Science versus Populism: Why and When Do Charlatans Beat Experts?

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

October 6, 2023

Abstract

Jim completes a technical questionnaire and then receives recommendations by two advisors. He is told that one advisor is an expert, who answers all questions correctly, while the other is a populist and answers correctly only some. Jim also knows he has X correct answers, Y common answers with one advisor and X common answers with the other. Who should he pick? We recruit a large sample (12,000) of laypeople, representative of the general population, to answer this question. Even though the task is very simple, with a simple heuristic achieving the rational benchmark, the majority of participants fail to identify the expert. We modify the difficulty of the task by changing the presentation of the information between table and text, and by varying the number of common answers with the two advisors. Most people fail to identify the expert. The fraction of participants who answer correctly increases with the difference in commons and with text. Socio-demographic characteristics also play a role.

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 London, UK. Maniadis: Department of Economics and SInnoPSis, University of Cyprus, and Department of Economics, School of Social Sciences, University of Southampton. Velias: Georganas: Department of Economics, City University London, UK.

1 Introduction

As science and society co-evolve, reliance on expertise is becoming increasingly important. From dealing with a pandemic to climate change to combating inflation, sound public policies can be formed in democratic countries only if laypeople heed the advice of experts. However, since public debate is contested by several parties, the ability of laypeople to distinguish experts from populists becomes a crucial factor for the sustainability of good governance. How can ordinary citizens evaluate expert opinion and distinguish it from unreliable sources? Consider a debate with two speakers on a major issue of global importance, like the pandemic. How is the public to detect the expert if laypeople lack that highly specific knowledge themselves? How can democracies produce efficient policies if the voters confuse populists with experts?

In this paper we develop and test, on a representative sample of 12,000 participants, a simple task that captures the essence of the expertise detection problem. Imagine the following situation. You are asked to answer a set of questions on a technical topic such as economics or epidemics. You are told the number of common answers with each of two independent advisors along with your number of correct answers to the questionnaire. You are also told one of the advisors is an *expert* on this and answers all questions correctly while the other is a *populist* and answers only a fraction of them correctly. You are tasked with identifying the expert. How and whom should you pick?

This minimal version of the *Identify-the-Expert* task (and the easiest we could think of) has a simple but counter-intuitive solution. Since the expert is always correct and you know how many questions you have answered correctly, then it must be the case that *the expert is the advisor who has as many common answers with you as you have correct answers.* The implication of this is that knowledgeable individuals on the topic will tend to have many common answers with the expert, while non-knowledgeable ones will tend to have few. The resulting recommendation is both strikingly simple and counter-intuitive: *if you are knowledgeable in a topic, choose advisors you agree with. If you are not knowledgeable, you are better served by advisors that tell you something that you could not think of by yourself.*

Clearly, our simple task does not capture the information richness of the settings in which humans usually evaluate expertise. However, this is a feature rather than a caveat. Previous work in this area has paid attention to the content of advice (Bonaccio and Dalal, 2006), the credentials of the advisor (Algan et al., 2021) or the characteristics of the advisee (Oliveros and Várdy, 2015).

Factors such as overconfidence, ego-rents, motivated reasoning and beliefs (Chopra et al., 2022; Bursztyn et al., 2022; Thaler, 2021; Alysandratos et al., 2023) also play a role. With our design, we intentionally abstract from all these elements to remove additional sources of noise so that we best highlight an aspect of the problem that has received less attention: *even in an ideal information environment most people would face difficulty in identifying the expert because it would be the advisor they disagree with the most.*

We argue that this is due to three interlocking reasoning obstacles. First, individuals have to accept their own lack of knowledge. In previous work (Alysandratos et al., 2023) we showed that indeed, giving feedback regarding the subject's knowledgeability improves their chances of detecting the expert. Second, being rational is sometimes very counter-intuitive in this set-up. Even if you accept your lack of knowledge, you have to go the extra step of accepting as an expert someone you expressly disagree with! For people who tend to think in a fast and shallow manner (Kahneman, 2011) choosing a seemingly wrong advisor can be hard.

The third and crucial obstacle to rational behaviour lies deeply in the mechanics of any expert-detection task involving laypeople. The exact same people who are lacking knowledge, are faced with the cognitively hardest task: knowledgeable people have to choose advisors they agree with, while non-knowledgeable have to choose the ones they disagree with. Since expert-finding skills and knowledgeability are presumably correlated (indeed, they are in our sample), the task demands stronger cognitive power from the people who have less of it. Our results re-affirm this hypothesis. We find that high human-capital subjects find it easier to apply the counter-intuitive strategy than others. Consequently, to be able to persuade the median citizen, experts face a double challenge. They have to lower the citizens' disbelief regarding the expert's advice, while convincing them that any lingering disbelief might be a good sign.

In a between-subjects design, we test experimentally the above conjectures with the simple task described in the beginning. In addition, we vary two information-processing parameters: (i) the presentation of the information between table format and text format, and (ii) the number of common answers with the two advisors. We obtain several interesting results. First, subjects often fail to find the expert, even under favourable conditions. For example, when the subject has two answers in common with the populist and eight common with the expert, only 42.2% select the expert as advisor against 18.2% who select the populist. 18.67% say the two are equally likely, while 20.4% claim there is not enough information to make a choice. This result obtains despite the fact that participants face the easy task of picking the advisor with the most common answers.

Accordingly, in the reverse case of two common answers with the expert and eight with the populist, the respective percentages are 32.5%, 24.7%, 21.4% and 21.4%. We obtain similar results for the vast majority of configurations.

This points to a more general second result, namely congruence impedes subjects from understanding the nature of the problem. The chance of subjects answering correctly significantly decreases when the difference between common answers with the populist and common answers with the expert is large and in favour of the populist. Moreover, subjects are more likely to select the wrong answer or to be confused on how to answer. This is an indication that the cognitive load of the task increases when it becomes counter-intuitive.

Thirdly, the salience of the information provided, and individual subject characteristics, matter significantly. For example, revisiting the case of two answers in common with the expert and eight with the populist, restricting the sample to subjects with performance in terms of education and income, and including only those passing a very simple numeracy tests, increases expert choices to 39.7% and reduces the populist to 20.7%. Correspondingly, the frequency of subjects claiming the two are the same also falls significantly, to 10.3%.

The effect of summarizing the number of common answers in text, instead of presenting the answers extensively in a table, is more subtle. In general, it tends to make behaviour more systematic. For example, when the expert and the populist have the same number of correct answers, subject are much more likely to consider them equal when presented with text, than when presented with a table.

2 Experimental Design

As mentioned in the introduction, our design is based on the simple task of selecting one of two advisors on the basis of the correct answers of the participant and the common answers with the two advisors. In particular, in the *text*-treatments participants saw the following message:

" In a 10-question questionnaire about economics, Jim answered X questions correctly. Each question had two possible answers, A or B, with only one being correct. There are two advisers, J and M. One of the two advisers (we do not know which one) is an expert on the subject and answered all questions correctly in the above questionnaire. The other is a populist and answered only 40% of questions correctly. Jim had Y common answers with adviser J and X common answers with

adviser M. Based on this information alone, which of the two advisers, J or M, is most likely to be the expert on the subject?"

The relevant parameters of manipulation for these treatments are X, which is the number of correct answers and also the number of common answers with the expert, and Y, the number of common answers with the populist. X takes values between 1 and 10, while Y takes values between 0 and 10. In addition, we mention to participants that one of the advisors answers with 40% accuracy. This percentage comes from a previous study, where the populist answers four out of ten questions correctly.[1] This percentage constraints the number of common answers that 'Jim' can have with the former. All-in-all, there are 34 different possible ways for X, Y and the accuracy of the populist to be consistent with each other, and for each we construct a corresponding treatment. All possible and consistent combinations of X and Y are provided in appendix A.

The *table*-treatments are identical to the *text*-treatments with only a single difference. Instead of giving participants the relevant numbers in text, they saw a table with the answers for each question given by 'Jim' and the two advisors. Subjects needed to count the common answers and then conduct the deductive reasoning for the correct heuristic. Clearly, this treatment contains additional effort in reaching the correct conclusion and a reasonable hypothesis is that this makes it harder for subjects to identify the expert. Figure 1 below shows the presentation of information across the two treatment categories.

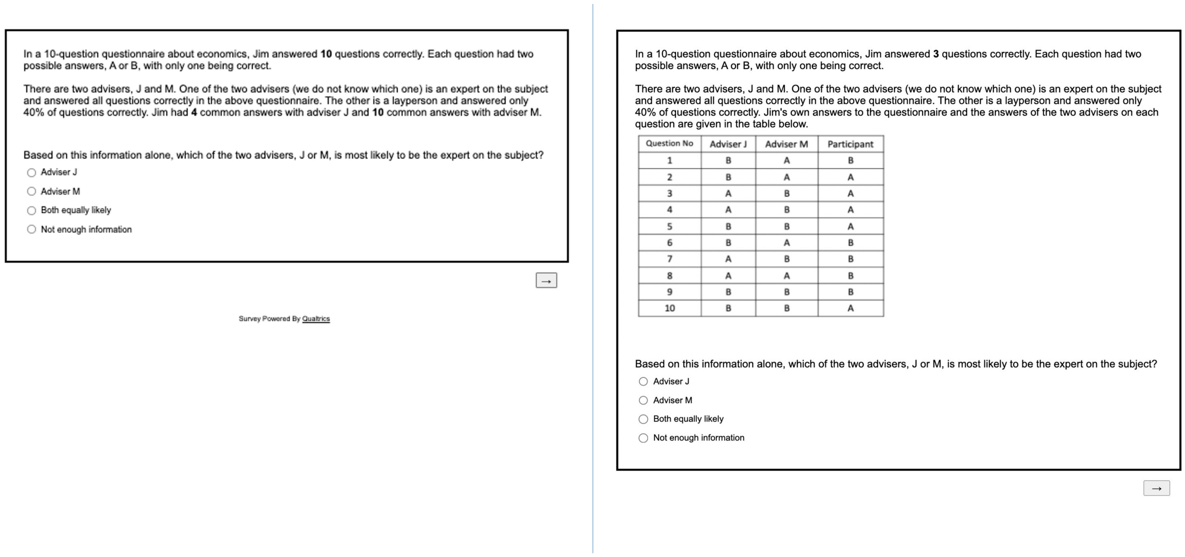
For both the *text*-treatment and the *table*-treatment category we created 34 sub-treatments based on the possible combinations of the relevant parameters, as discussed above. In addition, for robustness purposes, we created two high-incentives treatments, one with text and one with table presentation, where participants were allowed to join a lottery for winning the prize of 50 Euros. Overall, our design contains $2 \times 34 + 1 = 70$ different treatments.[2]

In all treatments, participants were given four response options. In addition to selecting advisor 'J' or 'M' as the expert, they could also answer 'Both are equally likely' or 'Not enough information'. The additional options are important. They allow us to separate those who are confused or unsure on how to understand the problem from those who make a choice because they believe they are applying the correct method. Besides, a significant factor for people mis-identifying experts is that they may not know how to evaluate informative signals. We want to rule out such cases in our

[1] See (Alysandratos et al., 2023) for details.

[2] See Table 6 in Appendix A for the full list of treatments.

experiment.



In a 10-question questionnaire about economics, Jim answered **10** questions correctly. Each question had two possible answers, A or B, with only one being correct.

There are two advisers, J and M. One of the two advisers (we do not know which one) is an expert on the subject and answered all questions correctly in the above questionnaire. The other is a layperson and answered only 40% of questions correctly. Jim had **4** common answers with adviser J and **10** common answers with adviser M.

Based on this information alone, which of the two advisers, J or M, is most likely to be the expert on the subject?

- Adviser J
- Adviser M
- Both equally likely
- Not enough information

→

Survey Powered By Qualtrics

In a 10-question questionnaire about economics, Jim answered **3** questions correctly. Each question had two possible answers, A or B, with only one being correct.

There are two advisers, J and M. One of the two advisers (we do not know which one) is an expert on the subject and answered all questions correctly in the above questionnaire. The other is a layperson and answered only 40% of questions correctly. Jim's own answers to the questionnaire and the answers of the two advisers on each question are given in the table below.

Question No	Adviser J	Adviser M	Participant
1	B	A	B
2	B	A	A
3	A	B	A
4	A	B	A
5	B	B	A
6	B	A	B
7	A	B	B
8	A	A	B
9	B	B	B
10	B	B	A

Based on this information alone, which of the two advisers, J or M, is most likely to be the expert on the subject?

- Adviser J
- Adviser M
- Both equally likely
- Not enough information

→

Figure 1: *The screens that participants saw in our experiment with the description of the task. The left screen presents the information in text and the right screen in a table format. Each subject participated in only one of the two presentation treatment categories.*

Note that in all treatments, advisor 'J' corresponds to the populist advisor and advisor 'M' corresponds to the expert advisor. Thus, 'Jim's number of correct answers is always equal to the number of common answers with 'M'. In treatments where $X \neq Y$, participants need to identify the correct heuristic, that is they need to understand that 'Jim' has as many common answers with 'M' as his own correct answers and so 'M' must be the advisor. For treatments where $X = Y$, both advisors have an equal number of common answers with 'Jim' so the heuristic does not apply. In these treatments we take the response 'both equally likely' to be the correct one, although we also do robustness analysis by taking the response 'not enough information' to also be correct.[3] The main hypothesis to be tested across all treatments is the following.

H1. *Less than half of all participants successfully identify the expert.*

Prima facie, the task is very simple, and if one realises what the appropriate heuristic is, then in treatments where $X \neq Y$ she can achieve perfect accuracy in identifying the expert. Moreover, the information setting is minimal with only three relevant numbers to process and the usual pitfalls, such as confirmation bias or bayesian updating, are absent from our design. However, our

[3]See also appendix A.

experience from our previous work (see Alysandratos et al., 2023) and the reasoning obstacles we mentioned in the introduction raise the prospect that the task is deeply counter-intuitive for most laypeople. Hence, we think it is important to formally check if laypeople can successfully perform this task despite its seeming simplicity or its abstract nature. To put it differently, if laypeople fail to identify the expert in this task, how likely is it that they will be able to perform the far more daunting task of spotting an expert in a real-world setting?

Consequently, it is worth testing whether the presentation of information or different parameter values make a difference on the task's difficulty. The presentation of the information in table format means that subjects must take the extra cognitive step of counting the common answers with the two advisors. If this leads to information overload or distracts away from considering the possibility of a different heuristic than the obvious one, then we should expect even worse performance in the table treatments.

H2. *The percentage of participants who correctly identify the expert in the table treatments is significantly less than the text treatments.*

Analogously, if the difference in commons answers between the two advisors is very small, the task becomes more difficult: it is harder to think counter-intuitively when the two advisors appear to be very similar in their responses. Subjects who face treatments with smaller differences in commons should have lower chances of using the correct heuristic.

H3. *The percentage of participants who correctly identify the expert is increasing in the absolute difference between the commons answers with the two advisors.*

3 Results

3.1 Data

Data collection took place between April 8th and 15th 2022, with 92% of the observations being collected in the first 4 days (April 8th-11th). The sample was provided by Lucid. A total of 12,000 responses were collected, including a 500 pilot run, and excluding responses that were thought of demonstrating insufficient attention. In addition, age and gender quotas targets were set so that the sample matches the demographic characteristics of the general population in the UK.[4]

[4]For more details see Appendix A.

The survey started with the experiment, either in text or table form and some questions regarding health status for a companion study. We then included a rough test of numeracy and some questions regarding sources of information and advisors that participants prefer in the field. In the final section we asked a series of demographics questions, including age, income and education.[5] The full distribution per demographic category is presented in appendix B. We briefly comment here that our demographic distributions are very close to the pre-set quota. The only exception is for individuals of the age group 75 and above, but since collection took place online this is not surprising. It is also inconsequential to the analysis of our main hypotheses.

3.2 Main Results

The first and most important result of the study is that people can not identify the expert. After aggregating across all treatments where $X \neq Y$, 3,300 out of a total of 9,870 participants or 33.4% answered correctly by selecting advisor M. 1,848 (18.7%) selected advisor J (the populist), 2,328 (23.6%) could not distinguish between the two and 2,394 (24.2%) responded that there was not enough information. Table 1 summarises this information. Thus, the fraction of subjects who identified the expert is statistically significantly lower than 50%. Although more participants selected advisor M than J, the inclusion of the categories 'both equally likely' and 'not enough information' is meaningful for comparison purposes, as it demonstrates that most people fail to identify the appropriate heuristic for the problem. Moreover, those who in the 'not enough information' category, which is a non-negligible fraction of all participants (24.2 % on average) believe that the problem is not well specified. This is a very strong indication that they consider it a mentally hard task.

Table 2 shows how many individuals answered correctly across all treatments once we include treatments where $X = Y$ and apply the appropriate correct response per treatment. In treatments where $X \neq Y$, a Bayesian decision maker would always correctly identify the expert as being the advisor who has as many common answers with 'Jim' as 'Jim' has correct answers. Hence, the correct response can be identified unambiguously. In treatments where $X = Y$, it is impossible to distinguish between the two advisors, so the correct response is 'both equally likely'. From Table 2 we observe that only 33.9% of subjects answered correctly, while close to 2/3 of participants gave incorrect responses. This finding is consistent with our interpretation that the task is cognitive hard. Note that even if one objects to our definition of a correct answer for treatments where $X = Y$, the fraction of correct responses are remarkably close between tables 1 and 2, where 33.4%

[5]The full questionnaire is available by the authors upon request.

of participants identify advisor M as the expert. Thus, our results are not driven by how we define correct responses.

Response	No of answers	Percentage
Advisor J	1,848	18.7
Advisor M	3,300	33.4
Both equally likely	2,328	23.6
Not enough information	2,394	24.2
Total	9,870	100.0

Table 1: *Number and percentage of answers per answer option. Pooled data across all treatments.*

Response	No of answers	Percentage
Correct	3,792	33.9
Incorrect	7,389	66.1
Total	11,181	100.0

Table 2: *Number and percentage of correct answers. Pooled data across all treatments.*

The same conclusions are preserved when we look at the data for each treatment. Figures 2, 3, 4 show the fractions of the four different responses across treatments for the case where we pool text and table format data together and for the case where we examine them separately. *Across all 84 different cells the percentage of subject is below* 50% *and this difference is statistically significant in all cases.*[6] This is evidence in support of hypothesis *H1*.

Nonetheless, there is substantial variation across parameter configurations. Above the diagonals when the task is relatively easy, as Jim has more answers in common with the expert than with the populist, the modal choice is the expert, followed by 'not enough information'. The populist is chosen least often, but still significantly more than zero.[7] Below the diagonals, the task is harder and in some cells the populist is chosen most often. This is striking, because it is direct evidence against subjects using the correct heuristic to reach decisions. Even more importantly, they are unaware of this mistake, because they select the populist instead of claiming there is not enough

[6]See appendix C for the summary of the statistical tests.
[7]See Table 12 in appendix C for the results of formal tests.

information or that they are equally likely. On the diagonals, where the expert and the populist have the same number of answers common with 'Jim', the modal choice is 'both equally likely', followed by 'not enough information'.

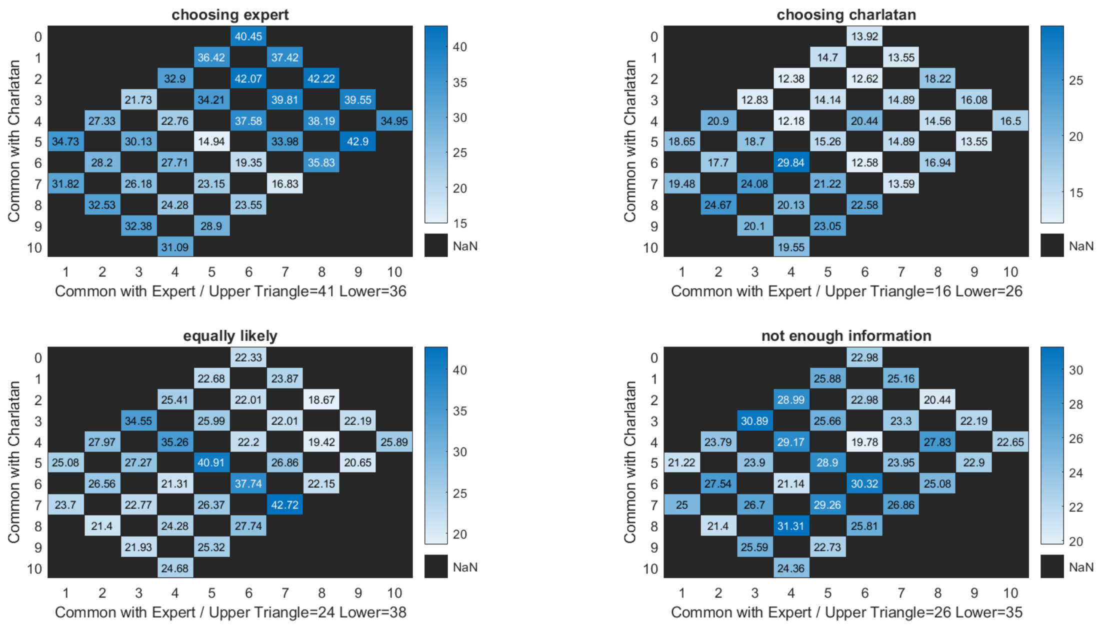


Figure 2: *Percentage of participants who gave each one of the four responses in each treatment. The horizontal axis shows 'Jim's number of correct answers and the vertical the number of common answers between 'Jim' and the populist advisor. Each cell represents a treatment across both formats, i.e. data for text-treatments and table-treatments are pooled. The heading 'choosing expert' corresponds to selecting advisor M, 'choosing charlatan' to selecting advisor J.*

In general, the difference in common answers can be thought of as the strength of a stimulus pushing subjects to correct choices, namely the further away the parameters are from the diagonal, the higher the chance that the participants will make correct choices. In accordance with hypothesis *H3*, we test this conjecture statistically. To be more precise, we exclude treatments where the number of common answers between 'Jim' and the advisors is equal, and we split the remaining treatments into two groups: those treatments where the absolute difference in the number of common answers with the two advisors is between one and three, and those treatments where the absolute difference is greater than three. We do this separately for *text*-treatments and *table*-treatments, so that in total we obtain for treatment groups: (i) *text*-treatments, one-to-three difference in common answers, (ii) *text*-treatments, more-than-three difference in common answers, (iii) *table*-treatments, one-to-three difference in common answers, (iv) *table*-treatments, more-than-three difference in common answers. We estimate the fraction of participants who identify the expert correctly for each one and compare the difference in the fractions between group (i) and (ii) and group (iii) and (iv). Both p-values for these differences are below 2%. Table 3 presents the exact values.

There is also evidence that many participants apply the wrong heuristic. Excluding the diagonals, where subjects choose the option 'equally likely' more often than any other response, the percentage of participants who choose the expert is higher than the percentage of those who choose the populist. Therefore, some individuals apply the correct heuristic, especially in the harder treatments of the lower triangle. However, it is also true that the difference between the percentage of those who choose the expert and those who choose the charlatan is larger for the upper triangle than the lower triangle of figures 2-4. This implies that some participants must be using the incorrect heuristic of selecting the expert with the highest number of commons when they should apply the opposite heuristic. To formalize this argument, we test whether the difference between the two percentages is greater in the upper triangle than the lower triangle of figure 2. Indeed, the corresponding p-value is 0, which is a strong evidence that a significant fraction of participants apply the wrong heuristic.

Pres	**NCor**$_{\leq 3}$	**Tot**$_{\leq 3}$	**NCor**$_{>3}$	**Tot**$_{>3}$	**Est**$_{\leq 3}$	**Est**$_{>3}$	**Pvalue**
Text	557	1845	1108	3139	0.3019	0.3530	0.00025
Table	536	1714	1099	3172	0.3127	0.3465	0.01856

Table 3: *Tests of hypothesis H3. Pres stands for presentation style, i.e. text or table. $NCor_{\leq 3}$ is the total number of participants who answer correctly in treatments where the difference in common answers is between one and three, while $Tot_{\leq 3}$ is the total number of participants in these treatments. $NCor_{>3}$ is the total number of participants who answer correctly in treatments where the difference in common answers is greater than three, while $Tot_{>3}$ is the total number of participants in these treatments. $Est_{\leq 3}$ and $Est_{>3}$ give the fraction of participants who answered correctly in the corresponding treatments groups. Finally, Pvalue shows the corresponding p-value for the difference between $Est_{\leq 3}$ and $Est_{>3}$.*

The final hypothesis we check is *H2* on whether the way information is presented matters. Initially, we conjectured and that text treatments, because the provide information in a much more concise format, may assist subjects in identifying the correct heuristic. Visual inspection of figures 3 and 4, however, does not provide a clear indication, and our formal tests reject the hypothesis.

Triangle	Com_1	Tot	Est	Pvalue
Upper	1,871	4,914	0.1650	0
Lower	1,090	4,956		

Table 4: *Test on the percentage of participants who selected the advisor with the most common answers between the upper and the lower triangle of figure 2. Com_1 stands for the total number of participants who selected the advisor with the highest number of common answers with 'Jim'. Tot stands for the total number of participants, Est for the estimated difference in percentages between the two triangles and Pvalue for the corresponding p-value.*

Specifically, the difference in the fraction of subjects who correctly identify the expert between *text*-treatments and *table*-treatments with different number of common and correct answers is not statistically significant for 28 out of 29 pairs (table 13 in appendix C). This indicates that the underlying heuristic is hard to discern for most laypeople even in extremely simplified settings.

Interestingly, there seems to be some effect when we consider treatments where the correct answer is 'both equally likely' (table 14 in appendix C). In three out of five configuration of answers, text-treatments had higher number of correct responses than table-treatments at the 1% significance level and in one configuration at the 5% level. Thus, the style of presentation does not seem to assist participants when $X \neq Y$, while it improves performance when $X = Y$. This is an indication that many subjects are at the limits of their cognitive abilities with the task. $X = Y$ is a symmetric problem, so presentation improves performance marginally for many subjects. Note that the average percentage of correct responses over all five configurations increases from 30.93% to 45.21%, an almost 50% increases in performance. This interpretation is corroborated further by the fact that presentation in text increases significantly the number of subject who report 'both equally likely' but has no impact on the number of subjects who report 'not enough information' (table 15 in appendix C). Clearly, subjects understand the difference between the two responses and systematically favor one over the other when provided with a slightly easier task.

3.3 The Role of Demographics

Individual subject characteristics have a strong effect on choices. As figure 5 shows, the frequency of subjects choosing the expert falls with age, and rises strongly with income and education. An outlier exists for education level 2, but the sample in that cell is very small. A question arising from the above facts is, how well do high-skill, high human-capital subjects perform? If we restrict the sample to subjects with university degree or higher education, below the age of 65 and individual

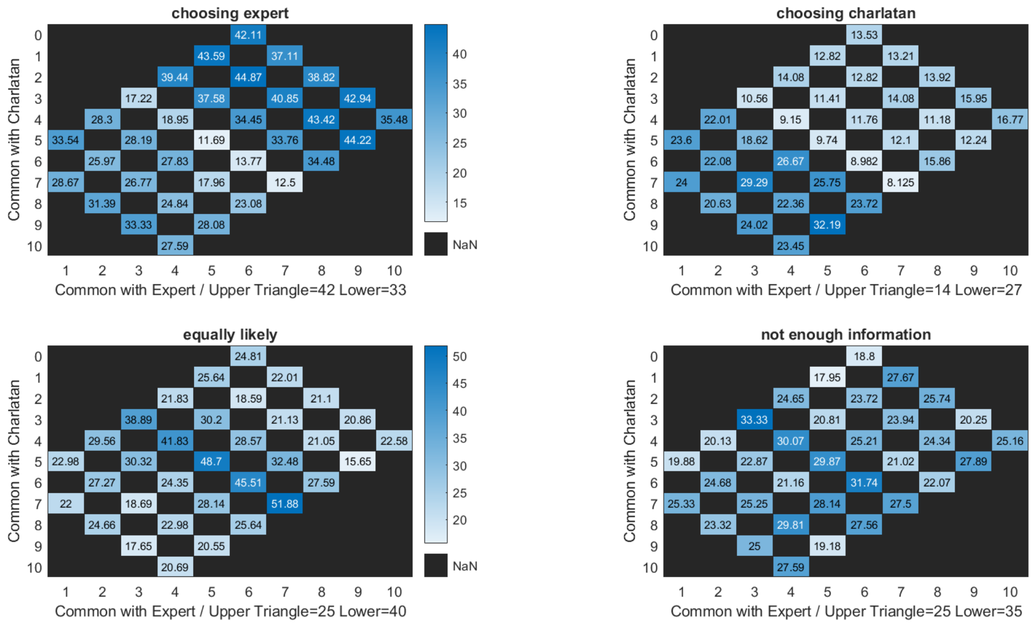


Figure 3: *Percentage of participants who gave each one of the four responses in each treatment. The horizontal axis shows 'Jim's number of correct answers and the vertical the number of common answers between 'Jim' and the populist advisor. Only text-treatments. The heading 'choosing expert' corresponds to selecting advisor M, 'choosing charlatan' to selecting advisor J.*

income of at least 45,000 Euros, we get the strongest performing subjects (figure 6). A full 50% of the subjects correctly identify the expert in the upper triangle, and 38% in the lower. In some cells the success rate is substantially higher, although note that the sample size is substantially smaller in these cases and the variability correspondingly higher. The flip side of this coin is that low-skill subjects do poorly. Restricting the sample to people with below university education certification and individual income less than 45,000 Euros while maintaining the age restriction under 65, we find that subjects identify the expert almost as (in)frequently as chance would suggest (figure 7). Moreover, note that in the lower triangle, which contains the counter-intuitive task of selecting the advisor with the least common answers, the populist is chosen more often than the expert.

3.4 Regressions

The analysis above points to some important findings: (i) Most laypeople either use the wrong heuristic or do not know how to evaluate the information in a task as simple as ours. (ii) The difference in common answers has significant impact on the expert-finding ability of subjects, indicating

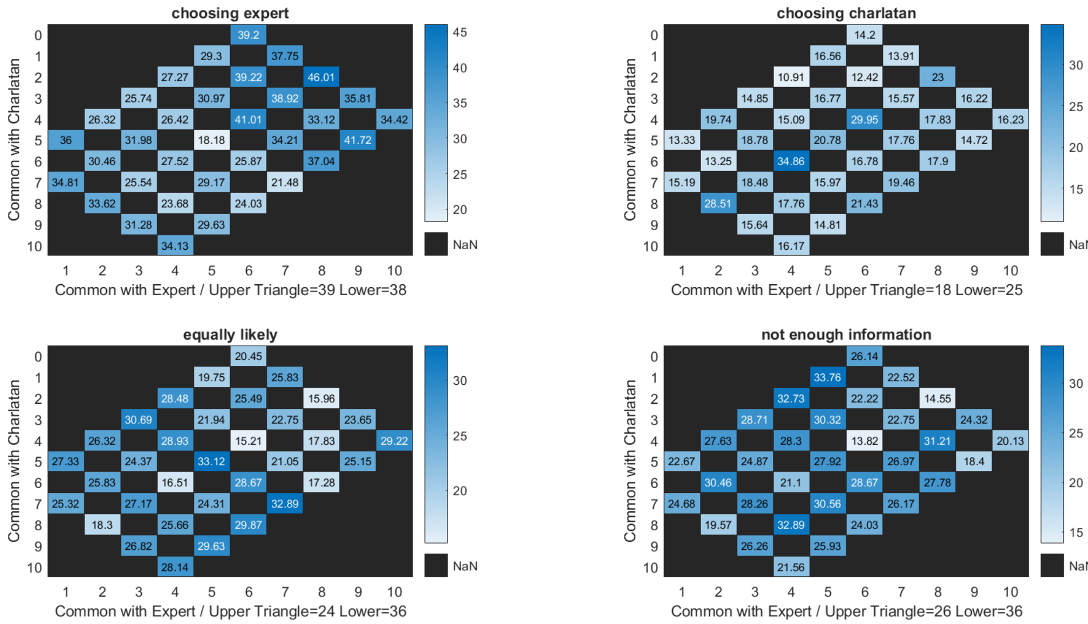


Figure 4: *Percentage of participants who gave each one of the four responses in each treatment. The horizontal axis shows 'Jim's number of correct answers and the vertical the number of common answers between 'Jim' and the populist advisor. Only table-treatments. The heading 'choosing expert' corresponds to selecting advisor M, 'choosing charlatan' to selecting advisor J.*

a problem where cognitive constraints bind easily. (iii) Demographic and social characteristics seem to be strongly correlated with expert-finding in our sample. We wish to test further these results through a series of regressions. We pool observations together from all treatments where $X \neq Y$ and we run several logit models with the dependent variable being a dummy for identifying the expert. It takes the value one if the subject correctly identifies the expert and zero for any other response. Our main explanatory variables include treatment characteristics, such as the number of correct answers and the common answers with the populist, and individual characteristics, such as the education background and the household income of participants. The full list of explanatory variables is in table 16 in appendix D. The regression results themselves are provided in table 5.

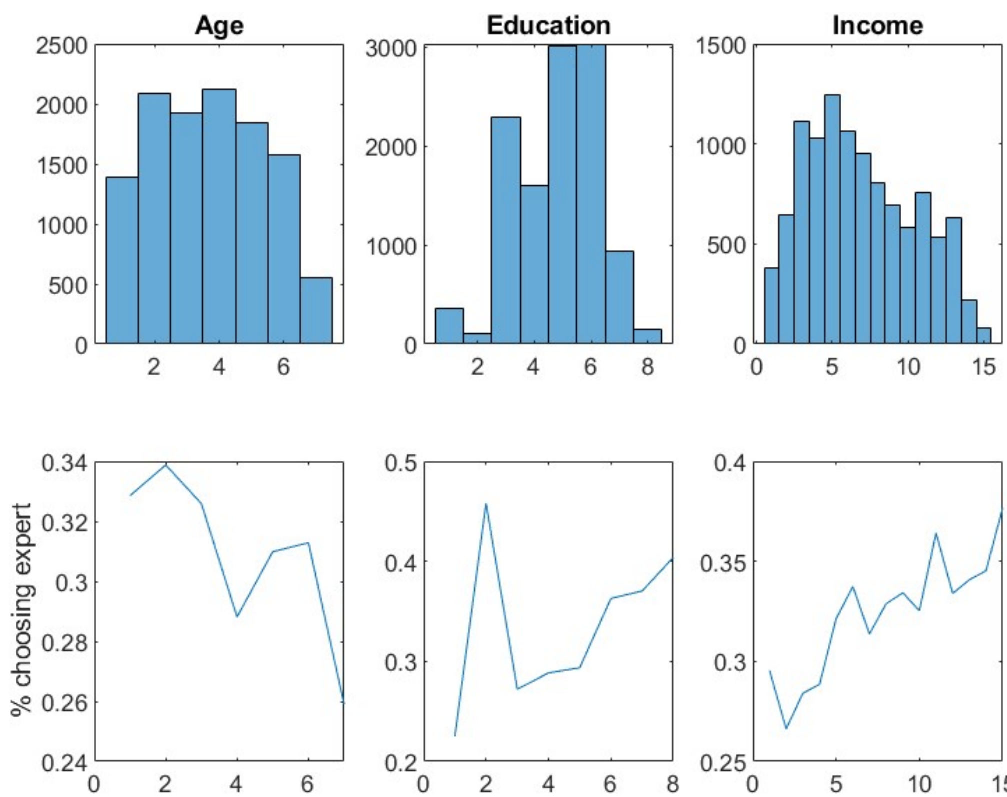


Figure 5: *Individual characteristics and expert-finding accuracy.*

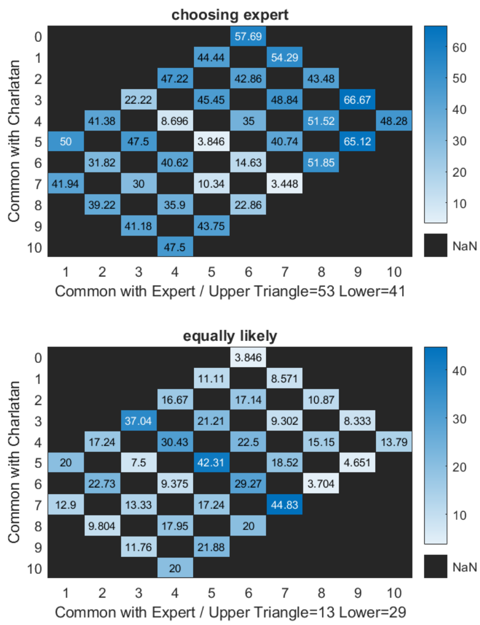


Figure 6: *Percentage of individuals giving a specific response when restricting the sample to vaccinated, top income and education quartile, below the age of 65. Responses from top-left to bottom-right: choosing expert - choosing populist - equally likely - not enough information. Pooled data across text and table presentation.*

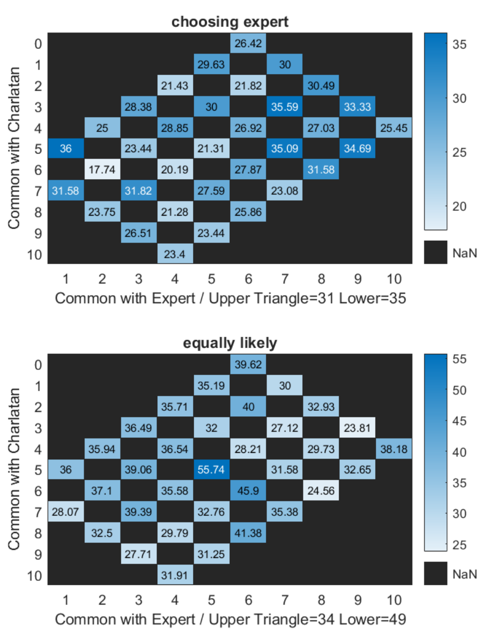


Figure 7: *Percentage of individuals giving a specific response when restricting the sample to bottom income and education quartile. Responses from top-left to bottom-right: choosing expert - choosing populist - equally likely - not enough information.*

	Dependent variable: FoundExpert					
	(1)	(2)	(3)	(4)	(5)	(6)
Diff	0.06***	0.05***	0.06***	0.06***	0.06***	0.06***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Table	-0.001	-0.003	-0.01	0.001	0.002	0.01
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
HI	0.18	0.30*	0.25	0.22	0.23	0.25
	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)	(0.17)
NCor		0.06***	0.04***	0.03***	0.03***	0.03***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
NCom			-0.05***	-0.05***	-0.05***	-0.05***
			(0.01)	(0.01)	(0.01)	(0.01)
Ed				0.10***	0.10***	0.08***
				(0.02)	(0.02)	(0.02)
Inc					0.01	0.01*
					(0.01)	(0.01)
Age						-0.05***
						(0.01)
Gndr						-0.26***
						(0.04)
Num						0.19***
						(0.05)
Vacc						0.07***
						(0.03)
SQS						-0.03*
						(0.02)
SQI						0.03***
						(0.01)
Prp						0.06**
						(0.03)
WSt						0.06***
						(0.01)
Constant	-0.93***	-1.21***	-0.84***	-1.24***	-1.18***	-1.24***
	(0.06)	(0.07)	(0.10)	(0.17)	(0.17)	(0.19)
Observations	9,870	9,870	9,870	9,192	9,192	8,161
Log Likelihood	-6,278.35	-6,255.93	-6,241.08	-5,770.71	-5,765.95	-5,148.47
Akaike Inf. Crit.	12,564.69	12,521.86	12,494.17	11,565.42	11,559.90	10,328.95

Note: $p^{*} < 0.1; p^{**} < 0.05; p^{***} < 0.01$

Table 5: *Results of logit regressions of the ability to identify the expert on a set of treatment and individual characteristics under different specifications. Data are taken from treatments where the number of common answers is different from the number of correct answers.*

The results support the evidence we highlighted previously. Subjects' performance improves with the difficulty of the problem. This happens when the number of correct answers (*NCor*) increases or the number of common answers with the populist (*NCom*) decreases. Both variables have a statistically significant coefficient and in the expected direction in all specifications. However, the presentation format (*Table*) does not seem to play a role, and neither does the inclusion of monetary incentives (*HI*) in order to increase subject attention. It is also noteworthy that the constant coefficient is statistically significant and negative at the 1% level, implying that most subject fail to identify the expert. Therefore, the findings from the logit regressions are in agreement with hypothesis *H1* and *H3*, but not *H2*.

In addition, we see that demographics play an important role. Education (*Ed*) and numeracy (*Num*) have particularly large positive coefficients while household income has little, if any, explanatory power. This means that the findings we reported in section 3.3 reflect primarily differences in human capital via education and cognitive ability. Gender is also an important determinant, with women performing significantly worse than men in this task. Coefficients for other demographics, such as age, property ownership (*Prp*) and work status (*WSt*), are also statistically significant, but with lower magnitude.

We also run regressions on other dependent variables. We created the variable *'Correct'*, which identifies correct answers through the full set of treatments. This means that it takes the value one if a subject gave the answer 'Advisor M' whenever $X \neq Y$ or 'Both equally likely' when $X = Y$. We run the same set of regressions as above using the new dependent variable and we obtained similar results (see appendix D). The coefficients of *NCor* are positive, while those of *NCom* are negatives. Both sets of coefficients are statistically significant. Education and numeracy have a positive effect and significant effect, while age and gender have a negative effect. Some of the other demographics though, such as vaccination status, either become insignificant or they exhibit a lower magnitude.

4 Conclusions

Identifying experts is important, from everyday situations to the functioning of democracy itself. We have shown, in a large-scale experiment with 12,000 subjects, that identifying an expert is not easy, even for the extreme case where there are just two possible advisors, the expert is always right, and complete information is provided regarding the populist's modus operandi. It is hard to imagine a simpler expert identification task than the text treatment, where finding the expert could be easier (but not trivial). Still less than 40% of the subjects identify the expert correctly.

High skills and human capital help - people belonging to the top quartile do a significantly better job of identifying the expert, and are fooled less often by the populist. Only individuals with high skills have a consistently higher than 50% chance of identifying the expert. Still their performance is far from perfect. Focussing on the bottom 20% of the sample in terms of skill, the frequency of correct answers approaches 25% which is the same as random. For hard cases (unfavourable parameter configurations), choosing the expert is more scarce than random choice, while the charlatan is more likely than not to beat the expert.

We also find that the exact evidence given to the subjects matters. People tend to choose the advisor they have the most in common with. When this aligns with the correct expert, they do relatively well and identify her frequently. However, subjects find it hard to follow the normative prescription when it involves choosing an advisor they have less answers in common with.

Note that the last two results combined lead to a deep issue in the field, which we call the paradox of advisor selection. Low skilled people find it harder to identify the true expert, especially when they have few correct answers and many in common with the populist. But these are exactly the people who find themselves most frequently in such a position, as having correct answers correlates (positively) with human capital.

The implications of these results for the functioning of parliamentarian democracies are consequential. We have abstracted from practically all factors that would add noise to decision making, such as the identity and natural characteristics of the advisors, political affiliations, rhetoric and charisma. People still find it hard to identify the true expert. Even when the expert is always right (which is obviously unusual), we inform the subjects about the populist's exact modus operandi, and we give perfect information about the number of their own correct answers, finding the expert is still hard.

Our findings provide a novel explanation why low-skilled citizens are susceptible to populism. It is not about preferences, incentives, or simple motivated reasoning. Some citizens just cannot distinguish populists from true experts, based purely on the objective correctness of the advice. While we have identified a fundamental mechanism that allows charlatans to win, and the target audience for this sort of populism, suggesting comprehensive solutions to overcome it is left for further research.

Appendix A

Treatment Design

The table below presents summary information for the design of all treatments we run. There are four answer options, which apply to all treatments: (i) 'Advisor J'. (ii) 'Advisor M'. (iii) 'Both equally likely'. (iv) 'Not enough information'.

No	Format	Incentives	NCor	NCom	CorAns
1	Text	No	1	5	Advisor M
2	Text	No	1	7	Advisor M
3	Text	No	2	4	Advisor M
4	Text	No	2	6	Advisor M
5	Text	No	2	8	Advisor M
6	Text	No	3	3	Both equally likely
7	Text	No	3	5	Advisor M
8	Text	No	3	7	Advisor M
9	Text	No	3	9	Advisor M
10	Text	No	4	2	Advisor M
11	Text	No	4	4	Both equally likely
12	Text	No	4	6	Advisor M
13	Text	No	4	8	Advisor M
14	Text	No	4	10	Advisor M
15	Text	No	5	1	Advisor M
16	Text	No	5	3	Advisor M
17	Text	No	5	5	Both equally likely
18	Text	No	5	7	Advisor M
19	Text	No	5	9	Advisor M
20	Text	No	6	0	Advisor M
21	Text	No	6	2	Advisor M
22	Text	No	6	4	Advisor M

–continued on next page–

–continued from the previous page–

No	Format	Incentives	NCor	NCom	CorAns
23	Text	No	6	6	Both equally likely
24	Text	No	6	8	Advisor M
25	Text	No	7	1	Advisor M
26	Text	No	7	3	Advisor M
27	Text	No	7	5	Advisor M
28	Text	No	7	7	Both equally likely
29	Text	No	8	2	Advisor M
30	Text	No	8	4	Advisor M
31	Text	No	8	6	Advisor M
32	Text	No	9	3	Advisor M
33	Text	No	9	5	Advisor M
34	Table	No	10	4	Advisor M
35	Table	No	1	5	Advisor M
36	Table	No	1	7	Advisor M
37	Table	No	2	4	Advisor M
38	Table	No	2	6	Advisor M
39	Table	No	2	8	Advisor M
40	Table	No	3	3	Both equally likely
41	Table	No	3	5	Advisor M
42	Table	No	3	7	Advisor M
43	Table	No	3	9	Advisor M
44	Table	No	4	2	Advisor M
45	Table	No	4	4	Both equally likely
46	Table	No	4	6	Advisor M
47	Table	No	4	8	Advisor M
48	Table	No	4	10	Advisor M
49	Table	No	5	1	Advisor M
50	Table	No	5	3	Advisor M
51	Table	No	5	5	Both equally likely

–continued on next page–

–continued from the previous page–

No	Format	Incentives	NCor	NCom	CorAns
52	Table	No	5	7	Advisor M
53	Table	No	5	9	Advisor M
54	Text	No	6	0	Advisor M
55	Table	No	6	2	Advisor M
56	Table	No	6	4	Advisor M
57	Table	No	6	6	Both equally likely
58	Table	No	6	8	Advisor M
59	Table	No	7	1	Advisor M
60	Table	No	7	3	Advisor M
61	Table	No	7	5	Advisor M
62	Table	No	7	7	Both equally likely
63	Table	No	8	2	Advisor M
64	Table	No	8	4	Advisor M
65	Table	No	8	6	Advisor M
66	Table	No	9	3	Advisor M
67	Table	No	9	5	Advisor M
68	Table	No	10	4	Advisor M
69	Text	Yes	3	5	Advisor M
70	Table	Yes	3	5	Advisor M

Table 6: *Number and key features of each treatment in our experiment. Format refers to whether the information on common answers was conveyed to participants via text or table. The variable incentives takes the value no if participants receive no monetary reward for a correct answer and yes if the treatment gave the possibility of monetary reward in case of correct answer. NCor presents the number of correct answers that Jim gave to the questionnaire, which is equal to the number of common answers with the expert advisor by construction. NCom is the number of common answers with the populist advisor. CorAns is the correct answer, which participants should have provided for the treatment.*

Appendix B

Demographic Representation

Our sample is representative of the UK general population. To achieve this we used the official statistics provided by ONS (Office for National Statistics) and the 2011 Census data.[8] The tables below report the sample quotas along with the actual fractions, which resulted from the collection.

Variable	Levels	No of Obs	Realised (%)	Quota (%)
Gender	Female	5789	50.3	50.6
	Male	5713	49.7	49.4
Age	18-24	1391	12.1	11.1
	25-34	2090	18.2	17.2
	35-44	1927	16.8	15.9
	45-54	2117	18.4	17.6
	55-64	1846	16.1	15.3
	65-74	1579	13.7	12.6
	75+	552	4.8	10.4
Ethnicity	White	9864	85.8	85.6
	Black African	301	2.6	2.3
	Black Caribbean	136	1.2	1.4
	Mixed	259	2.2	2.3
	East Asia	138	1.2	0.7
	Other Asia	104	0.9	1.5
	Pakistan-Bangladesh	347	3.0	2.8
	India	243	2.1	2.5
	Latin American	0	0.0	0.0
	Other group	110	1.0	1.0

Table 7: *Realised distribution of characteristics and target quotas for gender, age and ethnicity.*

[8] The original sources are available here (ONS) and here (2011 census).

Variable	Levels	No of Obs	Realised (%)
Education	No formal	360	3.1
	Primary school	107	0.9
	Secondary school	2293	19.9
	High school	1598	13.9
	College	3014	26.2
	Bachelors	3034	26.4
	Masters	945	8.2
	PhD	151	1.3
Marital status	Married	4833	42.0
	Living as married	1516	13.2
	Separated	1204	10.5
	Widowed	468	4.1
	Never married	3187	27.7
	Civil partnership	294	2.6
Work type	Professional	1744	15.2
	Manager	2586	22.5
	Clerical	1733	15.1
	Sales/Services	871	7.6
	Foreman/Supervisor	447	3.9
	Manual skilled	1334	11.6
	Semi-skilled/Unskilled manual	1568	13.6
	Other	759	6.6
	Never worked	460	4.0
Property	State-sponsored	767	6.9
	Renting	3370	30.4
	Mortgage	2678	24.1
	Own	4281	38.6
Household income (£)	Under 20,000	3167	29.5
	20,000-40,000	4059	37.8
	40,000-60,000	2039	19.0
	60,000-100,000	1164	10.8
	100,000 and above	300	2.8

Table 8: *Distribution of demographic characteristics for other collected categories.*

Appendix C

Additional Results

For each one of the treatments where $X \neq Y$ we test if the fraction of the participants who identify advisor ‘M’ and the expert is significantly less than 50%. The results of these tests are shown on Tables 9, 10, and 11 below, grouped by treatment presentation.

No	NCor	NCom	Chose Expert	Total	Estimate	Pvalue
1	1	5	108	310	0.34839	0
2	1	7	98	308	0.31818	0
3	2	4	85	311	0.27331	0
4	2	6	86	305	0.28197	0
5	2	8	149	458	0.32533	0
6	3	5	116	385	0.3013	0
7	3	7	100	381	0.26247	0
8	3	9	124	383	0.32376	0
9	4	2	101	306	0.33007	0
10	4	6	156	563	0.27709	0
11	4	8	76	313	0.24281	0
12	4	10	97	311	0.3119	0
13	5	1	114	312	0.36538	0
14	5	3	104	304	0.34211	0
15	5	7	72	311	0.23151	0
16	5	9	89	308	0.28896	0
17	6	0	125	308	0.40584	0.00056
18	6	2	130	308	0.42208	0.00365
19	6	4	171	455	0.37582	0
20	6	8	73	309	0.23625	0
21	7	1	116	310	0.37419	0.00001
22	7	3	123	309	0.39806	0.0002

–continued on next page–

–continued from the previous page–						
No	**NCor**	**NCom**	**Chose Expert**	**Total**	**Estimate**	**Pvalue**
23	7	5	105	308	0.34091	0
24	8	2	190	448	0.42411	0.00076
25	8	4	118	309	0.38188	0.00002
26	8	6	110	307	0.35831	0
27	9	3	123	311	0.3955	0.00014
28	9	5	133	310	0.42903	0.00723
29	10	4	108	309	0.34951	0

Table 9: *Tests on the fraction of participants who selected the expert being below 50%. Pooled data from text-treatments and table-treatments. NCor presents the number of correct answers that Jim gave to the questionnaire, which is equal to the number of common answers with the expert advisor by construction. NCom is the number of common answers with the populist advisor. Chose Expert gives the number of participant in the treatment who correctly selected advisor M as the expert. Total gives the total number of participants in the treatment. Estimate is the fraction of the participants who selected the expert and Pvalue is the associated p-value.*

No	NCor	NCom	Chose Expert	Total	Estimate	Pvalue
1	1	5	54	161	0.3354	0.00002
2	1	7	43	150	0.28667	0
3	2	4	45	159	0.28302	0
4	2	6	40	154	0.25974	0
5	2	8	70	223	0.3139	0
6	3	5	53	188	0.28191	0
7	3	7	53	198	0.26768	0
8	3	9	68	204	0.33333	0
9	4	2	56	142	0.39437	0.00733
10	4	6	96	345	0.27826	0
11	4	8	40	161	0.24845	0
12	4	10	40	144	0.27778	0
13	5	1	68	156	0.4359	0.06397
14	5	3	56	149	0.37584	0.00153
15	5	7	30	167	0.17964	0
16	5	9	41	146	0.28082	0
17	6	0	56	132	0.42424	0.04891
18	6	2	70	156	0.44872	0.11483
19	6	4	82	238	0.34454	0
20	6	8	36	156	0.23077	0
21	7	1	59	159	0.37107	0.00072
22	7	3	58	142	0.40845	0.01777
23	7	5	53	156	0.33974	0.00004
24	8	2	92	236	0.38983	0.00043
25	8	4	66	152	0.43421	0.0615

–continued on next page–

–continued from the previous page–						
No	**NCor**	**NCom**	**Chose Expert**	**Total**	**Estimate**	**Pvalue**
26	8	6	50	145	0.34483	0.00012
27	9	3	70	163	0.42945	0.04227
28	9	5	65	147	0.44218	0.09338
29	10	4	55	155	0.35484	0.00019

Table 10: *Tests on the fraction of participants who selected the expert being below 50%. Data from text-treatments only. NCor presents the number of correct answers that Jim gave to the questionnaire, which is equal to the number of common answers with the expert advisor by construction. NCom is the number of common answers with the populist advisor. Chose Expert gives the number of participant in the treatment who correctly selected advisor M as the expert. Total gives the total number of participants in the treatment. Estimate is the fraction of the participants who selected the expert and Pvalue is the associated p-value.*

No	NCor	NCom	Chose Expert	Total	Estimate	Pvalue
1	1	5	54	149	0.36242	0.00049
2	1	7	55	158	0.3481	0.00008
3	2	4	40	152	0.26316	0
4	2	6	46	151	0.30464	0
5	2	8	79	235	0.33617	0
6	3	5	63	197	0.3198	0
7	3	7	47	183	0.25683	0
8	3	9	56	179	0.31285	0
9	4	2	45	164	0.27439	0
10	4	6	60	218	0.27523	0
11	4	8	36	152	0.23684	0
12	4	10	57	167	0.34132	0.00003
13	5	1	46	156	0.29487	0
14	5	3	48	155	0.30968	0
15	5	7	42	144	0.29167	0
16	5	9	48	162	0.2963	0
17	6	0	69	176	0.39205	0.00257
18	6	2	60	152	0.39474	0.00584
19	6	4	89	217	0.41014	0.00487
20	6	8	37	153	0.24183	0
21	7	1	57	151	0.37748	0.00163
22	7	3	65	167	0.38922	0.00259
23	7	5	52	152	0.34211	0.00006
24	8	2	98	212	0.46226	0.15145
25	8	4	52	157	0.33121	0.00001

–continued on next page–

–continued from the previous page–

No	NCor	NCom	Chose Expert	Total	Estimate	Pvalue
26	8	6	60	162	0.37037	0.0006
27	9	3	53	148	0.35811	0.00035
28	9	5	68	163	0.41718	0.02069
29	10	4	53	154	0.34416	0.00007

Table 11: *Tests on the fraction of participants who selected the expert being below 50%. Data from table-treatments only. NCor presents the number of correct answers that Jim gave to the questionnaire, which is equal to the number of common answers with the expert advisor by construction. NCom is the number of common answers with the populist advisor. Chose Expert gives the number of participant in the treatment who correctly selected advisor M as the expert. Total gives the total number of participants in the treatment. Estimate is the fraction of the participants who selected the expert and Pvalue is the associated p-value.*

Table 12 below presents that results of the tests on whether the fraction of participants who select the populist in each treatment is significantly higher than 0%. We consider treatments where $X \neq Y$ and we pool data from table and text treatments together.

No	NCor	NCom	Chose Populist	Total	Estimate	Pvalue
1	1	5	58	310	0.1871	0
2	1	7	60	308	0.19481	0
3	2	4	65	311	209	0
4	2	6	54	305	0.17705	0
5	2	8	113	458	0.24672	0
6	3	5	72	385	0.18701	0
7	3	7	92	381	0.24147	0
8	3	9	77	383	0.20104	0
9	4	2	38	306	0.12418	0
10	4	6	168	563	0.2984	0
11	4	8	63	313	0.20128	0
12	4	10	61	311	0.19614	0
13	5	1	46	312	0.14744	0
14	5	3	43	304	0.14145	0
15	5	7	66	311	0.21222	0
16	5	9	71	308	0.23052	0
17	6	0	43	308	0.13961	0
18	6	2	39	308	0.12662	0
19	6	4	93	455	0.2044	0
20	6	8	70	309	0.22654	0
21	7	1	42	310	0.13548	0
22	7	3	46	309	0.14887	0
23	7	5	46	308	0.14935	0
24	8	2	82	448	0.18304	0

–continued on next page–

–continued from the previous page–						
No	**NCor**	**NCom**	**Chose Populist**	**Total**	**Estimate**	**Pvalue**
25	8	4	45	309	0.14563	0
26	8	6	52	307	0.16938	0
27	9	3	50	311	0.16077	0
28	9	5	42	310	0.13548	0
29	10	4	51	309	0.16505	0

Table 12: *Tests on the fraction of participants who selected the populist being above 0%. Pooled data from treatments where number of commons answers with the two advisors is not the same. NCor presents the number of correct answers that Jim gave to the questionnaire, which is equal to the number of common answers with the expert advisor by construction. NCom is the number of common answers with the populist advisor. Chose Populist gives the number of participant in the treatment who selected advisor J as the expert. Total gives the total number of participants in the treatment. Estimate is the fraction of the participants who selected the expert and Pvalue is the associated p-value.*

No	NCor	NCom	Table %	Text %	Pvalue
1	1	5	0.36242	0.3354	0.70437
2	1	7	0.3481	0.28667	0.30084
3	2	4	0.26316	0.28302	0.79056
4	2	6	0.30464	0.25974	0.4569
5	2	8	0.33617	0.3139	0.68278
6	3	5	0.3198	0.28191	0.48475
7	3	7	0.25683	0.26768	0.90142
8	3	9	0.31285	0.33333	0.75047
9	4	2	0.27439	0.39437	0.03539
10	4	6	0.27523	0.27826	1.00
11	4	8	0.23684	0.24845	0.91444
12	4	10	0.34132	0.27778	0.27866
13	5	1	0.29487	0.4359	0.01355
14	5	3	0.30968	0.37584	0.27368
15	5	7	0.29167	0.17964	0.02776
16	5	9	0.2963	0.28082	0.86243
17	6	0	0.39205	0.42424	0.65112
18	6	2	0.39474	0.44872	0.39888
19	6	4	0.41014	0.34454	0.17826
20	6	8	0.24183	0.23077	0.92438
21	7	1	0.37748	0.37107	1.00
22	7	3	0.38922	0.40845	0.82
23	7	5	0.34211	0.33974	1
24	8	2	0.46226	0.38983	0.14618
25	8	4	0.33121	0.43421	0.08082
26	8	6	0.37037	0.34483	0.72878
27	9	3	0.35811	0.42945	0.24244

–continued on next page–

–continued from the previous page–

No	NCor	NCom	Table %	Text %	Pvalue
28	9	5	0.41718	0.44218	0.74204
29	10	4	0.34416	0.35484	0.93814

Table 13: *Tests on the difference of fraction of participants who selected the expert between text and table treatments within the same configuration of correct and common answers with $X \neq Y$. NCor presents the number of correct answers that Jim gave to the questionnaire, which is equal to the number of common answers with the expert advisor by construction. NCom is the number of common answers with the populist advisor. Table % is the fraction of participants that answered correctly in the table treatments and Text % is the fraction of participants that answered correctly in the text treatments. Pvalue is the associated p-value for the test.*

No	NCor	NCom	Table %	Text %	Pvalue
1	3	3	0.30693	0.38889	0.11557
2	4	4	0.29299	0.4183	0.02877
3	5	5	0.33117	0.48701	0.00769
4	6	6	0.28671	0.45509	0.00338
5	7	7	0.32886	0.51875	0.00113

Table 14: *Tests on the difference of fraction of participants who selected 'both equally likely' between text and table treatments within the same configuration of correct and common answers with $X = Y$. NCor presents the number of correct answers that Jim gave to the questionnaire, which is equal to the number of common answers with the expert advisor by construction. NCom is the number of common answers with the populist advisor. Table % is the fraction of participants that answered correctly in the table treatments and Text % is the fraction of participants that answered correctly in the text treatments. Pvalue is the associated p-value for the test.*

No	NCor	NCom	Table %	Text %	Pvalue
1	3	3	0.28713	0.33333	0.3872
2	4	4	0.28662	0.30065	0.88356
3	5	5	0.27922	0.2987	0.80149
4	6	6	0.28671	0.31737	0.64454
5	7	7	0.26174	0.2750	0.89321

Table 15: *Tests on the difference of fraction of participants who selected 'not enough information' between text and table treatments within the same configuration of correct and common answers with $X = Y$. NCor presents the number of correct answers that Jim gave to the questionnaire, which is equal to the number of common answers with the expert advisor by construction. NCom is the number of common answers with the populist advisor. Table % is the fraction of participants that answered correctly in the table treatments and Text % is the fraction of participants that answered correctly in the text treatments. Pvalue is the associated p-value for the test.*

Appendix D

Results from Regressions

Variable	**Description**
Diff	Absolute difference between number of correct answers and common answers with the populist.
Table	Dummy variable that takes value 1 if treatment is presented as a table.
HI	Dummy variable that takes value 1 if treatment provided monetary incentives for correct answers to participants.
NCor	Number of correct answers.
NCom	Number of common answers with the populist.
Ed	Education background.
Inc	Level of household income.
Age	Participant's age.
Gndr	Participant's gender
Num	Participant's numeracy.
Vacc	Participant's vaccination status.
SQS	Variable that categorises whether participant seeks quality sources.
SQI	Variable that categorises whether participant seeks quality information.
Prp	Property ownership of participant.
WSt	Work status of participant.

Table 16: *List of explanatory variables used in regression analysis.*

	Dependent variable: Correct Answer					
	(1)	(2)	(3)	(4)	(5)	(6)
Diff	0.01	0.004	0.004	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Table	-0.09**	-0.09**	-0.09**	-0.09**	-0.09**	-0.09**
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
HI	0.04	0.17	0.13	0.10	0.10	0.13
	(0.15)	(0.15)	(0.15)	(0.16)	(0.16)	(0.17)
NCor		0.06***	0.04***	0.04***	0.04***	0.04***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
NCom			-0.04***	-0.04***	-0.04***	-0.04***
			(0.01)	(0.01)	(0.01)	(0.01)
Ed				0.08***	0.08***	0.06***
				(0.02)	(0.02)	(0.02)
Inc				0.02**	0.02**	0.002
				(0.01)	(0.01)	(0.01)
Age				-0.02*	-0.03**	-0.04***
				(0.01)	(0.01)	(0.01)
Gndr				-0.17***	-0.17***	-0.18***
				(0.04)	(0.04)	(0.04)
Num				0.13***	0.12**	0.08
				(0.05)	(0.05)	(0.05)
Vacc				0.06**	0.05**	0.05*
				(0.02)	(0.02)	(0.03)
SQS					-0.03	-0.03
					(0.02)	(0.02)
SQI					0.02**	0.03***
					(0.01)	(0.01)
Prp						0.03
						(0.02)
WSt						0.04***
						(0.01)
Constant	-0.65***	-0.95***	-0.64***	-1.05***	-1.02***	-1.03***
	(0.04)	(0.06)	(0.08)	(0.15)	(0.15)	(0.17)
Observations	11,489	11,489	11,489	10,716	10,716	9,520
Log Likelihood	-7,369.02	-7,343.38	-7,330.69	-6,819.61	-6,816.73	-6,097.01
Akaike Inf. Crit.	14,746.04	14,696.76	14,673.38	13,663.21	13,661.46	12,226.03

Note: $p^{*} < 0.1; p^{**} < 0.05; p^{***} < 0.01$

Table 17: *Results of logit regressions of correct responses on a set of treatment and individual characteristics under different specifications. All treatments.*

	Dependent variable: Selected Populist					
	(1)	(2)	(3)	(4)	(5)	(6)
Diff	-0.03*	-0.02	-0.03*	-0.04**	-0.04**	-0.04**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Table	-0.03	-0.03	-0.03	-0.03	-0.03	-0.06
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)
HI	0.01	-0.10	-0.04	-0.13	-0.13	-0.13
	(0.19)	(0.19)	(0.19)	(0.20)	(0.20)	(0.21)
NCor		-0.06***	-0.03**	-0.03**	-0.03**	-0.03**
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
NCom			0.06***	0.06***	0.06***	0.06***
			(0.01)	(0.01)	(0.01)	(0.01)
Ed				-0.02	-0.02	-0.05**
				(0.02)	(0.02)	(0.02)
Inc				-0.002	-0.003	-0.01
				(0.01)	(0.01)	(0.01)
Age				-0.05***	-0.05***	-0.06***
				(0.02)	(0.02)	(0.02)
Gndr				-0.06	-0.06	-0.04
				(0.05)	(0.05)	(0.06)
Num				-0.12**	-0.12*	-0.10
				(0.06)	(0.06)	(0.07)
Vacc				0.01	0.01	-0.02
				(0.03)	(0.03)	(0.03)
SQS					0.02	0.01
					(0.03)	(0.03)
SQI					-0.01	-0.02
					(0.01)	(0.01)
Prp						0.03
						(0.03)
WSt						-0.002
						(0.02)
Constant	-1.34***	-1.08***	-1.52***	-0.98***	-0.99***	-0.78***
	(0.07)	(0.09)	(0.12)	(0.20)	(0.20)	(0.22)
Observations	9,870	9,870	9,870	9,192	9,192	8,161
Log Likelihood	-4,757.41	-4,744.23	-4,729.08	-4,422.81	-4,422.32	-3,930.78
Akaike Inf. Crit.	9,522.82	9,498.45	9,470.16	8,869.63	8,872.64	7,893.55

Note: $p^{*} < 0.1; p^{**} < 0.05; p^{***} < 0.01$

Table 18: *Results of logit regressions of selecting the populist on a set of treatment and individual characteristics under different specifications. All treatments.*

	Dependent variable: More Info					
	(1)	(2)	(3)	(4)	(5)	(6)
Diff	-0.04***	-0.04***	-0.04***	-0.04***	-0.04***	-0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Table	0.03	0.03	0.03	0.03	0.02	0.04
	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)
HI	-0.40**	-0.42**	-0.42**	-0.40**	-0.40**	-0.42**
	(0.18)	(0.19)	(0.19)	(0.19)	(0.19)	(0.20)
NCor		-0.01	-0.01	-0.01	-0.01	-0.01
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
NCom			0.004	0.01	0.01	0.004
			(0.01)	(0.01)	(0.01)	(0.01)
Ed				0.03*	0.03	0.04**
				(0.02)	(0.02)	(0.02)
Inc				-0.01*	-0.01*	-0.001
				(0.01)	(0.01)	(0.01)
Age				0.06***	0.04***	0.06***
				(0.01)	(0.01)	(0.02)
Gndr				0.22***	0.21***	0.20***
				(0.05)	(0.05)	(0.05)
Num				0.56***	0.47***	0.59***
				(0.05)	(0.06)	(0.06)
Vacc				-0.07***	-0.11***	-0.08***
				(0.03)	(0.03)	(0.03)
SQS					0.09***	0.10***
					(0.02)	(0.02)
SQI					0.04***	0.04***
					(0.01)	(0.01)
Prp					-0.05*	
					(0.03)	
WSt					-0.01	
					(0.01)	
Constant	-0.97***	-0.91***	-0.94***	-1.78***	-1.65***	-1.89***
	(0.05)	(0.07)	(0.09)	(0.17)	(0.17)	(0.19)
Observations	11,489	11,489	11,489	10,716	10,716	9,520
Log Likelihood	-6,447.72	-6,446.85	-6,446.76	-5,832.37	-5,814.48	-5,035.87
Akaike Inf. Crit.	12,903.44	12,903.69	12,905.51	11,688.74	11,656.95	10,103.74

Note: $p^{*} < 0.1; p^{**} < 0.05; p^{***} < 0.01$

Table 19: *Results of logit regressions of answering that more information is required to find the expert on a set of treatment and individual characteristics under different specifications. All treatments.*

References

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.

Alysandratos, T., A. Boukouras, S. Georganas, and Z. Maniadis (2023): "'Identify the Expert': An Experimental Study in Economic Advice," *Available at SSRN 4369623.*

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," Tech. rep., ECONtribute Discussion Paper.

Chopra, F., I. Haaland, and C. Roth (2022): "Do people demand fact-checked news? Evidence from US Democrats," *Journal of Public Economics*, 205, 104549.

Kahneman, D. (2011): *Thinking, Fast and Slow*, Macmillan.

Oliveros, S. and F. Várdy (2015): "Demand for slant: How abstention shapes voters' choice of news media," *The Economic Journal*, 125, 1327–1368.

Thaler, M. (2021): "The supply of motivated beliefs," *arXiv preprint arXiv:2111.06062.*