self-proclaimed experts often compete in offering advice to laypeople. In our design, no clues about the identity of the advisors is provided. This again matches many situations of online advice, where credentials and identity are not provided, nor are verifiable. We did not provide feedback mechanisms (such as "like" buttons), but given the success of the Populist, such mechanisms would possibly exacerbate the problem.

Given the applied setting of our experiments in terms of the expertise domain, the representative sample, and the natural online setting that corresponds to the target environment, we have reasons to believe that our results are likely to offer some general insights for the 'identify the Expert' problem. There is evidence that many laypeople have little knowledge of important economic concepts, mostly restricted to the ones they experience in their daily lives (Runge and Hudson, 2020). This is consistent with our findings and suggests that the majority of the population will likely have trouble identifying the Expert, rendering them susceptible to the strategy followed by the Populist. This can lead to "knowledge poverty traps", whereby citizens who are not knowledgeable select advisors who offer useless information, further perpetuating their lack of knowledge.

A key insight from our work concerns the inability of participants to choose advisors against their priors, even if they are ignorant regarding a given domain and are aware of this ignorance. Arguably, the required heuristic in this case ("admit your ignorance, choose advisors you disagree with") is very simple, but not intuitive or psychologically palatable. Extensions to other domains of expertise, ranging from questions of mere knowledge (e.g. geography) to more technical ones (e.g. medicine), are needed in order to examine the scope of this phenomenon. We conjecture that the difficulty of the topic will matter, but so will the extent to which average people think they *ought* to be knowledgeable, even if they are not. Potentially, what matters is the degree to which science provides answers that agree to laypeople's intuitions in a given domain. As seen from participant beliefs, identifying the Expert is particularly unlikely if questions exist where people are confidently choosing the wrong answer. In these cases the Populist has a clear advantage.

In addition, several features of the informational environment in modern public debates makes

[12] On a related topic, Biermann et al. (2022) find that humans cannot verify the quality of algorithms meant to help them in making decisions.

the applicability of the ‘identify the Expert’ problem extend in domains beyond economics, possibly to most topics where forming accurate opinions is counterintuitive. Specifically, the following four features are common in expert-detection problems. (a) Anonymity/pseudonymity in the social media. (b) Even when advisors are not anonymous, credentials do not matter as much as they used to. For example, people are far less likely to follow academics on social media platforms as opposed to celebrities. (c) Detecting what credentials are valuable is a major problem by itself. Laypeople cannot easily distinguish between academics of different quality who offer advice in their discipline. (d) Public debates often have academics supporting both sides of the argument.

Democracies, from ancient Greek cities to the modern world, encourage participation in the public discourse and are thus particularly susceptible to the expert misidentification problem. Modern public discourse involves increasing pluralism and a falling value of official credentials. Our results indicate that this changing nature of the public debate is making the expert detection problem more difficult for the average citizen, requiring the application of counterintuitive heuristics. Investigating how different demographic groups address this task, and what can be done to assist them, is left for future work.

Acknowledgments

Part of this study has been conducted while Maniadis was supported by the Horizon 2020 grant 857636 — SInnoPSis — H2020-WIDESPREAD-2018-2020. Maniadis and Boukouras were supported by BA/Leverhulme small research grant SG162362. We are grateful to John Hey, Paul J. Healy, Roberto Weber, Axel Ockenfels, David Cooper, Constantine Sedikides, Santiago Sanchez-Pages as well as seminar participants at the University of Southampton, City University, King’s College London, University of Leicester, University of Maastricht, University of Durham, University of Heidelberg, ESA Global Meeting in Los Angeles, CRETE 2019 in Tinos, Greece, HeiKaMax 2022, and ESA European Meeting 2022.

Appendix

Full regression tables from main text

In this appendix we shall present in detail our extended empirical specifications. Table 8 describes the full set of covariates used to test our pre-registered hypotheses. The variable **Political leaning** is constructed as the sum of the answers of each participant to the following questions: *In political matters, people talk of 'the left' and 'the right'. How would you place your views on this scale generally speaking?* (1 is leftmost, 7 is rightmost), *Please indicate your level of agreement with the following sentence: 'The government should take more responsibility to ensure that everyone is provided for'.* (1 is strong agreement, 7 is strong disagreement), *Please indicate your level of agreement with the following sentence: 'Competition is good. It stimulates people to work hard and develop new ideas.'* (1 is strong disagreement, 7 is strong agreement) and *Generally speaking, would you say that most people can be trusted or that you cannot be too careful in dealing with people?* (1 is strong agreement, 7 is strong disagreement).

The variable **Political participation** is a dummy variable constructed from answers to the question: *How often do you vote in the general elections?* (1 is never, 7 is always). Participants who answered 1, 2, or 3 are classified as low in **Political participation**. **Econ knowledge** is a dummy variable constructed from answers to the question: *When it comes to matters of public policy, such as the minimum wage, taxes, or public investment, how knowledgeable do you consider yourself?* (1 is not at all, 7 is very knowledgeable). Those who answered 1, 2, or 3 are classified as low in **Econ knowledge**.

Variable	Description
Treatment (control group: Baseline)	
No Overconfidence	1 if Treatment is Treatment 2
No Ego-Involvement	1 if Treatment is Treatment 3
Age	Self-reported age of the participant
Sex (control group: Female)	
Nationality (control group: UK)	
Marital status (control group: In a relationship)	
Income (control group: under £20,000)	
Political leaning	Sum of 4 Likert scale questions as per pre-registrations
Political participation (control group: High)	
Low	1 if voting frequency strictly below 4 in Likert scale
Econ knowledge (control group: High)	
Low	1 if self-reported knowledge strictly below 4 in Likert scale
Attention	Sum of time spent on Stages 1 and 2 as per pre-registration
Highest educational level (control group: Secondary school up to 16 years)	
Occupational sector (control group: Business and sales)	Categorized as per pre-registration

Table 8: *Description of control variables used in the regressions - Full table*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment: No Overconfidence	0.195	0.197	0.181	0.199	0.192	0.191	0.213	0.192	0.184
	(0.162)	(0.163)	(0.163)	(0.163)	(0.163)	(0.163)	(0.164)	(0.164)	(0.169)
Treatment: No Ego-involvement	0.220	0.221	0.219	0.227	0.223	0.230	0.265*	0.221	0.257
	(0.155)	(0.155)	(0.155)	(0.155)	(0.155)	(0.155)	(0.156)	(0.158)	(0.161)
Age	0.001	0.000	0.002	0.001	0.001	0.002	0.002	0.002	0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
Sex: Male	0.523***	0.525***	0.508***	0.473***	0.524***	0.514***	0.504***	0.515***	0.453***
	(0.132)	(0.132)	(0.133)	(0.135)	(0.132)	(0.132)	(0.138)	(0.140)	(0.151)
Sex: Other	1.633*	1.647**	1.654*	1.446*	1.622*	1.578*	1.550*	1.635*	1.371
	(0.847)	(0.833)	(0.870)	(0.868)	(0.836)	(0.806)	(0.831)	(0.869)	(0.902)
Nationality: EU	0.442	0.449	0.445	0.438	0.426	0.383	0.459*	0.434	0.384
	(0.272)	(0.273)	(0.274)	(0.272)	(0.273)	(0.278)	(0.275)	(0.269)	(0.281)
Nationality: Other	−0.289	−0.292	−0.278	−0.259	−0.306	−0.395	−0.295	−0.329	−0.392
	(0.373)	(0.377)	(0.374)	(0.383)	(0.371)	(0.383)	(0.380)	(0.384)	(0.410)
Marital status									
Married		0.099							0.076
		(0.182)							(0.192)
Single		−0.042							−0.043
		(0.167)							(0.173)
Income									
£20,000 - £30,000			0.123						0.074
			(0.185)						(0.194)
£30,001 - £44,000			0.200						0.180
			(0.192)						(0.201)
£44,001 and above			0.225						0.155
			(0.181)						(0.206)
Political leaning				−0.009					−0.006
				(0.021)					(0.022)
Political participation: Low				0.036					0.123
				(0.193)					(0.202)
Econ knowledge: Low				−0.423***					−0.394**
				(0.157)					(0.163)
Attention					0.000				0.000
					(0.000)				(0.000)
Highest educational level									
Primary school						1.579			1.449
						(4.007)			(3.037)
Higher or secondary						0.261			0.130
						(0.230)			(0.244)
College or university						0.232			0.062
						(0.213)			(0.231)
Postgraduate						0.548**			0.340
						(0.238)			(0.275)
Discipline studied									
Business, Management, and Econnomics							−0.203		−0.264
							(0.190)		(0.202)
None							−0.451		−0.438
							(0.347)		(0.371)
Sciences, Maths, and Engineering							0.027		−0.033
							(0.169)		(0.181)
Social Sciences							0.096		−0.022
							(0.219)		(0.233)
Occupational sector									
Health								−0.125	−0.163
								(0.234)	(0.249)
Other								−0.357	−0.413
								(0.317)	(0.327)
Sciences and Engineering								0.303	0.156
								(0.228)	(0.246)
Student								0.038	0.035
								(0.256)	(0.264)
Teaching and Protective service								−0.083	−0.223
								(0.197)	(0.214)
Constant	4.903***	4.935***	4.772***	5.201***	4.823***	4.598***	4.904***	4.895***	4.897***
	(0.208)	(0.215)	(0.241)	(0.352)	(0.236)	(0.287)	(0.229)	(0.232)	(0.502)
R^2	0.040	0.042	0.043	0.053	0.041	0.051	0.047	0.048	0.077
Adjusted R^2	0.029	0.027	0.027	0.037	0.028	0.033	0.030	0.028	0.030
Observations	600	600	600	600	600	600	600	600	600

Robust standard errors in brackets. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 9: *Regressions on the number of correct answers in Stage 1*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment: No Overconfidence	0.092*	0.093*	0.093*	0.096*	0.091*	0.092*	0.092*	0.093*	0.094*
	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(0.050)
Treatment: No Ego-involvement	0.009	0.009	0.009	0.013	0.010	0.009	0.013	0.010	0.016
	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)	(0.049)	(0.048)	(0.050)
Age	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	0.000	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Sex: Male	0.023	0.025	0.025	0.020	0.024	0.026	0.023	0.026	0.031
	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)	(0.041)	(0.041)	(0.044)
Sex: Other	0.417	0.417	0.422	0.349	0.414	0.419	0.409	0.405	0.350
	(0.284)	(0.286)	(0.289)	(0.251)	(0.285)	(0.280)	(0.295)	(0.319)	(0.297)
Nationality: EU	0.143*	0.141*	0.138	0.156*	0.139*	0.139*	0.144*	0.149*	0.153*
	(0.083)	(0.084)	(0.084)	(0.083)	(0.083)	(0.084)	(0.084)	(0.084)	(0.088)
Nationality: Other	0.066	0.062	0.063	0.083	0.062	0.067	0.066	0.061	0.077
	(0.118)	(0.119)	(0.121)	(0.120)	(0.118)	(0.120)	(0.121)	(0.125)	(0.136)
Marital status									
Married		0.027							0.045
		(0.053)							(0.055)
Single		0.014							0.010
		(0.052)							(0.054)
Income									
£20,000 - £30,000			0.022						0.029
			(0.057)						(0.060)
£30,001 - £44,000			0.003						0.020
			(0.056)						(0.061)
£44,001 and above			−0.041						−0.032
			(0.055)						(0.061)
Political leaning				−0.010					−0.010
				(0.006)					(0.007)
Political participation: Low				−0.069					−0.086
				(0.059)					(0.063)
Econ knowledge: Low				−0.076					−0.087*
				(0.046)					(0.047)
Attention					0.000				0.000
					(0.000)				(0.000)
Highest educational level									
Primary school						−0.351			−0.352
						(0.629)			(0.279)
Higher or secondary						0.018			−0.024
						(0.070)			(0.074)
College or university						−0.022			−0.052
						(0.063)			(0.071)
Postgraduate						−0.004			−0.060
						(0.072)			(0.084)
Discipline studied									
Business, Management, and Econnomics							−0.054		−0.028
							(0.059)		(0.063)
None							−0.071		−0.063
							(0.089)		(0.097)
Sciences, Maths, and Engineering							−0.024		−0.034
							(0.050)		(0.053)
Social Sciences							−0.062		−0.062
							(0.067)		(0.071)
Occupational sector									
Health								0.016	0.014
								(0.070)	(0.074)
Other								−0.019	−0.027
								(0.099)	(0.102)
Sciences and Engineering								0.035	0.039
								(0.068)	(0.073)
Student								0.124	0.114
								(0.082)	(0.087)
Teaching and Protective service								0.029	0.025
								(0.057)	(0.062)
Constant	0.333***	0.330***	0.341***	0.523***	0.315***	0.339***	0.362***	0.276***	0.521***
	(0.067)	(0.070)	(0.075)	(0.110)	(0.074)	(0.091)	(0.075)	(0.075)	(0.154)
R^2	0.020	0.021	0.023	0.034	0.021	0.022	0.023	0.025	0.047
Adjusted R^2	0.009	0.006	0.006	0.018	0.008	0.004	0.005	0.005	−0.001
Observations	600	600	600	600	600	600	600	600	600

Robust standard errors in brackets. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Table 10: *Regressions on choosing the Expert in Stage 2 - Full table*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment: Overconfidence	0.214***	0.213***	0.217***	0.217***	0.215***	0.213***	0.219***	0.209***	0.218***	0.198***
	(0.058)	(0.058)	(0.059)	(0.058)	(0.058)	(0.059)	(0.059)	(0.058)	(0.062)	(0.062)
Treatment: Ego-involvement	0.156***	0.156***	0.156***	0.158***	0.160***	0.156***	0.170***	0.153***	0.170***	0.154***
	(0.052)	(0.052)	(0.052)	(0.052)	(0.052)	(0.052)	(0.053)	(0.052)	(0.054)	(0.056)
Easy task	0.781***	0.783***	0.782***	0.783***	0.784***	0.779***	0.794***	0.781***	0.799***	0.616***
	(0.056)	(0.058)	(0.057)	(0.057)	(0.056)	(0.056)	(0.057)	(0.058)	(0.064)	(0.096)
Treatment: Overconfidence * Easy task	−0.510***	−0.509***	−0.507***	−0.509***	−0.512***	−0.508***	−0.524***	−0.521***	−0.536***	−0.553***
	(0.109)	(0.110)	(0.109)	(0.110)	(0.108)	(0.110)	(0.109)	(0.109)	(0.113)	(0.112)
Treatment: Ego-involvement * Easy task	−0.360***	−0.361***	−0.366***	−0.365***	−0.361***	−0.356***	−0.394***	−0.373***	−0.416***	−0.509***
	(0.123)	(0.125)	(0.126)	(0.122)	(0.122)	(0.125)	(0.122)	(0.125)	(0.128)	(0.136)
Age	−0.001	−0.000	−0.001	−0.000	−0.001	−0.001	−0.001	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Sex: Male	0.001	−0.000	0.004	0.010	0.001	0.002	−0.011	0.007	0.006	0.012
	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.044)	(0.046)	(0.047)
Sex: Other	0.430***	0.428**	0.437***	0.352**	0.423**	0.424***	0.427***	0.460**	0.378**	0.412**
	(0.165)	(0.170)	(0.162)	(0.160)	(0.175)	(0.162)	(0.147)	(0.180)	(0.166)	(0.165)
Nationality: EU	0.111	0.112	0.104	0.116	0.103	0.107	0.113	0.113	0.107	0.120
	(0.083)	(0.085)	(0.085)	(0.081)	(0.084)	(0.083)	(0.083)	(0.084)	(0.088)	(0.088)
Nationality: Other	0.156	0.160	0.156	0.147	0.145	0.155	0.158	0.147	0.133	0.148
	(0.143)	(0.143)	(0.147)	(0.141)	(0.142)	(0.144)	(0.146)	(0.156)	(0.160)	(0.158)
Marital status										
Married		−0.020							−0.015	−0.015
		(0.058)							(0.060)	(0.060)
Single		−0.009							−0.026	−0.022
		(0.057)							(0.059)	(0.059)
Income										
£20,000 - £30,000			−0.009						0.015	0.025
			(0.060)						(0.063)	(0.064)
£30,001 - £44,000			−0.010						0.027	0.051
			(0.060)						(0.065)	(0.065)
£44,001 and above			−0.078						−0.044	−0.024
			(0.057)						(0.063)	(0.064)
Political leaning				−0.011*					−0.010	−0.010
				(0.006)					(0.007)	(0.007)
Attention					0.000				0.000	0.000*
					(0.000)				(0.000)	(0.000)
Highest educational level										
Higher or secondary						0.019			−0.016	−0.011
						(0.074)			(0.077)	(0.077)
College or university						−0.003			−0.022	−0.018
						(0.066)			(0.074)	(0.075)
Postgraduate						0.022			−0.002	−0.012
						(0.076)			(0.089)	(0.088)
Discipline studied										
Business, Management, and Econnomics							−0.088		−0.069	−0.051
							(0.064)		(0.067)	(0.067)
None							−0.080		−0.077	−0.091
							(0.098)		(0.106)	(0.107)
Sciences, Maths, and Engineering							0.038		0.034	0.043
							(0.053)		(0.059)	(0.058)
Social Sciences							−0.089		−0.081	−0.071
							(0.065)		(0.071)	(0.074)
Occupational sector										
Health								0.013	−0.013	−0.023
								(0.074)	(0.080)	(0.080)
Other								−0.021	−0.034	−0.033
								(0.095)	(0.097)	(0.095)
Sciences and Engineering								0.081	0.039	0.037
								(0.076)	(0.085)	(0.085)
Student								0.175*	0.127	0.125
								(0.090)	(0.100)	(0.099)
Teaching and Protective service								0.045	0.028	0.025
								(0.058)	(0.064)	(0.063)
Total own										−0.145**
										(0.069)
Total own2										0.017**
										(0.007)
Average confidence										−0.003
										(0.002)
Constant	0.191***	0.192**	0.218***	0.350***	0.145*	0.183*	0.209**	0.111	0.270*	0.720***
	(0.072)	(0.077)	(0.078)	(0.115)	(0.081)	(0.096)	(0.082)	(0.078)	(0.162)	(0.252)
Additional controls	Nationality	Marital status	Income	Political leaning	Attentions	Educational level	Discipline studied	Occupational sector	All included	All included
R^2	0.193	0.193	0.197	0.199	0.196	0.193	0.204	0.203	0.222	0.236
Adj. R^2	0.175	0.172	0.174	0.179	0.176	0.170	0.179	0.176	0.170	0.179
Num. obs.	465	465	465	465	465	465	465	465	465	465
RMSE	0.441	0.442	0.442	0.440	0.441	0.443	0.440	0.441	0.443	0.440

Participants who have an equal number of common answers with the two advisers are excluded. Robust standard errors in brackets. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Table 11: *Exploratory regressions on choosing the Expert in Stage 2 - Full table*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment: No Ego-involvement	−0.061	−0.058	−0.065	−0.062	−0.057	−0.064	−0.057	−0.065	−0.060	−0.047
	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)	(0.062)	(0.062)	(0.061)	(0.065)	(0.066)
Age	−0.001	−0.000	−0.001	−0.001	−0.001	−0.001	−0.001	0.001	0.000	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Sex: Male	−0.052	−0.052	−0.047	−0.037	−0.056	−0.045	−0.067	−0.043	−0.041	−0.039
	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)	(0.062)	(0.061)	(0.064)	(0.067)	(0.069)
Sex: Other	0.639	0.655***	0.629	0.515	0.651	0.625***	0.664***	0.672	0.598	0.624
	(0.670)	(0.075)	(2.002)	(0.510)	(2.001)	(0.081)	(0.082)	(1.003)	(1.014)	(2.008)
Nationality: EU	0.108	0.110	0.104	0.105	0.107	0.096	0.113	0.126	0.132	0.144
	(0.127)	(0.131)	(0.132)	(0.125)	(0.127)	(0.129)	(0.124)	(0.131)	(0.141)	(0.144)
Nationality: Other	0.155	0.177	0.153	0.143	0.151	0.145	0.171	0.170	0.176	0.181
	(0.180)	(0.177)	(0.186)	(0.173)	(0.178)	(0.184)	(0.182)	(0.185)	(0.179)	(0.184)
Marital status										
Married		−0.087							−0.097	−0.097
		(0.084)							(0.092)	(0.094)
Single		−0.043							−0.081	−0.076
		(0.080)							(0.085)	(0.087)
Income										
£20,000 - £30,000			0.001						0.049	0.052
			(0.086)						(0.091)	(0.093)
£30,001 - £44,000			−0.017						0.031	0.049
			(0.089)						(0.093)	(0.094)
£44,001 and above			−0.090						−0.019	−0.002
			(0.082)						(0.092)	(0.094)
Political leaning				−0.014*					−0.014	−0.013
				(0.009)					(0.010)	(0.010)
Attention					0.000				0.000	0.000
					(0.000)				(0.000)	(0.000)
Highest educational level										
Higher or secondary						−0.029			−0.104	−0.089
						(0.103)			(0.114)	(0.114)
College or university						−0.080			−0.131	−0.120
						(0.094)			(0.114)	(0.114)
Postgraduate						0.004			−0.037	−0.035
						(0.116)			(0.142)	(0.144)
Discipline studied										
Business, Management, and Econnomics							−0.007		−0.001	0.002
							(0.085)		(0.092)	(0.093)
None							0.004		−0.031	−0.067
							(0.121)		(0.140)	(0.143)
Sciences, Maths, and Engineering							0.106		0.159*	0.155*
							(0.078)		(0.087)	(0.088)
Social Sciences							0.014		0.023	0.017
							(0.108)		(0.121)	(0.125)
Occupational sector										
Health								−0.150	−0.179	−0.183
								(0.100)	(0.117)	(0.116)
Other								−0.062	−0.088	−0.083
								(0.134)	(0.146)	(0.146)
Sciences and Engineering								−0.112	−0.212*	−0.199
								(0.113)	(0.127)	(0.135)
Student								0.179	0.111	0.106
								(0.122)	(0.131)	(0.133)
Teaching and Protective service								−0.027	−0.003	−0.018
								(0.091)	(0.099)	(0.100)
Total own correct answers										−0.182
										(0.184)
Total own2 correct answers^2										0.020
										(0.021)
Average confidence in own answers										−0.002
										(0.003)
Constant	0.442***	0.446***	0.478***	0.650***	0.389***	0.493***	0.412***	0.379***	0.646**	1.120**
	(0.101)	(0.107)	(0.117)	(0.168)	(0.111)	(0.136)	(0.113)	(0.120)	(0.251)	(0.456)
R^2	0.021	0.025	0.027	0.031	0.026	0.026	0.031	0.045	0.090	0.098
Adjusted R^2	−0.002	−0.005	−0.007	0.004	−0.000	−0.008	−0.007	0.004	−0.004	−0.008
Oberrvations	266	266	266	266	266	266	266	266	266	266

Robust standard errors in brackets. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Table 12: *Exploratory regressions on choosing the Expert in Stage 2 among those who observe Decision-Makers with strictly more common answers with the Populist - Comparison of Treatments 2 and 3*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment: No Overconfidence	0.106*	0.107*	0.112**	0.113**	0.106*	0.105*	0.104*	0.102*	0.106*
	(0.056)	(0.056)	(0.056)	(0.056)	(0.056)	(0.056)	(0.057)	(0.057)	(0.059)
Treatment: No Ego-involvement	0.034	0.034	0.033	0.037	0.034	0.030	0.042	0.030	0.032
	(0.056)	(0.056)	(0.056)	(0.056)	(0.056)	(0.057)	(0.058)	(0.057)	(0.059)
Age	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.000	−0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Sex: Male	0.014	0.014	0.017	0.023	0.014	0.017	0.014	0.008	0.021
	(0.046)	(0.047)	(0.046)	(0.046)	(0.046)	(0.046)	(0.048)	(0.049)	(0.053)
Sex: Other	0.320	0.320	0.321	0.272	0.319	0.339	0.316	0.295	0.293
	(0.391)	(0.401)	(0.414)	(0.362)	(0.391)	(0.391)	(0.415)	(0.443)	(0.446)
Nationality: EU	0.119	0.117	0.112	0.119	0.119	0.121	0.119	0.123	0.114
	(0.090)	(0.091)	(0.092)	(0.089)	(0.090)	(0.092)	(0.091)	(0.091)	(0.096)
Nationality: Other	0.186	0.185	0.186	0.184	0.184	0.200	0.189	0.192	0.204
	(0.139)	(0.139)	(0.144)	(0.138)	(0.139)	(0.140)	(0.142)	(0.150)	(0.156)
Marital status									
Married		−0.007							0.022
		(0.062)							(0.065)
Single		0.015							0.021
		(0.063)							(0.067)
Income									
£20,000 - £30,000			−0.008						0.004
			(0.067)						(0.072)
£30,001 - £44,000			−0.025						0.004
			(0.066)						(0.073)
£44,001 and above			−0.095						−0.063
			(0.065)						(0.072)
Political leaning				−0.012*					−0.012
				(0.007)					(0.008)
Attention					0.000				0.000
					(0.000)				(0.000)
Highest educational level									
Higher or secondary						0.066			0.036
						(0.080)			(0.083)
College or university						−0.009			−0.042
						(0.071)			(0.080)
Postgraduate						0.001			−0.041
						(0.082)			(0.095)
Discipline studied									
Business, Management, and Econnomics							−0.055		−0.039
							(0.070)		(0.076)
None							−0.114		−0.155
							(0.106)		(0.113)
Sciences, Maths, and Engineering							−0.034		−0.040
							(0.059)		(0.063)
Social Sciences							−0.052		−0.059
							(0.082)		(0.086)
Occupational sector									
Health								−0.053	−0.067
								(0.080)	(0.085)
Other								−0.014	−0.038
								(0.118)	(0.122)
Sciences and Engineering								0.054	0.056
								(0.080)	(0.086)
Student								0.112	0.045
								(0.107)	(0.114)
Teaching and Protective service								−0.025	−0.038
								(0.064)	(0.069)
Constant	0.335***	0.328***	0.375***	0.501***	0.321***	0.314***	0.370***	0.299***	0.551***
	(0.080)	(0.085)	(0.091)	(0.129)	(0.094)	(0.104)	(0.091)	(0.089)	(0.183)
R^2	0.024	0.024	0.030	0.030	0.024	0.027	0.027	0.030	0.048
Adjusted R^2	0.008	0.004	0.007	0.013	0.006	0.005	0.003	0.003	−0.010
Observations	450	450	450	450	450	450	450	450	450

Robust standard errors in brackets. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Table 13: *Regressions on choosing the Expert in Stage 2 - Excluding those at the bottom quartile of total attention*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment: No Overconfidence	0.112	0.114	0.112	0.120*	0.113	0.110	0.115	0.104	0.106
	(0.071)	(0.071)	(0.071)	(0.071)	(0.071)	(0.071)	(0.072)	(0.073)	(0.076)
Treatment: No Ego-involvement	0.035	0.035	0.035	0.039	0.035	0.032	0.037	0.023	0.025
	(0.070)	(0.070)	(0.070)	(0.069)	(0.070)	(0.071)	(0.071)	(0.071)	(0.074)
Age	−0.001	−0.000	−0.001	−0.000	−0.001	−0.001	−0.001	−0.000	−0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Sex: Male	−0.060	−0.068	−0.064	−0.052	−0.061	−0.057	−0.057	−0.078	−0.076
	(0.057)	(0.059)	(0.058)	(0.057)	(0.058)	(0.058)	(0.059)	(0.062)	(0.066)
Sex: Other	0.267	0.260	0.292	0.203	0.266	0.281	0.258	0.260	0.251
	(0.391)	(0.397)	(0.408)	(0.356)	(0.392)	(0.387)	(0.400)	(0.420)	(0.388)
Nationality: EU	0.094	0.097	0.082	0.090	0.095	0.088	0.093	0.091	0.059
	(0.112)	(0.114)	(0.114)	(0.109)	(0.113)	(0.112)	(0.115)	(0.119)	(0.121)
Nationality: Other	0.154	0.159	0.154	0.152	0.154	0.151	0.162	0.166	0.171
	(0.146)	(0.145)	(0.149)	(0.145)	(0.146)	(0.149)	(0.150)	(0.151)	(0.159)
Marital status									
Married		−0.062							−0.034
		(0.079)							(0.087)
Single		−0.014							0.002
		(0.080)							(0.087)
Income									
£20,000 - £30,000			0.044						0.056
			(0.079)						(0.085)
£30,001 - £44,000			0.039						0.080
			(0.082)						(0.092)
£44,001 and above			−0.053						−0.007
			(0.083)						(0.097)
Political leaning				−0.015*					−0.015
				(0.008)					(0.010)
Attention					−0.000				−0.000
					(0.000)				(0.000)
Highest educational level									
Higher or secondary						0.097			0.084
						(0.096)			(0.104)
College or university						0.005			−0.003
						(0.085)			(0.099)
Postgraduate						0.053			0.051
						(0.101)			(0.123)
Discipline studied									
Business, Management, and Econnomics							−0.054		−0.081
							(0.089)		(0.099)
None							0.069		0.008
							(0.166)		(0.187)
Sciences, Maths, and Engineering							−0.032		−0.043
							(0.072)		(0.078)
Social Sciences							0.022		−0.005
							(0.102)		(0.111)
Occupational sector									
Health								−0.083	−0.086
								(0.102)	(0.114)
Other								0.067	0.062
								(0.147)	(0.162)
Sciences and Engineering								−0.018	−0.028
								(0.092)	(0.101)
Student								0.013	−0.080
								(0.135)	(0.150)
Teaching and Protective service								−0.058	−0.101
								(0.084)	(0.095)
Constant	0.380***	0.389***	0.387***	0.594***	0.397***	0.344***	0.391***	0.400***	0.639***
	(0.105)	(0.110)	(0.121)	(0.159)	(0.129)	(0.131)	(0.119)	(0.118)	(0.244)
R^2	0.026	0.028	0.031	0.036	0.026	0.032	0.029	0.030	0.057
Adjusted R^2	0.002	−0.002	−0.002	0.010	−0.001	−0.002	−0.008	−0.010	−0.033
Observations	300	300	300	300	300	300	300	300	300

Robust standard errors in brackets. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Table 14: *Regressions on choosing the Expert in Stage 2 - Excluding those at the bottom half of total attention*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment: No Overconfidence	0.088^{*}	0.088^{*}	0.088^{*}	0.094^{*}	0.087^{*}	0.087^{*}	0.089^{*}	0.089^{*}	0.093^{*}
	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.052)
Treatment: No Ego-involvement	−0.007	−0.007	−0.006	−0.004	−0.006	−0.007	−0.002	−0.006	0.000
	(0.049)	(0.050)	(0.050)	(0.049)	(0.050)	(0.050)	(0.050)	(0.050)	(0.052)
Age	−0.001	−0.001	−0.001	−0.000	−0.001	−0.001	−0.001	−0.000	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Sex: Male	0.012	0.013	0.015	0.022	0.013	0.016	0.016	0.013	0.035
	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.042)	(0.043)	(0.045)
Sex: Other	0.422	0.421	0.430	0.363	0.418	0.424	0.411	0.414	0.358
	(0.288)	(0.291)	(0.291)	(0.269)	(0.290)	(0.281)	(0.300)	(0.311)	(0.297)
Nationality: EU	0.151^{*}	0.149^{*}	0.147^{*}	0.156^{*}	0.145^{*}	0.147^{*}	0.153^{*}	0.154^{*}	0.145
	(0.085)	(0.086)	(0.087)	(0.084)	(0.086)	(0.086)	(0.086)	(0.087)	(0.090)
Nationality: Other	0.079	0.075	0.076	0.078	0.072	0.082	0.075	0.076	0.067
	(0.118)	(0.119)	(0.120)	(0.117)	(0.117)	(0.120)	(0.121)	(0.122)	(0.128)
Marital status									
Married		0.018							0.034
		(0.055)							(0.057)
Single		0.015							0.019
		(0.054)							(0.056)
Income									
£20,000 - £30,000			0.032						0.047
			(0.059)						(0.063)
£30,001 - £44,000			0.001						0.029
			(0.058)						(0.063)
£44,001 and above			−0.027						0.002
			(0.057)						(0.063)
Political leaning				-0.012^{*}					-0.013^{*}
				(0.006)					(0.007)
Attention					0.000				0.000
					(0.000)				(0.000)
Highest educational level									
Primary school						−0.324			−0.205
						(0.878)			(0.144)
Higher or secondary						0.035			−0.003
						(0.072)			(0.077)
College or university						−0.027			−0.077
						(0.065)			(0.075)
Postgraduate						−0.000			−0.059
						(0.074)			(0.086)
Discipline studied									
Business, Management, and Econnomics							−0.051		−0.040
							(0.062)		(0.066)
None							−0.136		-0.170^{*}
							(0.089)		(0.098)
Sciences, Maths, and Engineering							−0.054		−0.059
							(0.052)		(0.055)
Social Sciences							−0.060		−0.060
							(0.069)		(0.072)
Occupational sector									
Health								0.011	0.003
								(0.072)	(0.075)
Other								−0.004	−0.015
								(0.103)	(0.105)
Sciences and Engineering								0.018	0.024
								(0.069)	(0.075)
Student								0.077	0.047
								(0.087)	(0.092)
Teaching and Protective service								0.019	0.004
								(0.058)	(0.062)
Constant	0.334^{***}	0.330^{***}	0.337^{***}	0.497^{***}	0.304^{***}	0.336^{***}	0.372^{***}	0.301^{***}	0.525^{***}
	(0.070)	(0.073)	(0.079)	(0.112)	(0.076)	(0.093)	(0.078)	(0.078)	(0.156)
R^2	0.022	0.022	0.024	0.029	0.024	0.026	0.027	0.024	0.043
Adjusted R^2	0.010	0.006	0.006	0.015	0.010	0.006	0.008	0.003	−0.005
Observations	563	563	563	563	563	563	563	563	563

Robust standard errors in brackets. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Table 15: *Robustness check - Excluding participants with more than 1 inconsistency between choices and stated confidence.*

References

Aksoy, C., B. Eichengreen, and O. Saka (2020): "Revenge of the Experts: Will COVID-19 Renew or Diminish Public Trust in Science?" *Available at SSRN 3737580.*

Algan, Y., D. Cohen, E. Davoine, M. Foucault, and S. Stantcheva (2021): "Trust in Scientists in Times of Pandemic: Panel Evidence from 12 Countries," *Proceedings of the National Academy of Sciences*, 118, e2108576118.

Allcott, H., L. Braghieri, S. Eichmeyer, and M. Gentzkow (2020): "The Welfare Effects of Social Media," *American Economic Review*, 110, 629–76.

Allcott, H., M. Gentzkow, and C. Yu (2019): "Trends in the Diffusion of Misinformation on Social Media," *Research & Politics*, 6, 2053168019848554.

Alysandratos, T., A. Boukouras, S. Georganas, and Z. Maniadis (2020): "The Expert and the Charlatan: an Experimental Study in Economic Advice," *Available at SSRN 3644653.*

Alysandratos, T., A. Boukouras, S. Georganas, Z. Maniadis, and A. Velias (2023): "Science versus Populism: Why and When Do Charlatans Beat Experts?" *Working Paper.*

Amaral-Garcia, S., M. Nardotto, C. Propper, and T. Valletti (2022): "Mums Go Online: Is the Internet Changing the Demand for Health Care?" *Review of Economics and Statistics*, 104, 1157–1173.

Andre, P., C. Pizzinelli, C. Roth, and J. Wohlfart (2022): "Subjective Models of the Macroeconomy: Evidence from Experts and Representative Samples," *The Review of Economic Studies*, 89, 2958–2991.

Angeletos, G.-M., Z. Huo, and K. A. Sastry (2021): "Imperfect Macroeconomic Expectations: Evidence and Theory," *NBER Macroeconomics Annual*, 35, 1–86.

Autor, D., D. Dorn, G. Hanson, and K. Majlesi (2020): "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure," *American Economic Review*, 110, 3139–3183.

Bailey, P. E., T. Leon, N. C. Ebner, A. A. Moustafa, and G. Weidemann (2022): "A Meta-Analysis of the Weight of Advice in Decision-Making," *Current Psychology*, 1–26.

Biermann, J., J. Horton, and J. Walter (2022): "Algorithmic Advice as a Credence Good," *ZEW Discussion Paper*, 22.

Blendon, R. J., J. M. Benson, M. Brodie, R. Morin, D. E. Altman, D. Gitterman, M. Brossard, and M. James (1997): "Bridging the Gap Between the Public's and Economists' Views of the Economy," *Journal of Economic Perspectives*, 11, 105–118.

Bonaccio, S. and R. S. Dalal (2006): "Advice Taking and Decision-Making: An Integrative Literature Review, and Implications for the Organizational Sciences," *Organizational behavior and human decision processes*, 101, 127–151.

Bursztyn, L., A. Rao, C. Roth, and D. Yanagizawa-Drott (2022): "Opinions as Facts," *The Review of Economic Studies*, https://doi.org/10.1093/restud/rdac065.

Caplan, B. (2011): *The Myth of the Rational Voter: Why Democracies Choose Bad Policies-New Edition*, Princeton University Press.

Chakraborty, A., P. Ghosh, and J. , Roy (2020): "Expert Captured Democracies," *American Economic Review*, 110, 1713–51.

Charness, G., R. Oprea, and S. Yuksel (2021): "How Do People Choose Between Biased Information Sources? Evidence from a Laboratory Experiment," *Journal of the European Economic Association*, 19, 1656–1691.

Chopra, F., I. Haaland, and C. Roth (2022): "Do People Demand Fact-Checked News? Evidence From US Democrats," *Journal of Public Economics*, 205, 104549.

Eskreis-Winkler, L. and A. Fishbach (2019): "Not Learning From Failure—the Greatest Failure of All," *Psychological Science*, 30, 1733–1744.

——— (2022): "You Think Failure Is Hard? So Is Learning From It," *Perspectives on Psychological Science*, 174569162110598.

Falk, A. and F. Zimmermann (2017): "Information Processing and Commitment," *The Economic Journal*, 128, 1983–2002.

Fetzer, T. (2019): "Did Austerity Cause Brexit?" *American Economic Review*, 109, 3849–3886.

FUNKE, M., M. SCHULARICK, AND C. TREBESCH (2020): "Populist Leaders and the Economy," *CEPR Discussion Paper No. DP15405.*

GENTZKOW, M. AND J. M. SHAPIRO (2006): "Media Bias and Reputation," *Journal of Political Economy*, 114, 280–316.

GILLESPIE, T. (2018): *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*, Yale University Press.

GORDON, R. AND G. B. DAHL (2013): "Views Among Economists: Professional Consensus or Point-Counterpoint?" *American Economic Review*, 103, 629–35.

GURIEV, S. AND E. PAPAIOANNOU (2022): "The Political Economy of Populism," *Journal of Economic Literature*, 60, 753–832.

KRISHNA, V. AND J. MORGAN (2001): "A Model of Expertise," *The Quarterly Journal of Economics*, 116, 747–775.

LONG, J. S. AND L. H. ERVIN (2000): "Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model," *The American Statistician*, 54, 217–224.

MELOSO, D., S. NUNNARI, AND M. OTTAVIANI (2023): "Looking into Crystal Balls: A Laboratory Experiment on Reputational Cheap Talk," *Management Science.*

MUDDE, C. AND C. R. KALTWASSER (2017): *Populism: A Very Short Introduction*, Oxford University Press.

OLIVEROS, S. AND F. VÁRDY (2015): "Demand for Slant: How Abstention Shapes Voters' Choice of News Media," *The Economic Journal*, 125, 1327–1368.

PENNYCOOK, G., J. MCPHETRES, Y. ZHANG, J. G. LU, AND D. G. RAND (2020): "Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention," *Psychological Science*, 31, 770–780.

RONAYNE, D. AND D. SGROI (2018): "Ignoring Good Advice," *Competitive Advantage in the Global Economy (CAGE), Working Paper No. 359.*

RUNGE, J. AND N. HUDSON (2020): "Public Understanding of Economics and Economic Statistics," *EScoE Occasional Paper*, 3.

Sapienza, P. and L. Zingales (2013): "Economic Experts Versus Average Americans," *American Economic Review*, 103, 636–42.

Schotter, A. (2003): "Decision Making with Naive Advice," *American Economic Review*, 93, 196–201.

Thaler, M. (2021): "The Supply of Motivated Beliefs," *arXiv preprint arXiv:2111.06062*.

Trautmann, S. T. and G. van de Kuilen (2015): "Belief Elicitation: A Horse Race Among Truth Serums," *The Economic Journal*, 125, 2116–2135.