

Discussion Paper Series

IZA DP No. 18453

March 2026

Expertise and Prediction Accuracy

Elisabeth Grewenig

Kreditanstalt für Wiederaufbau (KfW)

Klaus Gründler

University of Kassel, ifo Institute
and CESifo

Philipp Lorgetporer

ifo Institute, CESifo, Technical
University of Munich and IZA@LISER

Niklas Potrafke

ifo Institute, University of
Munich (LMU) and CESifo

Katharina Werner

ifo Institute, CESifo and
Business School Pforzheim

Helen Zeidler

Technical University of Munich

The IZA Discussion Paper Series (ISSN: 2365-9793) ("Series") is the primary platform for disseminating research produced within the framework of the IZA@LISER Network, an unincorporated international network of labour economists coordinated by the Luxembourg Institute of Socio-Economic Research (LISER). The Series is operated by LISER, a Luxembourg public establishment (établissement public) registered with the Luxembourg Business Registers under number J57, with its registered office at 11, Porte des Sciences, 4366 Esch-sur-Alzette, Grand Duchy of Luxembourg.

Any opinions expressed in this Series are solely those of the author(s). LISER accepts no responsibility or liability for the content of the contributions published herein. LISER adheres to the European Code of Conduct for Research Integrity. Contributions published in this Series present preliminary work intended to foster academic debate. They may be revised, are not definitive, and should be cited accordingly. Copyright remains with the author(s) unless otherwise indicated.



Expertise and Prediction Accuracy*

Abstract

Public support for policy interventions depends on citizens' beliefs about their likely effects. We examine how individuals form such beliefs by studying their predictions of experimental outcomes in a policy-relevant setting, and why their predictions differ from expert benchmarks. We elicit forecasts from 127 professional economists and a representative sample of 6,200 German households about a large-scale behavioral experiment on education policy (N = 3, 133). Nonexperts predict both average outcomes and treatment effects far less accurately than experts. Prediction accuracy improves with calibrated priors, self-reported effort, and the use of structured reasoning, but remains well below expert levels. We show that scalable design features, including the provision of well-calibrated numerical anchors and monetary incentives to rise effort, improve non-expert predictions, with effects comparable in magnitude to tertiary education or structured reasoning. Our findings have important implications for bridging the 'expertise gap' in public discourse.

JEL classification

A11, D83, H52, I22

Keywords

expert forecasts, lay predictions, belief formation, expertise gap, policy support, behavioral experiments

Corresponding author

Helen Zeidler

helen.zeidler@tum.de

* We would like to thank Stefano DellaVigna and Eva Vivalt for highly valuable comments. We would also like to thank participants at the 4th Workshop on Field Experiments in Economics and Business at Technical University of Munich (TUM) for insightful feedback and discussions. AEA Registry: This experiment has been pre-registered at the AEA RCT Registry under RCT ID AEARCTR0006674.

1 Introduction

In recent years, economists have increasingly turned to *forecasting*—asking individuals to predict the results of experiments—as a tool for improving research design, benchmarking priors, and evaluating the credibility and external validity of empirical findings (DellaVigna et al., 2019, 2020). Forecast elicitation is now central to experimental economics, replication projects, and the study of how scientific knowledge is produced and communicated (Camerer et al., 2016; DellaVigna and Pope, 2018). A growing body of work shows that academic experts are often remarkably accurate in forecasting treatment effects and behavioral responses. At the same time, persistent and substantial gaps between the predictions of experts and those of non-experts have been documented across a wide range of settings (DellaVigna and Pope, 2018; Chu et al., 2024).

Understanding expert-lay differentials in prediction accuracy is important for both economic science and policy. Many policy interventions rely on public support, which in turn depends on citizens’ beliefs about their likely effects. When public perceptions systematically diverge from expert benchmarks, support for otherwise effective, evidence-based measures may be weakened. Despite their importance, however, the mechanisms generating prediction gaps between experts and non-experts remain poorly understood. Do these gaps reflect immutable differences in knowledge and experience, or do they arise from how beliefs are formed and elicited—and therefore admit scope for improvement?

In this paper, we study the origins of divergences in forecasts between experts and non-experts and explore whether predictive accuracy among non-experts can be improved through elicitation design features. We conduct a pre-registered survey experiment centered on public education spending, a highly salient and policy-relevant domain in which individuals often hold strong intuitions but imperfect information. Our approach combines two parallel surveys in Germany: one with 127 economics professors (*experts*) and another with a nationally representative sample of roughly 6,200 adults (*non-experts*). Respondents in both surveys are asked to predict the outcomes of a separate behavioral information experiment among the German adult population ($N = 3,133$), which elicits preferences for education spending across three intervention groups. Participants are either informed about actual per-student spending, given the opportunity to acquire this information, or remain uninformed in a pure control group.

In our main prediction task, we ask experts and non-experts to forecast both base-

line support for increased education spending and the effects of experimental interventions. To explore whether non-expert forecasts can be improved, we embed a prediction elicitation experiment that introduces three features associated with expert prediction into the non-expert environment: monetary incentives for accuracy, numerical anchoring based on prior survey benchmarks, and prior experience with the decision context. This design allows us to address three central questions. First, how large is the gap in predictive accuracy between experts and non-experts, and how does accuracy vary across tasks and within each group? Second, what cognitive strategies shape prediction formation, and to what extent do differences in priors, reasoning strategies, and engagement explain forecasting performance? Third, can simple, scalable interventions reduce the expert–lay gap, and how do their effects compare to persistent heterogeneity driven by socio-economic characteristics and reasoning strategies?

We report three main results. The first main result is that experts consistently outperform non-experts in predicting both average support for increased education spending and the effects of informational treatments. Experts’ forecasts closely align with realized outcomes (65% vs. 77%), whereas non-experts substantially underpredict baseline support (47%). Experts also more accurately anticipate treatment effects across both the information provision and acquisition conditions. While non-experts with higher educational attainment perform better, they remain significantly less accurate than experts, suggesting that observable characteristics only partially mitigate the expert–lay gap. Importantly, experts’ advantage is largely unrelated to specialization in education economics, indicating that it reflects general forecasting skill rather than narrow domain expertise.

The second main result is that the gap between expert and non-expert predictions reflects systematic differences in how information is processed, not just in knowledge or experience. Non-experts’ forecasts are heavily shaped by their prior beliefs. Respondents with more accurate priors produce predictions closer to realized outcomes, whereas miscalibrated priors directly translate into larger errors. Among non-experts, self-reported sources of information, including familiarity with related studies or online searches, are not consistently associated with higher accuracy, suggesting that additional information alone is insufficient when it is not effectively integrated. By contrast, self-reported effort strongly predicts better performance, indicating that careful reflection and structured processing, rather than mere exposure to information, are key. Open-ended responses reveal that non-experts predominantly rely on intuition or

gut feeling (57%), whereas experts more frequently apply deductive reasoning (48%) and knowledge-based approaches (27%). Non-experts who report using structured reasoning achieve substantially higher accuracy, although even their forecasts remain below expert levels. Taken together, these findings show that prediction errors among non-experts arise from a combination of poorly calibrated priors, ineffective use of information, and reliance on intuitive strategies.

The third main result is that interventions targeting key sources of non-expert prediction errors—poorly calibrated priors, limited cognitive effort, and lack of prior experience—can substantially improve forecasting accuracy. Both monetary incentives and numerical anchors raise accuracy, with anchoring proving particularly effective. Providing respondents with a plausible benchmark reduces prediction errors by an amount comparable to, or exceeding, the gains associated with deductive reasoning or higher education. In the control condition, anchoring raises non-experts’ forecasts from 47% to 53% (true value: 77%), narrowing—but not closing—the gap with experts (65%). These findings demonstrate that simple, scalable design interventions can improve non-expert predictions as much as, if not more than, persistent individual traits. At the same time, even the most effective interventions leave a substantial gap, highlighting enduring limits in non-expert quantitative judgment.

Our findings carry important implications for how economic research is interpreted and communicated. Building on prior evidence of expert–lay prediction gaps, we document substantial divergence between expert and non-expert beliefs about experimental outcomes. These gaps can weaken public support for evidence-based policies across multiple domains. Importantly, however, much of the divergence reflects differences in prior knowledge, cognitive strategy, and task engagement rather than specialization or expertise *per se*. Consequently, non-expert prediction errors can be reduced through interventions that shape reasoning and information processing, narrowing the gap with expert predictions. These results highlight the potential—and the need—for future research on interventions that systematically improve how non-experts interpret, process, and act on economic information.

The main contribution of this paper is to unpack the mechanisms behind expert–lay prediction gaps and show how these gaps can be narrowed through structured interventions. As such, our paper contributes to several strands of literature on belief formation, forecasting, and expert–lay differences in economic reasoning. Most directly, it adds to a growing body of work studying how well experts and non-experts

predict empirical research findings and policy-relevant outcomes (Camerer et al., 2016; DellaVigna et al., 2019, 2020; Banuri et al., 2019; Hjort et al., 2021; Chu et al., 2024). While this literature documents systematic expert advantages, it remains largely silent on the mechanisms generating expert-lay prediction gaps. We contribute by studying expert and non-expert forecasts in a common experimental environment and by jointly analyzing accuracy, elicitation design, and self-reported prediction strategies. This integrated approach allows us to move beyond documenting gaps to identifying the *channels*—knowledge, cognitive strategy, and task engagement—through which expert-lay prediction gaps arise.

We also contribute to the literature on expectation formation and belief heterogeneity among non-experts (Manski, 2004; DellaVigna, 2009; Coibion and Gorodnichenko, 2015; Roth and Wohlfart, 2018; Dräger et al., 2025; Gründler et al., 2026). Existing work emphasizes differences in information, priors, and numeracy, but provides limited evidence on how these factors interact with forecasting performance in complex, policy-relevant prediction tasks. We show that educational attainment among non-experts is strongly associated with forecast accuracy, yet does not eliminate the gap relative to experts. At the same time, we find little role for narrow domain specialization among experts. Together, these findings suggest that general forecasting skills and structured reasoning, rather than topic-specific expertise, are central to expert performance, contributing to clarify an open question in prior work on expert accuracy.

Our paper also connects to research on belief elicitation and survey design (Delavande et al., 2011; Prior et al., 2015; Grewenig et al., 2022). While earlier studies document mixed effects of incentives and framing on belief accuracy, we provide experimental evidence that simple design features, particularly numerical anchoring, substantially improve non-expert predictions. Importantly, we benchmark these design effects against individual characteristics and cognitive strategies, showing that elicitation choices can matter as much as education or reasoning style. This comparison highlights that the elicitation design can be a practical and scalable tool for improving belief measurement.

Finally, we contribute to the emerging literature that links cognitive processes to economic expectations (e.g., DellaVigna, 2018). By classifying open-ended responses, we provide direct evidence on the strategies individuals use when forming forecasts. We show that non-experts disproportionately rely on intuitive reasoning, which is strongly associated with lower accuracy, whereas experts predominantly use deduc-

tive and knowledge-based approaches that yield consistently high performance. This evidence complements existing models of belief formation by identifying prediction strategy itself, rather than information alone, as a key mechanism underlying expert-lay differences.

2 Experimental design and data

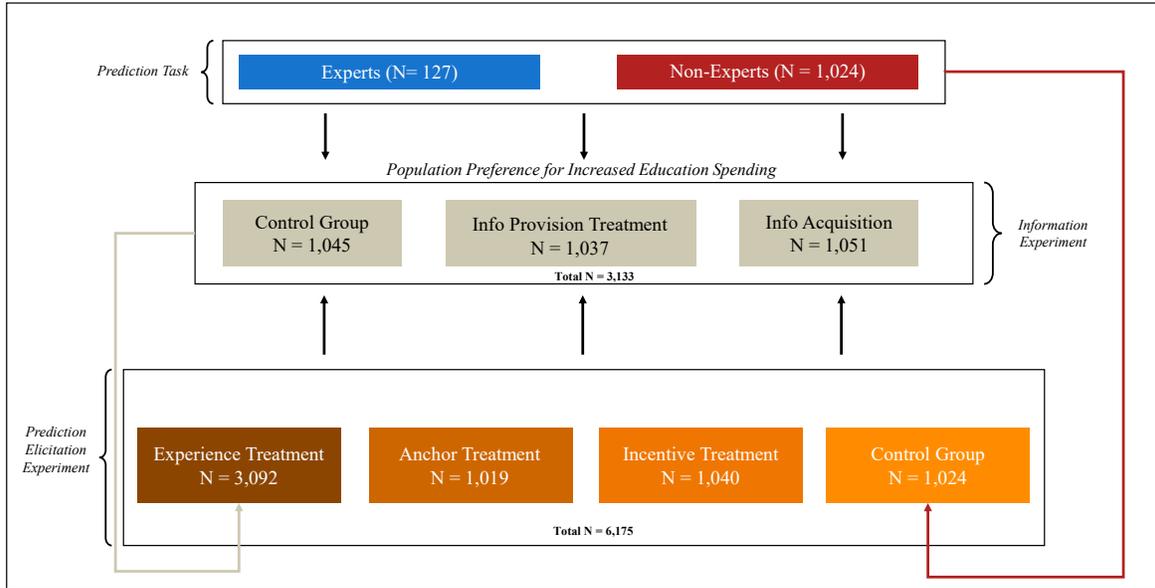
Our goal is to explore differentials in how experts and non-experts predict policy-relevant research outcomes and to understand their origins. To this end, we first conduct a behavioral information experiment and then ask experts and non-experts to predict the outcome of this experiment. We employ several experimental interventions to study whether non-expert predictions can be improved by specific features of the elicitation design. Figure (1) provides an overview of our experimental setup. We now describe the details of this setting and the samples that we are using in our project.

2.1 The information experiment

As a basis for the prediction experiment, we first implemented an information experiment among a representative sample ($N = 3,133$) of the German population (*non-experts*). The experiment is inspired by [Lergetporer et al. \(2018\)](#) and involves the elicitation of preferences for public school spending. Participants in the (uninformed) control group ($N = 1,045$) are asked about their support for increases or decreases in average school spending per student and year in Germany on a five-point Likert scale.

The information experiment includes two randomly assigned treatment conditions. The first group (*information provision*, $N = 1,037$) is informed about the actual per-student annual spending on public education in Germany (€7,300). The second group (*information acquisition*, $N = 1,051$) can acquire the identical information by clicking on a link. We focus on government spending on education, as this spending category accounts for a large share of public sector budgets around the world. This highlights the policy relevance of understanding belief formation in this area, as citizens' support for education spending is a critical factor in shaping public budgets.

Figure 1 Experimental Setup



Note: The figure shows the experimental setup. We conduct a survey among a large sample of around 6,200 participants representative for the population (non-experts) in Germany. We ask survey participants to predict outcomes of an information experiment. The information experiment involves eliciting population preferences towards increased education spending (N = 3,133). Before answering the outcome question, participants in the information experiment are randomly assigned to one of three groups: participants in one treatment group (information provision, N = 1,037) are informed about current education spending levels (the average public school spending per student in the year 2020 was €7,300). Participants in the other treatment group (information acquisition, N = 1,051) have the opportunity to retrieve information about current spending levels. Participants in the control group (N = 1,045) do not receive any information. In the prediction task (N = 1,024), participants first learn about the study design of the information experiment. Subsequently, we measure their predictions for support to increase education expenditure in the control group, information provision and information acquisition treatment group. We also ask 127 economics experts to predict the outcomes in the three groups. To examine how the elicitation design influences predictions, we run a prediction elicitation experiment among non-experts (N = 6,175) and randomize three design features at the individual level. First, we randomly incentivize participants (N=1,040) with one additional Euro if answers to the prediction experiment are correct. Second, we randomly provide participants (N=1,019) with an anchor by informing them about the control group support for increased public school spending in a similar survey conducted in 2019. Third, we assign participants of the information experiment to an experience condition (N=3,092). These respondents participated in the information experiment first before being asked to predict the outcomes of the experiment. We compare predictions of the prediction elicitation experiment to the predictions made by participants in the prediction task (control group, N = 1,024).

2.2 The prediction task

We ask participants in an expert ($N = 127$) and a non-expert ($N = 1,024$) sample to predict the outcomes of the information experiment. Participants in the prediction task are first informed about the design of the information experiment, including the exact wording used in each experimental arm. They also receive detailed information on the outcome questions and the corresponding answer categories.

As their primary task, participants were asked to forecast the outcomes of the information experiment by answering the question “*What do you think which share of respondents answers ‘strongly increase’ or ‘increase’? Keep in mind that the questions were posed in a survey among people aged between 18 and 69 years old*”. For each of the three experimental groups, participants could then indicate their predictions in an open field, allowing for answers between 0 and 100. We further elicit participants’ confidence with their predictions on a 7-point-scale.

Prediction formation We use open-ended text questions to study how participants form their predictions. Specifically, respondents are asked to explain, in their own words, how they arrived at their forecast using a free-text response. This elicitation approach allows us to capture individuals’ first-order considerations and underlying cognitive strategies without priming specific modes of reasoning (see, e.g., [Stantcheva, 2021](#); [Andre et al., 2022](#); [Dräger et al., 2025](#)). In addition, we use responses to structured questions on the informational inputs participants rely on, including whether they engage in online search, exert particular effort to investigate the task, or draw on knowledge of related or prior studies.

Participant characteristics The survey further collects detailed information on participants’ socio-economic characteristics. This includes standard demographics such as age, gender, and education, as well as measures of income, employment status, household composition, and political and economic attitudes. These data allow us to characterize heterogeneity in prediction accuracy across socio-economic groups and to examine how individual background characteristics correlate with forecasting performance.

2.3 The prediction elicitation experiment

To examine how the elicitation design influences predictions of non-experts, we run a prediction elicitation experiment ($N = 6,175$) as part of the prediction task of non-experts. This experiment randomizes three design features at the individual level.

Incentive treatment As a first intervention, we randomly incentivize participants ($N = 1,040$) with the prospect of an additional payoff of one Euro if answers to the prediction elicitation experiment are correct.¹ This design feature is intended to increase respondents' cognitive effort and allows us to test whether greater engagement improves prediction accuracy, consistent with prior work on incentivized belief elicitation (e.g., [Schlag et al., 2015](#); [Trautmann and van de Kuilen, 2015](#)).

Anchor treatment Our second intervention provides a randomly chosen subset of participants ($N = 1,019$) with an anchor, informing them about the control group support for increased public school spending in a similar survey conducted in 2019.² By providing an anchor as a salient reference-point, we can study to what extent reduced uncertainty, disciplined guesswork, and a closer calibration of predictions to plausible values improves non-expert predictions. Relatedly, [DellaVigna and Pope \(2018\)](#) find that non-experts create well-informed predictions about the direction of treatment effects, but often scale them around an incorrect baseline belief. The anchor treatment allows us to investigate the extent to which predictions become more accurate when being provided with a reasonable prior.

Experience treatment Finally, we assign a third randomly chosen group of subjects to an experience condition ($N = 3,092$). Respondents allocated to this intervention group first participated in the initial information experiment before predicting the outcome of this experiment.³ The experience condition allows us to study whether a

¹We consider a prediction correct if it falls within the 10% percentile of the corresponding outcome in the population sample. After the survey, one of the four predictions provided by each participant was randomly selected, and incentives were paid if the chosen prediction met this accuracy criterion.

²The participants learned that in 2019, 78% of 18–69 year old participants—who did not receive information about accurate school spending levels—were in favor of increasing public school spending.

³Due to the contemporaneous structure of the information experiment and the prediction elicitation experiment, participants in the experienced condition participated in the initial baseline experiment first and were subsequently asked to predict its outcomes. Consequently, the experienced groups consist of participants who (i) never received any information about actual spending levels, (ii) always received information about actual spending levels and (iii) could choose to receive information about

potentially better understanding of a task achieved through prior personal engagement with the topic results in more accurate predictions.

2.4 Samples

2.4.1 Expert sample

We exploit the unique infrastructure of the “Economists panel” (German: “*Ökonomen-panel*”) conducted jointly by the ifo Institute in Munich and the *Frankfurter Allgemeine Zeitung* (FAZ) to recruit subjects for our expert panel.⁴ The Economists panel regularly surveys the universe of tenured economics professors at German universities, asking them about their evaluation of recent economic and political events, as well as their recommendations regarding economic policy. The results of the Economists panel are regularly published in academic outlets of the ifo Institute and in press articles in the FAZ, making them transparent to a broad audience in the economic and political sphere. Data collected by the Economists panel has previously been used in related studies that required high-quality responses from professional economic experts (e.g., [Dräger et al., 2025](#) and [Gründler et al., 2025](#)). The survey was conducted as a special module of the 29th Economists panel (tenured professors). We also run a complementary survey among assistant professors that were non-tenured by the time of the survey in May 2020.

We received responses from 127 economics professors, a sample size comparable to previous and subsequent waves of the Economist panel. Respondents are informed about the special module only after participating in the first (regular) part of the survey to mitigate a potential selection into survey. [Table \(A4\)](#) list characteristics of the subjects in our expert sample. For instance, the average expert participating in our study is 53 years old and received their PhD in 2002. A share of 22% of respondents are female. Approximately 8% have previously published in the field of education economics, 21% have experience conducting experimental analyses, and 90% have published an empirical study prior to the survey. These characteristics are broadly consistent with the population of economics professors in Germany and reflect prior analyses revealing that participants of the Economists panel are broadly representative

actual spending levels. For sake of simplicity, we pool all experienced participants into one experienced group when investigating how experience influences predictions.

⁴With a circulation of about 200,000 printed copies daily, the FAZ is Germany’s second-largest high-profile newspaper.

of the universe of economics professors at German universities in terms of academic age, gender, regional distribution of affiliations, and scholarly impact as measured by Google Scholar citations (Geißendörfer et al., 2025).

2.4.2 Non-expert sample

The information experiment, the prediction task and the prediction elicitation experiment were conducted as part of the annually collected ifo Education Survey (for methodological details of this survey, see Freundl et al., 2023). Each annual survey release is widely covered in major German newspapers, and the large coverage of households and the high data quality have made it a popular tool for experimental education research (e.g. Lergetporer et al., 2018; Lergetporer et al., 2020). We included our experiments in the 2020 wave of the ifo Education Survey, which was collected in June 2020 and covers a nationally representative sample of more than 6,000 participants aged between 18 and 65 years. The sample was drawn to match the German population in terms age, gender, federal state, education degree, and employment status.

Table (A1) compares the distribution of key demographic and socioeconomic characteristics in our two survey experiments to the German Microcensus. Across both the information experiment and the prediction elicitation experiment, observable characteristics closely track official population benchmarks. Gender shares, educational attainment, employment rates, and regional composition are very similar across the two experiments and align well with the Microcensus. Respondents in both experiments are slightly younger and less likely to report high household income than in the Microcensus, reflecting common patterns in online survey data. The share of second-generation migrants is modestly higher in the survey samples, while the fraction born in Germany is correspondingly lower. Overall, the comparison indicates that the two experimental samples are broadly representative of the German population along core observable dimensions.

Quality of data The questionnaire is extensively pre-tested to reduce measurement error. To limit low-effort responding, the survey includes several quality-control measures, including attention checks, response-time monitoring, and internal consistency checks across related items. Observations failing these criteria are flagged or excluded according to pre-specified rules. The median completion time is 24 minutes, and respondents receive monetary compensation.

Implementation Respondents were randomly assigned to one of two groups prior to the survey. Half of the subjects were allocated to the information experiment, whose results provide the baseline for the subsequent prediction task and the prediction elicitation experiment (Section 2.1). The other half were directly assigned to the prediction task. The contemporaneous survey structure ensures that no experimental results could leak during the intervention period and that outcomes were not publicly accessible online or elsewhere.

Integrity of randomization To assess the integrity of our randomization, we examine whether participants’ observable sociodemographic characteristics differ across experimental groups. Tables (A2) and (A3) in the appendix present a detailed balancing analysis. The results indicate that covariates are well balanced: Out of 49 pairwise comparisons, only one is statistically significant at the 10% level, consistent with what would be expected under random variation. The numbers provide strong evidence that our random assignment was successful and that systematic differences across treatment groups are unlikely to confound our results.

3 Prediction accuracy: Experts vs. non-experts

We begin our analysis by examining the differences in prediction accuracy between experts and non-experts. In this section, we present the results of the information experiment and compare it with forecasts made by experts and non-experts. This comparison provides a baseline understanding of the magnitude and structure of the expert–lay prediction gap before considering the effects of subject heterogeneity and experimental design features.

3.1 Results of the information experiment

Figure (2) reports the baseline outcomes of the information experiment. In the control group, 77% of participants favor increasing public school spending (left beige bar, left panel), in line with previous evidence (Lergetporer et al., 2018). Provision of information about last year’s per-student spending significantly lowers support to 59% (left beige bar, center panel), while giving participants the option to acquire additional spending information yields 72% support (left beige bar, right panel). The 5-percentage-point decline compared to the control group (77% vs. 72%) is statistically

significant at the 1% level ($p = 0.005$).⁵ These patterns indicate that, although public support for higher education spending remains broadly strong, exposure to objective information moderates this support. This finding aligns with a substantial body of evidence demonstrating that information interventions can significantly influence public preferences, both in education policy and beyond (e.g. [Lergetporer and Woessmann, 2023](#); [Haaland et al., 2023](#)).

3.2 Results of the prediction task: Experts versus non-experts

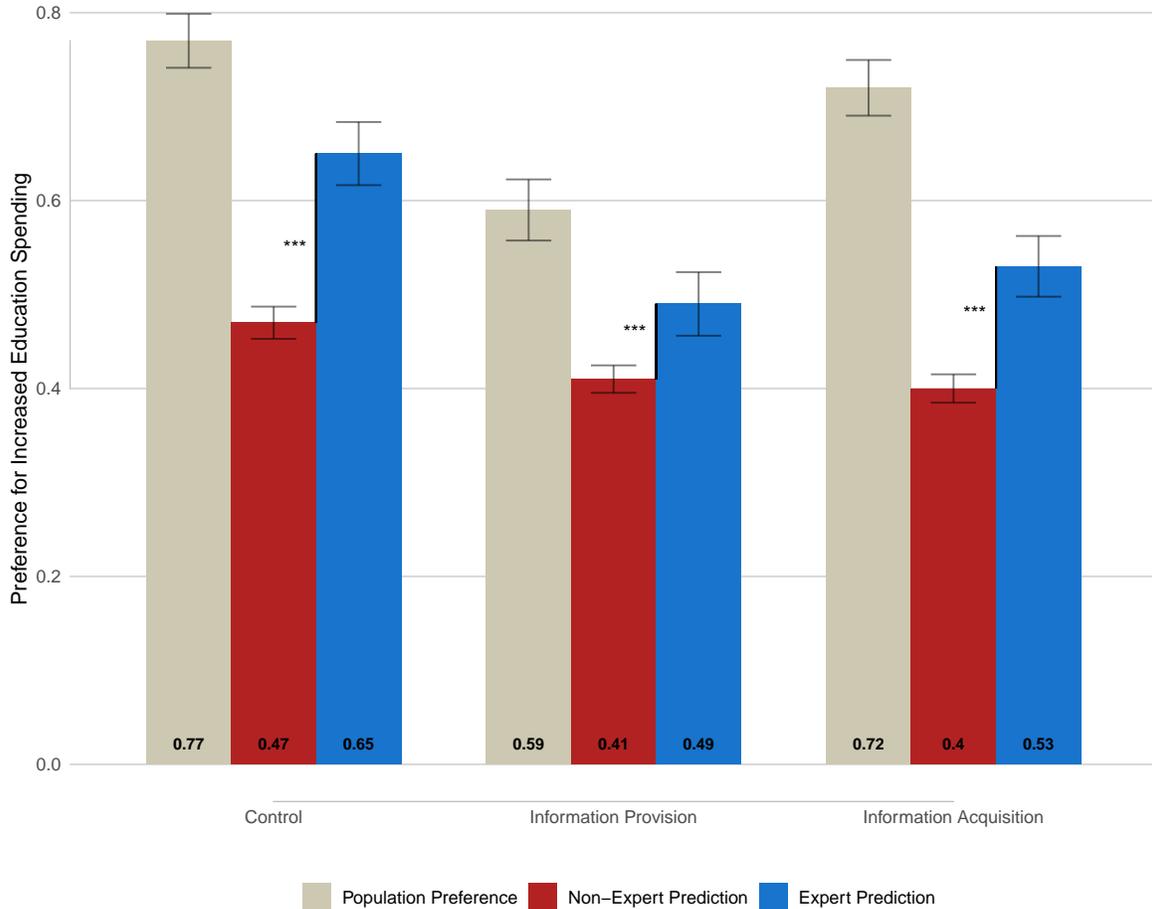
How well do experts and non-experts predict the results of the information experiment? Figure (2) compares the forecasts from the prediction task with the observed outcomes of the information experiment (numerical effects are reported in Table A5). The main result of this comparison is that expert forecasts consistently and significantly outperform those of non-experts across all prediction tasks.

Non-experts substantially underestimate participants' preferences for increased education spending, as illustrated by the red bars in Figure (2). For instance, non-experts forecast that only 47% of participants in the control group would favor an increase (red bar, left panel), a 30 percentage point deviation from the true value of 77%. For the information provision group, non-experts predict 41% support, falling 18 percentage points short of the observed 59% (red bar, center panel). Finally, for the information acquisition group, non-experts estimate that 40% of subjects favor an increase, underpredicting the true 72% by 32 percentage points (red bar, right panel). Non-experts also misjudge the relative strength of the treatment effects, anticipating that the information acquisition task would reduce support slightly more than the information provision treatment, contrary to the observed outcomes.

By contrast, experts generate substantially more accurate predictions, as shown by the blue bars in Figure (2). Quantitatively, experts predict 65% support in the control group (true value: 77%), 49% under information provision (true value: 59%), and 53% under information acquisition (true value: 72%). Across all conditions, expert predictions differ significantly from those of non-experts at the 1% level. Experts also outperform non-experts in ranking treatments by their effects. In particular, experts correctly anticipate that information acquisition shifts preferences less than information provision, and the predicted difference in effect sizes between the control and

⁵Only 22.8% of participants opted to obtain the information; among these, 64% continued to favor spending increases.

Figure 2 Population Preferences and (Non)-Expert Predictions



Note: The Figure shows means and confidence intervals for population preferences and non-expert and expert predictions. Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) are asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the information provision condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context: "The aim now is to estimate the answers of other participants in this survey as accurately as possible. Some other participants are randomly divided into three groups and asked about government spending on schools." Then, the three questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The participants are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?. Population and non-expert sample weighted by survey weights. Significance levels: *p<0.1; **p<0.05; ***p<0.01. Data sources: ifo Education Survey, ifo Economists Panel 2020

information acquisition conditions is statistically significant at the 1% level. This pattern stands in sharp contrast to non-expert predictions, which are largely compressed toward the mean, understating both baseline support and cross-treatment variation.

Distribution of predictions Across all three prediction tasks, the distribution of non-experts’ predictions is markedly more dispersed than that of experts, pointing to a high variability in non-expert beliefs (Figure B1 in the appendix). These results reinforce the conclusion that experts consistently predict research outcomes more accurately than non-experts, in line with prior findings (e.g., [Chu et al., 2024](#)).

3.3 What determines prediction accuracy?

We next explore heterogeneity in prediction accuracy to study whether and to what extent forecasting performance among non-experts and experts can be explained by individual characteristics. Specifically, we examine whether prediction accuracy varies with educational attainment, field of expertise, and self-reported certainty. Across the board, expert forecasts show very little dispersion, while we uncover substantial heterogeneity in the group of non-experts.

3.3.1 Heterogeneity among non-experts

Education and cognitive skills We begin by examining whether non-experts with higher educational attainment make more accurate predictions. We classify respondents based on whether they hold a university entrance degree and compare their forecasts across the control and treatment groups (Figure B2 and Table A7 in the appendix). Non-experts with a university entrance degree consistently provide more accurate predictions than those without. For instance, in the control group, non-experts with a university entrance degree predict that 52% of participants favor higher education spending, a share much closer to the true value (77%) than expected by other participants in the control group (44%). Differences persist in the information provision and information acquisition conditions, though attenuated: Non-experts with and without a university entrance degree predict 43% versus 40% (information provision) and 43% versus 38% (information acquisition), with statistical significance at the 10% and 1% levels. Overall, higher educational attainment narrows the expert–lay prediction gap, yet non-experts remain less accurate than experts even when holding a university entrance degree.

Certainty in predictions We examine the role of self-reported certainty in shaping prediction accuracy. Non-experts who express high certainty (ratings 5–7 on a 7-point scale) predict more accurately than those with low certainty (ratings 1–4) across all experimental conditions (Figure B3 and Table A7). Differences are statistically significant at the 1% level in the control and information provision groups, and at the 5% level in the information acquisition group. However, even highly certain non-experts do not reach the accuracy of experts.

3.3.2 Heterogeneity among experts

Field of specialization Next, we assess whether experts’ predictions vary with their field of specialization. Experts in our sample cover a range of domains, including education economics, experimental economics, microeconomics, and macroeconomics. Figures (B4)–(B7) in the appendix show that experts’ forecasts are remarkably uniform across specializations.⁶ Experts with a focus on education economics make predictions nearly identical to those in other fields. Importantly, all experts outperform the group of non-experts, regardless of their specialization. These findings align with DellaVigna and Pope (2018), who show that prediction accuracy among experts does not depend on field-specific expertise.

Certainty in predictions Contrary for the results we found for non-experts, experts’ predictions are largely invariant to self-reported certainty (Figure B8). For example, in the control group, the values predicted by experts with high (65%) and low (64%) certainty are very close together and statistically indistinguishable.

4 How do experts and non-experts form predictions?

To better understand why experts outperform non-experts in forecasting experimental outcomes, we investigate *how* predictions are formed. Forecasting an experimental result is a cognitively demanding task that requires individuals to start from prior beliefs about the outcome of interest, incorporate available information, and translate both into a coherent quantitative prediction.

In this section, we shed light on these components. We start by exploring *prior beliefs*, which provide the baseline reference point from which respondents form expect-

⁶Numerical treatment effects are presented in Table (A8) in the appendix.

tations about the experimental outcome. We then turn to the *inputs* used to generate forecasts, including information about related evidence, active information search, and self-reported effort. Finally, we examine the *cognitive strategies* through which these inputs are mapped into quantitative predictions. This distinction allows us to disentangle whether expert superiority reflects better-calibrated priors and access to relevant information, or fundamentally different ways of reasoning about and integrating experimental evidence.

4.1 Baseline beliefs about education spending

Spending beliefs of experts and non-experts Figure (B9) plots the distribution of baseline beliefs about education spending in Germany among experts and non-experts. This analysis reveals two distinct facts. First, both groups exhibit dispersion in their beliefs, indicating considerable heterogeneity in perceptions of public spending even for a salient and well-documented policy domain. Second, the expert distribution is substantially more concentrated, with beliefs clustering more tightly around smaller absolute deviations, whereas non-experts display a much wider spread and heavier tails. On average, experts (mean: 7,646 Euros) are remarkably close to the true spending level (7,300 Euros), while non-experts significantly under-predict education spending (mean: 4,684 Euros).

Overall, the results indicate that expert beliefs exhibit significantly higher precision and lower dispersion than those of the general public. Consequently, the magnitude of misperceptions regarding average education spending is substantially lower among the expert sample.

Spending beliefs and prediction accuracy Among non-experts, we find a clear relationship between the accuracy of prior beliefs and forecasting performance (Figure B10). Predictions of individuals with better-calibrated priors are substantially closer to the outcome of the information experiment. This pattern suggests that non-experts rely heavily on their initial beliefs when forming expectations, and that misperceptions at the outset translate directly into less accurate forecasts. By contrast, among experts we detect no meaningful difference in predictive accuracy across baseline spending beliefs (Figure B11). This absence of a gradient is consistent with the lower dispersion of experts' baseline beliefs. It also suggests more generally that experts are less constrained by their baseline beliefs, either because they draw on a broader set of in-

formational inputs or because they apply more structured reasoning when mapping priors into quantitative predictions.

4.2 Prediction inputs

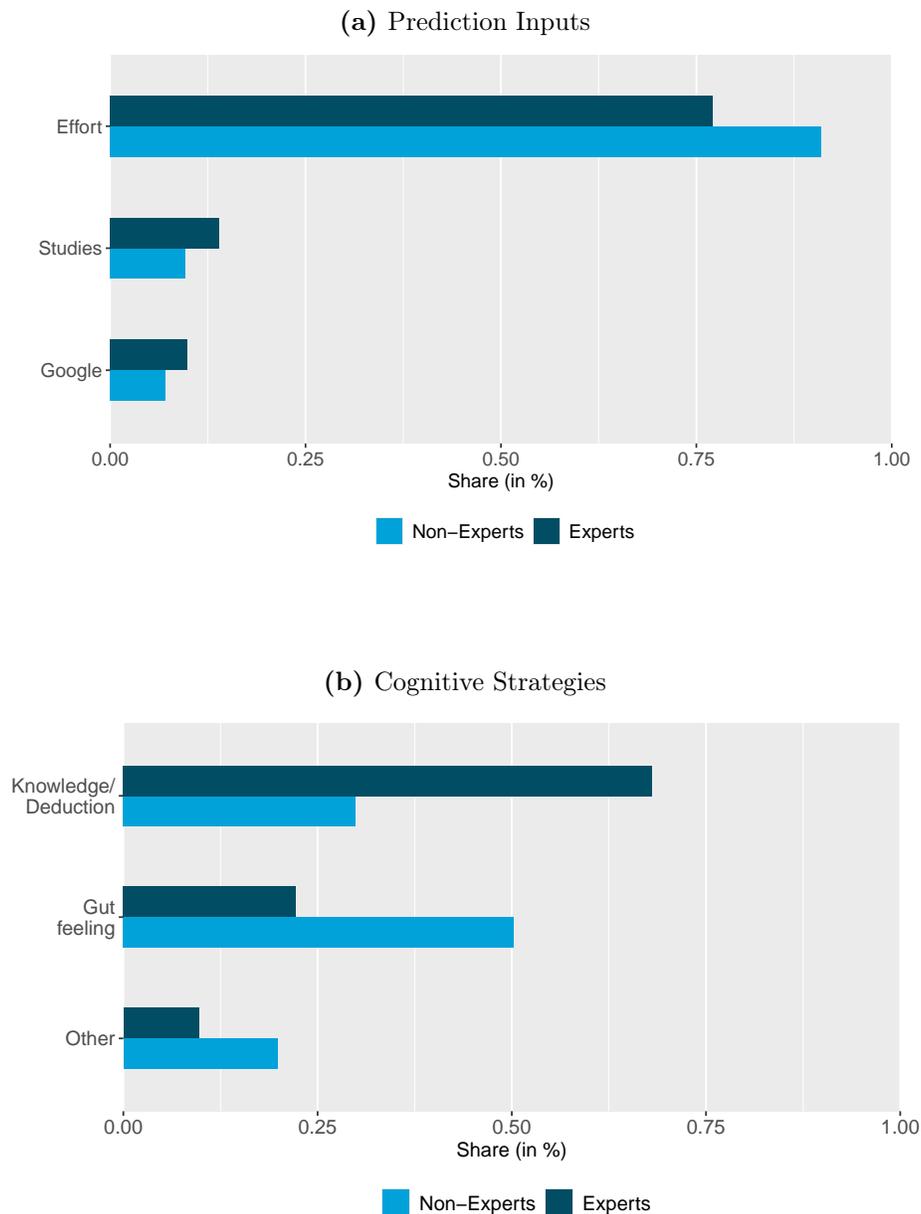
We now study whether experts and non-experts differ in the *inputs* they draw upon when forming predictions. These inputs reflect the resources individuals employ to discipline their baseline spending beliefs, such as exerting greater cognitive effort, searching for external information, or relying on prior exposure to similar scientific studies. We elicited the inputs respondents used to form their forecasts through closed-ended questions in both the expert and non-expert surveys, using identical wording to ensure comparability across groups.

Inputs used by experts and non-experts Panel (a) of Figure (3) summarizes the self-reported inputs respondents used to form their predictions. Large majorities in both groups report exerting effort in arriving at their forecasts (91% among non-experts and 77% among experts). By contrast, only 7% of non-experts and 10% of experts report conducting an internet search, while 10% of non-experts and 14% of experts indicate drawing on knowledge of prior or related studies. Overall, these patterns point to substantial overlap in the self-reported inputs underlying expert and non-expert forecasts, suggesting that differences in performance are unlikely to be explained by informational inputs alone.

Inputs and accuracy of predictions We next examine how different self-reported inputs into the forecasting process relate to prediction accuracy. Two patterns stand out. First, non-experts who report knowing academic studies related to the experimental setting make significantly *less* accurate predictions. We obtain similar results for non-experts who report searching for information online. Despite actively acquiring additional information, their forecasts are farther from the realized outcomes on average. These findings suggest that non-experts face substantial difficulties in translating raw or fragmented information into accurate quantitative predictions.

By contrast, self-reported effort is strongly positively associated with forecasting accuracy. Non-experts who state that they invested particular effort into the task make substantially more accurate predictions. This pattern indicates that improvements in non-expert accuracy can primarily be achieved by more careful processing, integration,

Figure 3 Inputs and Cognitive Strategies to Form Predictions, Experts versus Non-Experts



Note: The figure shows inputs and cognitive strategies used to form predictions, comparing experts and non-experts. Expert and non-expert surveys used the identical question wording to maximize comparability. Non-expert observations are weighted using survey weights. Panel (a) visualizes responses to closed-ended questions on inputs used for the prediction task. Panel (b) depicts the shares of code categories in the open-text question analysis. We separately report shares for experts ($N = 72$) and non-experts ($N = 940$). Answers to the open-text question are assigned to one of the following categories: gut feeling, deduction, knowledge, personal situation, news and other. We summarize the codes 'knowledge' and 'deduction' in the category 'knowledge/deduction', and report the codes 'News', 'Personal situation', 'Other' and 'Don't know' in the category 'Other'. Data sources: ifo Education Survey 2020 and ifo Economists Panel 2020.

and reflection on the available information, rather than from expanding the information set per.

4.3 Cognitive reasoning strategies

Finally, we turn to the cognitive strategies respondents use to transform available inputs into predictions. These strategies capture qualitative differences in reasoning, ranging from deliberate deduction and knowledge-based inference to intuitive or preference-driven judgments. Unlike prediction inputs, which concern the amount and source of information, cognitive strategies describe how individuals process that information. Examining these strategies allows us to assess whether differences in forecast accuracy stem from distinct modes of reasoning rather than from disparities in effort or information alone.

Retrieving cognitive strategies of subjects To classify the cognitive strategies used by experts and non-experts in generating their predictions, we ask participants to explain in a free-text entry box how they arrived at the reported value. This method allows us to measure what’s on top of people’s heads without priming them towards answers they would not have naturally thought of (e.g. [Haaland et al., 2025](#), [Ferrario and Stantcheva, 2022](#)). In order to classify responses, we reviewed a subset of non-expert answers and developed a tailored coding scheme that distinguishes between distinct prediction strategies. This scheme was then consistently applied to all non-expert responses, as well as to the corresponding answers provided by experts. [Table A9](#) details our classification framework, which partitions open-ended responses into six categories: gut feeling, deduction, knowledge, personal situation, news, and other. The table provides representative examples for each category alongside their respective coding assignments.

Differences in cognitive strategies between experts and non-experts [Figure \(3\)](#) summarizes the strategies respondents report using when forming predictions about experimental outcomes, based on a coding of open-ended responses. We group responses referring to personal circumstances, news exposure, or other unspecified considerations into a single residual category. Non-experts predominantly rely on intuitive reasoning, with a majority indicating that they follow their “gut feeling” when making predictions (51%). In contrast, experts are substantially less likely to rely on intuition.

Instead, they most frequently report forming predictions through deductive reasoning or by drawing on relevant domain knowledge and experience (68%). This type of systematic reasoning is substantially less prevalent among non-expert (29%).

Together, the observed between-group differences suggest that experts rely more systematically on structured reasoning and informational inputs, whereas non-experts tend to rely on heuristics and intuition when faced with predictive tasks.

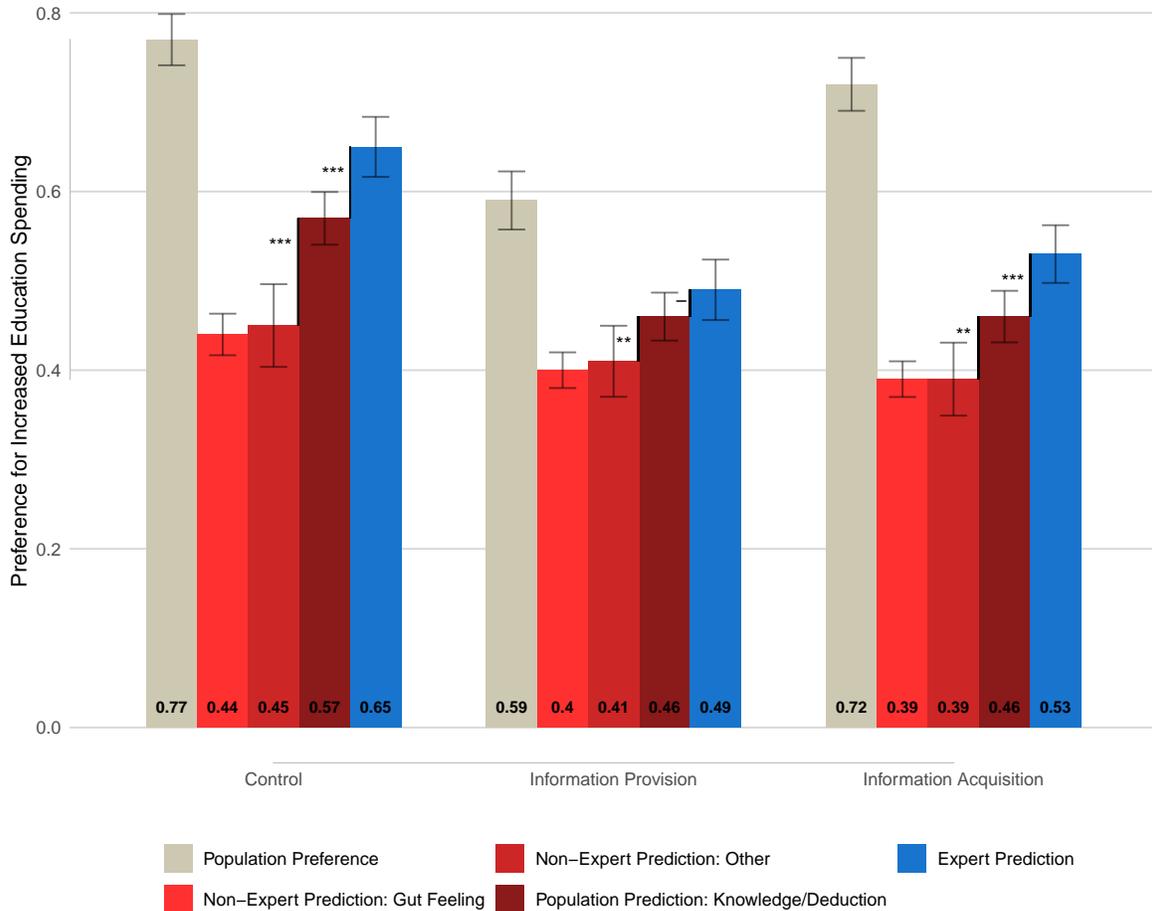
Cognitive strategies and accuracy of predictions Having established that experts and non-experts rely on systematically different strategies to form forecasts, we explore whether these strategies are also associated with predictive accuracy. Specifically, we compare non-experts who report forming predictions based on knowledge or deductive reasoning to those who rely on intuition or other heuristic strategies. Figure (4) presents the results (Table A10 reports numerical outcomes). As in the baseline analysis, the beige and blue bars depict population preferences from the information experiment and expert predictions, respectively. The red bars show non-expert predictions, disaggregated by reported prediction strategy: gut feeling (light red), knowledge or deduction (dark red) and other strategies (medium red).

Non-experts who report relying on knowledge or deductive reasoning make significantly more accurate predictions than those who rely on gut feeling or other heuristic strategies. Their predictions are systematically closer to realized population preferences, with differences statistically significant at the 1% level. For instance, in the control group, non-experts using knowledge or deduction predict that 57% of the population favors higher education spending, compared to a true value of 77%. By contrast, non-experts relying on gut feeling predict a substantially lower share of 44%.

Despite this improvement, non-experts who report using knowledge or deduction remain significantly less accurate than experts. For both the control group and the information acquisition treatment group, the difference in prediction accuracy between experts and highly educated non-experts is statistically significant at the 1% level.

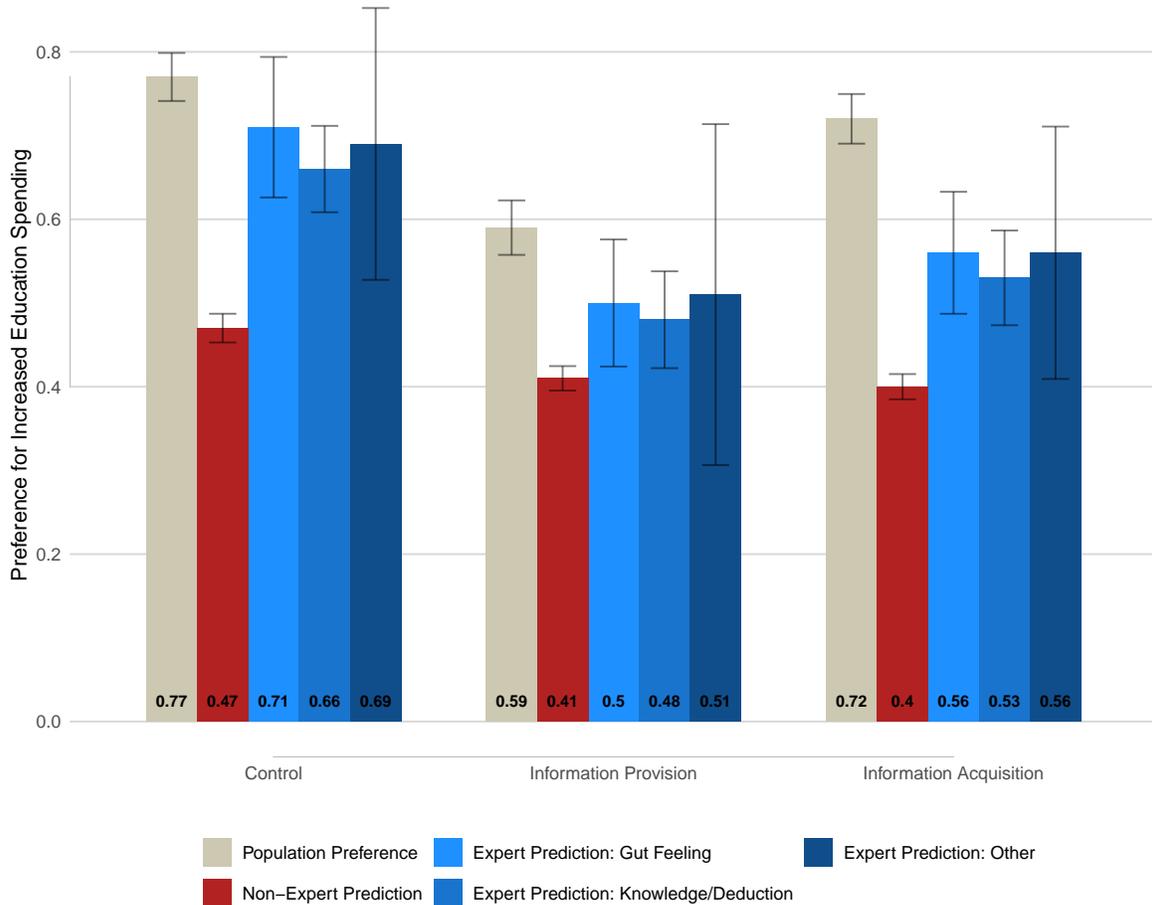
By contrast, variation in reported prediction strategies among experts is weakly related to predictive accuracy. As shown in Figure (5), experts who report relying on knowledge or deduction make similar predictions to those who report relying on intuition (see Table A11 for numerical results). In the control group, experts using knowledge or deduction and those using gut feeling predict population support for higher education spending of 66% and 71%. In the information provision treatment

Figure 4 Non-Expert Predictions split by Answer Strategy



Note: The Figure shows means and confidence intervals for population preferences as well as non-expert and expert predictions. We split the population prediction sample according to respondents' answer strategy (N = 940). We code and summarize answers to the following three categories: knowledge/deduction, gut feeling and other (news, personal situation, other). Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) were asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the information provision condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context: "The aim now is to estimate the answers of other participants in this survey as accurately as possible. Some other participants are randomly divided into three groups and asked about government spending on schools." Then, the three different questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The people are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?. Population and non-expert sample weighted by survey weights. Significance levels: $\bar{p}>0.1$; * $p<0.1$; ** $p<0.05$; *** $p<0.01$. Data sources: ifo Education Survey, ifo Economists Panel 2020

Figure 5 Expert Predictions split by Answer Strategy



Note: The Figure shows means and confidence intervals for population preferences as well as non-expert and expert predictions. We split the expert predictions sample based on different answer strategies (N = 71). We code and summarize answers to three categories: knowledge/deduction, gut feeling and other (news, personal situation, other). Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) are asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the information provision condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context: "The aim now is to estimate the answers of other participants in this survey as accurately as possible. Some other participants are randomly divided into three groups and asked about government spending on schools.". Then, the three different questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The people are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?. Population and non-expert sample weighted by survey weights. Significance levels: *p<0.1; **p<0.05; ***p<0.01. Data sources: ifo Education Survey, ifo Economists Panel 2020

group, experts predict 48% and 50%.

5 Can non-expert predictions be improved?

We examine whether the design of the prediction task can narrow the gap between expert and non-expert forecasts. Specifically, we study whether features of the elicitation environment improve non-experts’ predictive accuracy and bring their forecasts closer to those of experts. We then compare the effects of these design interventions to other determinants of predictive performance (e.g. socio-economic characteristics and cognitive reasoning strategies) in order to assess the relative importance of survey-based elicitation choices versus individual-level factors in shaping prediction accuracy.

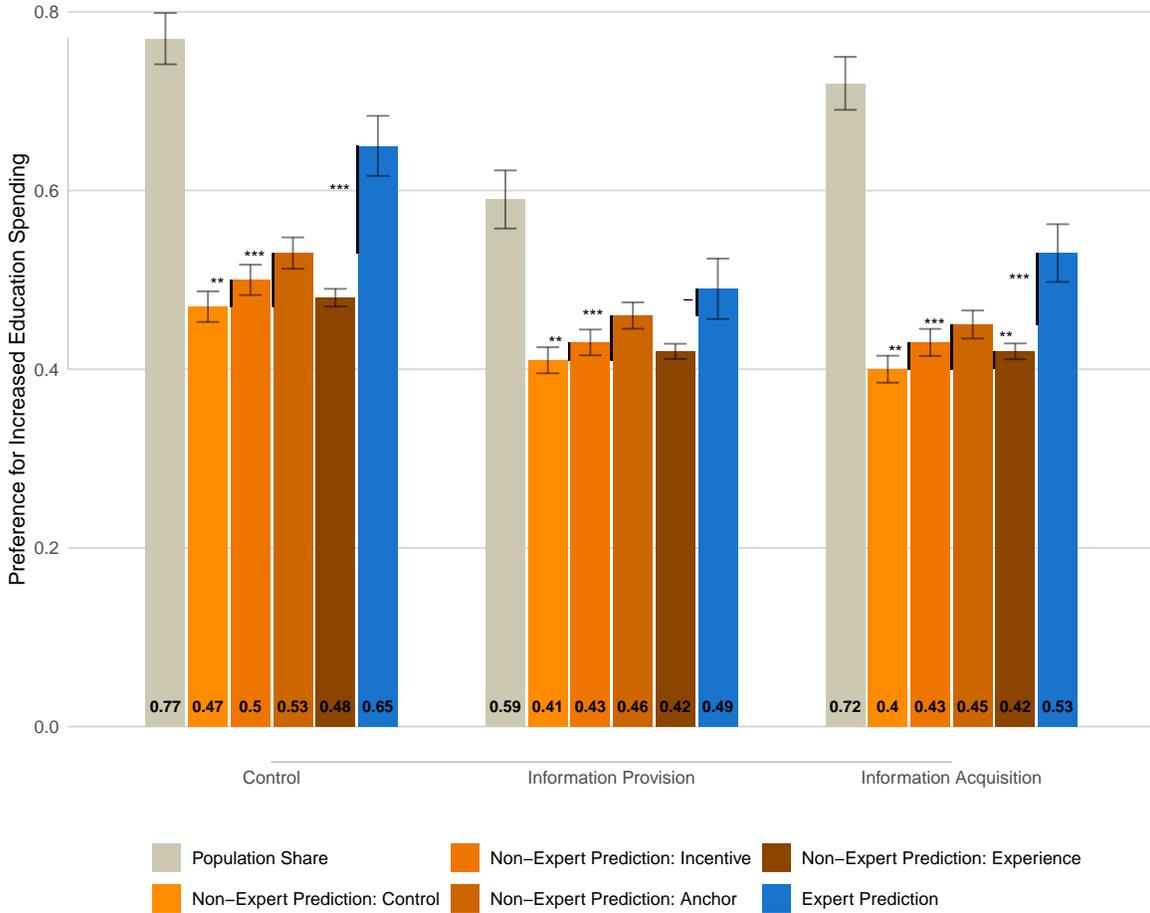
5.1 The impact of elicitation design on prediction accuracy

Our previous results suggest that experts’ superior predictive performance is plausibly rooted in a combination of well-calibrated baseline anchors, rigorous scientific training that fosters deduction-based reasoning strategies, accumulated experience with forecasting in professional settings (e.g. advising policymakers or supervising research projects), and institutional incentives that reward accuracy.

To assess whether analogous mechanisms enhance non-expert forecasting, we implement three experimental interventions that selectively introduce core elements of expert prediction into the non-expert elicitation environment, including focal reference points, disciplined reasoning, and experience with prediction tasks. The *incentive treatment* awards participants a monetary reward (1 Euro) for correct predictions. The *anchor treatment* provides participants with the population share favoring increased education spending from a prior survey (2019, 78%). Finally, in the *experience treatment*, participants generate firsthand experience with the decision context by first taking part in the information experiment. These treatments are implemented identically to predict preferences for education spending in the control group, the information provision group, and the information acquisition group. We rationalize the results regarding design features in our simple framework in Appendix (A).

Results Figure (6) presents the results. We find that both the incentive and anchor treatments significantly improve non-expert accuracy. For example, in the control group, non-experts predict the population share favoring increased education spending

Figure 6 Non-Expert Predictions by Belief Elicitation Treatments



Note: The Figure shows means and confidence intervals for population predictions under individual treatments: incentive, anchor and experience treatment. In the incentive condition, before showing the prediction questions, participants are informed about the incentive of 1 Euro for a correct answer. In the anchor condition, participants are informed about the population share for increased education spending for the year 2019 (78%). In the experience treatment, participants took part in the first experiment and indicated their preference for or against increased education spending in either one of the three conditions: control, information provision, information acquisition. Each treatment arm is a different sample: incentive treatment (1,040), anchor treatment (1,019), experience treatment (3,094), control group (1,024). Population preferences and expert predictions are included as references. Participants in all treatment arms make predictions for the three conditions in the information provision experiment: control, information provision and information acquisition. Population and non-expert sample weighted by survey weights. Significance levels: $\bar{\ }p>0.1$; * $p<0.1$; ** $p<0.05$; *** $p<0.01$. Data sources: ifo Education Survey, ifo Economists Panel 2020

to be 48% without treatment, 50% under the incentive treatment, and 53% under the anchor treatment (true value: 77%). Across all three groups of the prediction task, the anchor treatment is consistently the most effective in enhancing prediction accuracy, while the incentive treatment also yields statistically significant but smaller improvements. By contrast, the experience treatment has little effect in the control and information provision groups, and only a minor effect in the information acquisition group (42% vs. 40%).

Despite the improvements generated by the elicitation design, none of the treatments is sufficient to fully close the gap between non-experts and experts. In the control group, experts predict 65% (true value: 77%), whereas the closest non-expert prediction (the anchor treatment) reaches 53%. Even under the anchor treatment in the information provision task, which yields the smallest observed divergence, non-experts' predictions still fall 3 percentage points short of experts' forecasts.

Main takeaway Our findings imply that the design of the prediction task can significantly improve non-expert forecasts, yet intrinsic differences between experts and non-experts continue to sustain the expert-lay prediction gap.

5.2 Elicitation design versus individual determinants of prediction accuracy

The previous analyses show that both individual characteristics (such as education), answer strategies (such as knowledge and deduction), and features of the elicitation design (such as incentives and anchoring) are associated with higher accuracy, but it remains unclear how these forces compare quantitatively. Next, we provide a unified assessment of the relative importance of characteristics, answer strategies, and experimental treatments for forecast accuracy. This approach allows us to identify which levers, task design or respondent attributes, are most effective in narrowing the expert-lay prediction gap and to what extent improvements in non-expert forecasting can be achieved through survey design.

Model To quantify which factors improve the accuracy of non-expert forecasts, we estimate a series of linear regression models that relate individual prediction errors to respondent characteristics, self-reported prediction strategies, and experimental treatments. The empirical specification takes the following form:

$$|\widehat{y}_{ig} - y_g| = \alpha + \mathbf{X}'_i \boldsymbol{\beta} + \mathbf{S}'_i \boldsymbol{\gamma} + \mathbf{T}'_i \boldsymbol{\delta} + \varepsilon_{ig}, \quad (1)$$

where $|\widehat{y}_{ig} - y_g|$ denotes the absolute difference between individual i 's prediction of the population share in experimental group g and the realized population share. Lower values of the dependent variable therefore indicate more accurate forecasts.

The vector \mathbf{X}_i captures baseline respondent characteristics and prior beliefs, including an indicator for tertiary education, self-reported certainty about the prediction, and a measure of the respondent's misperception about actual education spending. The vector \mathbf{S}_i contains indicators for self-reported prediction strategies, distinguishing between knowledge- or deduction-based reasoning, gut feeling (the omitted category), and other heuristic approaches. The vector \mathbf{T}_i includes indicators for the experimental treatments in the prediction elicitation task, namely monetary incentives, anchoring, and prior forecasting experience. The error term ε_{ig} captures unobserved determinants of forecast accuracy.

All covariates are dummy variables, facilitating a direct comparison of magnitudes across characteristics, strategies, and treatments. The regressions are estimated on the sample of non-experts participating in the prediction elicitation experiment.

Results Table (1) reports the results. Column (1) focuses on baseline characteristics and prior beliefs. Non-experts with a tertiary education degree make significantly more accurate predictions, reducing the absolute prediction error by 4 percentage points. Above-median self-reported certainty is also strongly associated with improved accuracy. By contrast, respondents whose own beliefs about education spending are further from the true spending level in absolute terms make systematically worse predictions, highlighting the role of biased priors in shaping forecast errors.

Column (2) adds self-reported prediction strategies. Relative to respondents relying on gut feeling, non-experts who report using knowledge or deductive reasoning make substantially more accurate predictions, with a reduction in absolute error of nearly 0.05 points. In contrast, respondents relying on other heuristic strategies perform significantly worse. These results underscore that how non-experts approach the prediction task is a key determinant of accuracy, above and beyond observable characteristics.

Column (3) adds the experimental treatments. Among these, anchoring has a large and precisely estimated effect, improving prediction accuracy by about 0.04 points.

Table 1 Drivers of Good Predictions Among Non-Experts

<i>Dependent variable: Absolute Difference Non-Expert Predictions - Population Share</i>			
Prediction Elicitation Experiment			
	(1)	(2)	(3)
Difference in spending beliefs	0.028*** (0.005)	0.025*** (0.005)	0.024*** (0.005)
Tertiary education degree	-0.040*** (0.005)	-0.028*** (0.005)	-0.028*** (0.005)
Above median certainty	-0.024*** (0.005)	-0.020*** (0.005)	-0.020*** (0.005)
Answer Strategy:			
Knowledge/Deduction		-0.049*** (0.005)	-0.048*** (0.005)
Other		0.068*** (0.008)	0.069*** (0.008)
Experimental Treatments:			
Incentive			-0.013* (0.008)
Anchor			-0.041*** (0.008)
Experience			-0.004 (0.006)
Constant	0.294*** (0.004)	0.296*** (0.004)	0.307*** (0.007)
Observations	16,872	16,872	16,872

Note: Survey-weighted linear regression estimates. The dependent variable is the absolute difference between non-expert predictions and average population preferences. The lower the absolute difference, the closer a prediction is to the population average. We focus on the prediction elicitation experiment sample (5,624 individuals make three predictions in three treatment and the control group), and regress the dependent variable on the 'Difference in spending beliefs' dummy equaling 1 if the absolute difference between individual education spending beliefs and factual educations spending is above the median absolute difference, a tertiary education dummy equaling 1 if a respondent has a tertiary education degree, an 'Above median certainty' dummy equaling 1 if individual prediction certainty is above median certainty, and the answer strategy with 'Gut feeling' as reference category. Standard errors are clustered at individual level. Significance levels: *p<0.1; **p<0.05; ***p<0.01. Data source: ifo Education Survey 2020.

Monetary incentives also reduce prediction errors, though the effect is smaller and only marginally significant. Prior forecasting experience, by contrast, has no statistically detectable effect once other factors are controlled for. Notably, the estimated anchoring effect is roughly twice as large as the effect associated with tertiary education and closely aligned with the magnitude of knowledge-based reasoning.

Implications for Survey Design The findings have direct implications for the design of surveys that elicit beliefs or forecasts from non-experts. First, raw predictions should not be interpreted in isolation: Heterogeneity in respondents’ priors and reasoning strategies generates systematic and predictable forecast errors. Second, survey design can partially substitute for expertise. Simple design choices—most notably providing numerical anchors—can substantially improve forecast accuracy by guiding respondents toward more structured reasoning and mitigating the influence of diffuse or biased priors. Third, encouraging cognitive effort through incentives yields additional, albeit smaller, gains, suggesting diminishing returns to purely monetary motivation. More broadly, our results imply that surveys aimed at aggregating non-expert expectations or beliefs can achieve better predictions. The scope for improvement is, however, limited. Even when improving forecasts through survey-based approaches, non-expert forecasts cannot catch up with the accuracy of expert predictions.

5.3 A simple framework to understand the treatment effects

The patterns identified in our experiments can be rationalized within a simple model of prediction formation, presented in Appendix (A). The framework formalizes how heterogeneous prior beliefs and costly, structured reasoning generate a persistent expert–lay gap in forecasts. Experts enter with better calibrated priors and exert greater cognitive effort, while non-experts rely more heavily on heuristics. Anchors and incentives affect predictions through two distinct channels: by shifting priors toward more accurate baselines and by increasing the returns to cognitive effort. The model provides a parsimonious account of the observed treatment effects and clarifies some of the mechanisms underlying belief formation in our setting.

6 Conclusion

Our analysis documents large and persistent gaps between expert and non-expert forecasts, even in a well-defined and policy-relevant setting. In our data, these differences are explained less by narrow domain expertise than by systematic heterogeneity in priors, reasoning strategies, and task engagement. Among non-experts, simple elicitation features—most notably numerical anchors and incentives for accuracy—substantially improve performance, although they do not eliminate the expert–lay gap.

The differences in predictive beliefs uncovered in our experiments may have first-order implications outside the survey context. Misperceptions about the effects of policy interventions can weaken support for evidence-based reforms, from education and social insurance to broader structural policies. An important task for future research is to evaluate which scalable communication and elicitation designs (e.g. credible quantitative benchmarks, accuracy incentives, and structured reasoning prompts) can be deployed in real policy communication, and whether they durably improve belief formation in the field.

References

- Andre, P., Pizzinelli, C., Roth, C., and Wohlfart, J. (2022). Subjective models of the macroeconomy: Evidence from experts and representative samples. *Review of Economic Studies*, 89(6):2958–2991.
- Banuri, S., Dercon, S., and Gauri, V. (2019). Biased policy professionals. *World Bank Economic Review*, 33(2):310–327.
- Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M., and Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, 351(6280):1433–1436.
- Chu, J. Y., Voelkel, J. G., Stagnaro, M. N., Kang, S., Druckman, J. N., Rand, D. G., and Willer, R. (2024). Academics are more specific, and practitioners more sensitive, in forecasting interventions to strengthen democratic attitudes. *Proceedings of the National Academy of Sciences*, 121(3):e2307008121.
- Coibion, O. and Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, 105(8):2644–2678.
- Delavande, A., Giné, X., and McKenzie, D. (2011). Eliciting probabilistic expectations with visual aids in developing countries: how sensitive are answers to variations in elicitation design? *Journal of Applied Econometrics*, 26(3):479–497.
- DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature*, 47(2):315–372.
- DellaVigna, S. (2018). Structural behavioral economics. In DellaVigna, S. and Malmendier, U., editors, *Handbook of Behavioral Economics: Applications and Foundations 1*, volume 1, pages 613–723. Elsevier, Amsterdam.
- DellaVigna, S., Otis, N., and Vivaldi, E. (2020). Forecasting the results of experiments: Piloting an elicitation strategy. *AEA Papers and Proceedings*, 110:75–79.
- DellaVigna, S. and Pope, D. (2018). Predicting experimental results: Who knows what? *Journal of Political Economy*, 126(6):2410–2456.
- DellaVigna, S., Pope, D., and Vivaldi, E. (2019). Predict science to improve science. *Science*, 366(6464):428–429.
- Dräger, L., Gründler, K., and Potrafke, N. (2025). Political Shocks and Inflation Expectations: Evidence from the 2022 Russian Invasion of Ukraine. *Journal of International Economics*, 153:104029.
- Ferrario, B. and Stantcheva, S. (2022). Eliciting people’s first-order concerns: Text analysis of open-ended survey questions. *AEA Papers and Proceedings*, 112:163–69.
- Freundl, V., Grewenig, E., Kugler, F., Lergertporer, P., Schüler, R., Wedel, K., Werner, K., Wirth, O., and Woessmann, L. (2023). The ifo education survey 2014–2021: A new dataset on public preferences for education policy in Germany. *Jahrbücher für Nationalökonomie und Statistik*, 243(6):699–710.
- Geißendörfer, L., Gründler, K., Potrafke, N., and Schlepper, M. (2025). Wie groß ist der wirtschaftspolitische Konsens unter Ökonominen und Ökonomen in Deutschland?

- Evidenz aus 74.000 Antworten im Ökonomenpanel. *Perspektiven der Wirtschaftspolitik*, (0).
- Grewenig, E., Lorgetporer, P., Werner, K., and Woessmann, L. (2022). Incentives, search engines, and the elicitation of subjective beliefs: Evidence from representative online survey experiments. *Journal of Econometrics*, 231(1):304–326.
- Gründler, K., Lamla, M., Potrafke, N., and Wochner, T. (2026). Experts as intermediaries. *CESifo Working Paper 12492*.
- Gründler, K., Potrafke, N., and Wochner, T. (2025). Outside employment and parliamentary priorities. *Journal of Economic Behavior & Organization*, 237:107185.
- Haaland, I., Roth, C., Stantcheva, S., and Wohlfart, J. (2025). Understanding economic behavior using open-ended survey data. *Journal of Economic Literature*, 63(4):1244–80.
- Haaland, I., Roth, C., and Wohlfart, J. (2023). Designing information provision experiments. *Journal of Economic Literature*, 61(1):3–40.
- Hjort, J., Moreira, D., Rao, G., and Santini, J. F. (2021). How research affects policy: Experimental evidence from 2,150 Brazilian municipalities. *American Economic Review*, 111(5):1442–80.
- Lorgetporer, P., Schwerdt, G., Werner, K., West, M. R., and Woessmann, L. (2018). How information affects support for education spending: Evidence from survey experiments in Germany and the United States. *Journal of Public Economics*, 167:138–157.
- Lorgetporer, P., Werner, K., and Woessmann, L. (2020). Educational inequality and public policy preferences: Evidence from representative survey experiments. *Journal of Public Economics*, 188:104226.
- Lorgetporer, P. and Woessmann, L. (2023). Earnings information and public preferences for university tuition: Evidence from representative experiments. *Journal of Public Economics*, 226:104968.
- Manski, C. F. (2004). Measuring expectations. *Econometrica*, 72(5):1329–1376.
- Prior, M., Sood, G., and Khanna, K. (2015). You cannot be serious: The impact of accuracy incentives on partisan bias in reports of economic perceptions. *Quarterly Journal of Political Science*, 10(4):489–518.
- Roth, C. and Wohlfart, J. (2018). How do expectations about the macroeconomy affect personal expectations and behavior? *Review of Economics and Statistics*, 100(4):731–748.
- Schlag, K. H., Tremewan, J., and van der Weele, J. J. (2015). A penny for your thoughts: A survey of methods for eliciting beliefs. *Experimental Economics*, 18(3):457–490.
- Stantcheva, S. (2021). Understanding tax policy: How do people reason? *Quarterly Journal of Economics*, 136:2309–2369.
- Trautmann, S. T. and van de Kuilen, G. (2015). Belief elicitation: A horse race among truth serums. *Economic Journal*, 125(589):2116–2135.

A Conceptual Framework

This section describes a parsimonious framework to interpret (i) why experts forecast experimental outcomes more accurately than non-experts, (ii) why accuracy among non-experts is strongly related to education and engagement, and (iii) why simple elicitation features—anchors, incentives, and experience—can improve non-expert accuracy without fully eliminating the gap in forecasting accuracy between experts and non-experts. The framework treats the prediction task as an inference problem in which agents combine a prior with task information, but differ in (a) the accuracy of their priors and (b) the degree to which they engage in structured reasoning when mapping information into a quantitative forecast.

A.1 Object of prediction and information in the task

Let $\theta \in \mathbb{R}$ denote the unknown experimental object respondents are asked to forecast. In the application, θ can be baseline support for increased education spending or a treatment effect (e.g., the difference in support between an information arm and control). Respondents observe a common description of the experimental design, denoted by D , which is informative about θ but does not reveal the outcome itself.

We assume that D can be summarized by a “structural” or “design-implied” signal $s(D)$, capturing information about the object respondents are asked to forecast via

$$s(D) = \theta + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma_s^2),$$

where σ_s^2 captures the remaining uncertainty about the object θ even after learning the experimental design (e.g., because treatment effects depend on behavioral responses that are difficult to anticipate).

A.2 Priors and miscalibration

Each individual i begins with a subjective prior over θ , which we define as

$$\theta \sim \mathcal{N}(\mu_i, \sigma_{0,i}^2),$$

where μ_i captures prior beliefs, and $\sigma_{0,i}^2$ captures perceived uncertainty. We allow μ_i to vary across experts and non-experts and within each group.

A key mechanism uncovered by our experiments is that errors in priors translate into errors in forecasts. Define the prior miscalibration term $b_i \equiv \mu_i - \theta$. Cross-sectional heterogeneity in μ_i therefore generates persistent heterogeneity in forecast accuracy even when all individuals observe the same task description D .

A.3 The formation of forecasts

We model prediction as a two-step process: Individuals (i) decide how much cognitive effort to devote to the task, and (ii) produce a point forecast by combining their prior with the information contained in D .

Effort choice Individuals translate the experimental description D into a usable quantitative signal. This process involves a comprehension error $v - (e_i)$ which depends on costly cognitive effort (attention, careful reading, deliberate calculation), denoted by $e_i \geq 0$. Greater effort reduces the noise with which the individual extracts the signal from the design description.

Specifically, the individual’s internal representation of the design-implied signal is for a chosen effort e_i is

$$\tilde{s}_i = s(D) + \nu_i(e_i) = \theta + \varepsilon + \nu_i(e_i), \quad \nu_i(e_i) \sim \mathcal{N}(0, \sigma_{\nu,i}^2(e_i)),$$

with $\sigma_{\nu,i}^2(\cdot)$ decreasing in e_i . In principle, $\sigma_{\nu,i}^2(\cdot)$ captures a broad class of errors: misreading the design, failing to condition on treatment arms, confusion about baseline versus treatment effects, and difficulty translating qualitative reasoning into a numerical forecast.

Intuitive processing corresponds to low e_i (large $\sigma_{\nu,i}^2$), while structured reasoning corresponds to higher e_i (smaller $\sigma_{\nu,i}^2$).

Individuals choose effort to maximize expected payoff from accuracy net of effort costs:

$$\max_{e_i \geq 0} \underbrace{\lambda_i \cdot \mathbb{E}[-(\hat{\theta}_i - \theta)^2]}_{\text{Gain of accuracy}} - \underbrace{c_i(e_i, r_i)}_{\text{Cost of accuracy}},$$

where $c_i(\cdot)$ covers all costs of effort, including mental fatigue, opportunity cost, or time cost, and is increasing and convex. We allow the cost of cognitive effort to depend on an individual’s reasoning strategy. Let $r_i \in \{I, D\}$ index whether respondent i primarily relies on intuition (I) or on more structured, deductive reasoning (D), as

measured in our open-ended responses. For a given effort level e , structured reasoners face lower marginal (and hence total) costs of deliberation $c'_i(e; D) < c'_i(e; I)$. This formulation captures that deductive reasoning is not merely “more effort”, but a different *mode* of processing. Individuals trained in structured thinking can translate a complex task description into quantitative implications using established mental routines (e.g., decomposition into baseline and treatment effects, back-of-the-envelope calculations, consistency checks). As a result, a given increase in effort is less exhausting and less time-consuming for deductive reasoners than for intuitive reasoners, for whom sustained reflection is more cognitively demanding. The parameter λ_i scales the private return to accuracy (i.e. how much the person cares about being accurate).

Considering the loss function $L_i(e_i) = (\hat{\theta}_i - \theta)^2$, which is decreasing in e_i , the optimum satisfies

$$\lambda_i(-L'_i(e_i^*)) = c'_i(e_i^*, r_i).$$

This suggests that the optimal effort, e^* , is chosen such that the marginal *benefit* of effort (reduction in expected squared error), valued at λ_i , equals the marginal *cost* of effort.

Forecasting The effort chosen in the previous step determines the accuracy of the extracted signal \tilde{s}_i that individuals will use as input for the forecasting task. Given this signal, the Bayesian posterior mean under the Gaussian structure (i.e. priors and noise terms are Normal (Gaussian), and the signals are linear with additive Normal noise) is a weighted combination of priors and the extracted signal

$$\hat{\theta}_i^B = w_i \tilde{s}_i + (1 - w_i) \mu_i, \quad w_i = \frac{\sigma_{0,i}^2}{\sigma_{0,i}^2 + \sigma_s^2 + \sigma_{\nu,i}^2(e_i)}.$$

From this formation of posteriors, we can identify three sources of forecasting errors:

1. *Prior miscalibration*: if μ_i is biased, forecasts remain biased whenever $w_i < 1$.
2. *Noisy processing and low effort*: if $\sigma_{\nu,i}^2(e_i)$ is large, w_i is small and forecasts stick closely to the prior; moreover, conditional on w_i , noise in \tilde{s}_i inflates the variance of forecast errors.
3. *Cognitive reasoning strategy*: reasoning strategies affect forecast errors by shifting the cost of deliberation and therefore the chosen effort level e_i^* . Intuitive respon-

dents face higher effective costs of structured processing and thus optimally rely more on their priors, whereas deductive reasoners find it less costly to engage with the task and to extract the implications of the experimental design. In the model, this is captured by allowing the effort cost to depend on strategy r_i , which generates systematic differences in forecast accuracy even when all respondents observe the same task description.

A.4 Forecasting accuracy and elicitation design

Our experiments reveal that features of the elicitation design (especially monetary incentives and numerical anchors) impact the prediction gap. The effects of these features can be shown even in this simple conceptual framework.

Effort Monetary incentives directly raise λ_i , which should increase the optimal level of effort that is chosen. According to the process underlying the formation of posteriors, this should directly translate into lower forecast errors.

Numerical anchors Providing a salient and plausible benchmark a (e.g., drawn from prior survey evidence) can partially substitute for a noisy prior by supplying an externally grounded reference point. In this case, the anchor induces a *prior shift*

$$\mu_i^A = (1 - \alpha_i)\mu_i + \alpha_i a, \quad \alpha_i \in [0, 1],$$

where α_i captures the degree a respondent uses the anchor. The posterior forecast then takes the same form as before, but with the anchor-adjusted prior mean

$$\hat{\theta}_i^{B,A} = w_i \tilde{s}_i + (1 - w_i)\mu_i^A.$$

Anchoring improves accuracy whenever the anchor is, on average, better calibrated than the individual’s original prior, i.e., when $|a - \theta| < |\mu_i - \theta|$ for a substantial share of respondents. The gains are predicted to be largest for respondents with initially miscalibrated priors and for those who place relatively little weight on the design-implied signal (low w_i).

Experience In principle, our framework also offers an explanation for why experience might matter, even though the experimental results have demonstrated that this

channel is of little practical relevance. Prior exposure to a related task can be captured as lowering the cost of effort (or, equivalently, increasing processing efficiency). In the model, experience either (i) shifts down the effort cost $c_i(e^*, r_i)$, or (ii) improves the mapping from D into \tilde{s}_i for a given effort level (a downward shift in $\sigma_{\nu,i}^2(e)$). Either mechanism raises effective w_i and reduces reliance on the prior.

A.5 Experts versus non-experts

The expert–lay gap can be represented by differences in two key determinants. First, experts tend to have better-calibrated priors (smaller bias in μ_i). Second, experts more often rely on structured reasoning, raising the optimally chosen effort e^* invested in the task, reflecting scientific training and familiarity with structured, logical thinking processes.

Among non-experts, education can shift the same determinants, although to a lesser extent. More educated respondents may have priors that are closer to the true state (smaller bias in μ_i) and may also find it less costly to process experimental descriptions, for instance due to higher numeracy or greater familiarity with abstract reasoning. Still, these effects are likely to be smaller than for experts, for whom structured reasoning of this kind is routine and embedded in their day-to-day professional work.

B Supplementary Tables

Table A1 Descriptive Statistics: Comparison of Microcensus and Survey Data

	Information Experiment	Prediction Elicitation Experiment	Microcensus
Female	0.494	0.498	0.500
Age (years)	44.802	44.634	46.292
Tertiary education degree	0.409	0.401	0.402
Born in Germany	0.941	0.942	0.993
2nd generation migrant	0.089	0.091	0.043
Employed	0.717	0.714	0.742
High income	0.361	0.367	0.587
North	0.093	0.088	0.099
South	0.289	0.287	0.238
East	0.174	0.173	0.213
West	0.444	0.452	0.450
N	3,133	6,177	266,200

Note: This table displays the mean value of basic covariates from the Microcensus in Germany 2020 and the representative dataset we collected in 2020, split by the two experiments and using survey weights. "Female" is a binary variable with value one for female respondents. "Age (years)" is the age of the respondent. "Tertiary education degree" is a binary variable taking value one if the respondent has completed tertiary education (German "Abitur"). "Born in Germany" is a dummy variable taking the value 1 if the respondent is born in Germany. "2nd Generation migrant" is a binary variable taking value 1 if both parents of the respondent were not born in Germany. "Employed" is a dummy variable taking the value one if the respondent is employed full-time, part-time, or self-employed. "High income" is a dummy variable taking value one if the respondent reports a net household income at or above EUR 3,000 per month. "North", "West", "South", and "East" are binary variables taking value one if the respondent lives in the respective region. Data sources: ifo Education Survey, Microcensus Germany 2020

Table A2 Balancing Table: Information Experiment

	No Information		Information		Diff. in Means	p-value
	Mean	Std. Dev.	Mean	Std. Dev.		
Information Provision						
Being female	0.49	0.50	0.50	0.50	0.01	0.725
Age	44.40	14.15	45.00	14.23	0.60	0.394
Tertiary education degree	0.39	0.49	0.42	0.49	0.02	0.279
Living in West Germany	0.80	0.40	0.81	0.39	0.01	0.561
Net household income	10.39	4.75	10.66	4.74	0.27	0.223
Born in Germany	0.94	0.23	0.94	0.24	-0.01	0.482
Migration background	0.14	0.35	0.16	0.37	0.02	0.302
Being a parent	0.56	0.50	0.58	0.49	0.02	0.403
Being a teacher	0.05	0.22	0.04	0.19	-0.01	0.135
Being employed	0.73	0.44	0.71	0.46	-0.02	0.289
Information Acquisition						
Being female	0.49	0.50	0.49	0.50	-0.01	0.830
Age	44.40	14.15	45.01	14.36	0.62	0.377
Tertiary education degree	0.39	0.49	0.42	0.49	0.02	0.304
Living in West Germany	0.80	0.40	0.82	0.39	0.02	0.305
Net household income	10.39	4.75	10.36	4.73	-0.03	0.897
Born in Germany	0.94	0.23	0.94	0.23	0.00	0.974
Migration background	0.14	0.35	0.16	0.37	0.02	0.322
Being a parent	0.56	0.50	0.54	0.50	-0.02	0.428
Being a teacher	0.05	0.22	0.04	0.20	-0.01	0.349
Being employed	0.73	0.44	0.72	0.45	-0.01	0.525

Note: Table summarizes weighted group means and standard deviations for two comparisons: control vs. information provision condition, control vs. information acquisition condition. The net household income is indicated in income classes ranging from below €400 up to €5,000 and more (in steps of €250). Average income class 10 corresponds to a net household income ranging from €2,250 to €2,500. West Germany is a dummy equaling 1 if a respondent lives in a former West German state, and 0 if a respondent lives in a former East German state. Migration background is a dummy equaling 1 if both parents of respondents are not born in Germany, and 0 otherwise. Univ. entrance qualification is a dummy equaling 1 if a respondent has a higher education school degree, and 0 otherwise. Significance levels: *p<0.1; **p<0.05; ***p<0.01. Data source: ifo Education Survey 2020.

Table A3 Balancing Table: Prediction Experiment

	Control		Treatment		Diff. in Means	p-value
	Mean	Std. Dev.	Mean	Std. Dev.		
Incentive						
Being female	0.50	0.50	0.50	0.50	-0.01	0.724
Age	44.53	14.24	44.90	14.48	0.37	0.611
Tertiary education degree	0.38	0.48	0.41	0.49	0.03	0.132
Living in West Germany	0.81	0.39	0.81	0.39	0.00	0.962
Net household income	10.41	4.72	10.64	4.73	0.22	0.325
Born in Germany	0.95	0.22	0.94	0.23	0.00	0.652
Migration background	0.15	0.36	0.16	0.36	0.01	0.634
Being a parent	0.56	0.50	0.55	0.50	-0.01	0.636
Being a teacher	0.04	0.20	0.04	0.18	-0.01	0.494
Being employed	0.70	0.46	0.71	0.45	0.01	0.509
Anchor						
Being female	0.50	0.50	0.50	0.50	0.00	0.960
Age	44.53	14.24	44.01	14.25	-0.52	0.475
Tertiary education degree	0.38	0.48	0.39	0.49	0.01	0.560
Living in West Germany	0.81	0.39	0.80	0.40	-0.01	0.690
Net household income	10.41	4.72	10.69	4.87	0.28	0.241
Born in Germany	0.95	0.22	0.94	0.24	-0.01	0.306
Migration background	0.15	0.36	0.16	0.36	0.01	0.585
Being a parent	0.56	0.50	0.53	0.50	-0.03	0.244
Being a teacher	0.04	0.20	0.05	0.21	0.01	0.560
Being employed	0.70	0.46	0.72	0.45	0.02	0.370
Experience						
Being female	0.50	0.50	0.50	0.50	-0.01	0.615
Age	44.53	14.24	44.79	14.28	0.25	0.668
Tertiary education degree	0.38	0.48	0.41	0.49	0.03*	0.067
Living in West Germany	0.81	0.39	0.81	0.40	-0.01	0.693
Net household income	10.41	4.72	10.49	4.74	0.08	0.664
Born in Germany	0.95	0.22	0.94	0.23	-0.01	0.497
Migration background	0.15	0.36	0.16	0.36	0.01	0.571
Being a parent	0.56	0.50	0.56	0.50	0.00	0.937
Being a teacher	0.04	0.20	0.04	0.20	0.00	0.819
Being employed	0.70	0.46	0.72	0.45	0.02	0.288

Note: Table summarizes weighted group means and standard deviations for three comparisons: control group vs. incentive treatment condition, control group vs. anchor treatment condition and control group vs. experience treatment condition. The net household income is indicated in income classes ranging from below €400 up to €5,000 and more (in steps of €250). Average income class 10 corresponds to a net household income ranging from €2,250 to €2,500. West Germany is a dummy equaling 1 if a respondent lives in a former West German state, and 0 if a respondent lives in a former East German state. Migration background is a dummy equaling 1 if both parents of respondents are not born in Germany, and 0 otherwise. Univ. entrance qualification is a dummy equaling 1 if a respondent has a higher education school degree, and 0 otherwise. Significance levels: *p<0.1; **p<0.05; ***p<0.01. Data source: ifo Education Survey 2020.

Table A4 Expert Sample

	N	Mean	St. Dev.	Min	Max
Being female	127	0.220	0.416	0	1
Age	66	53.197	11.554	32	74
Year of PhD	121	2002.347	11.378	1977	2019
<i>Field of Study</i>					
Education economics	116	0.078	0.269	0	1
Empirical economics	127	0.906	0.294	0	1
Experimental economics	127	0.205	0.405	0	1

Note: Table reports means and standard deviations for the following expert characteristics: gender, age, year of PhD and field of study. Data source: ifo Economists Panel 2020.

Table A5 Population Preferences and (Non)-Expert Predictions

	<i>Dep. var.: Preference for Increased Education Spending</i>				
	<i>Population Share</i>		<i>Prediction</i>		
	(1)	(2)	(3)	(4)	(5)
Information Provision	−0.183*** (0.022)	−0.185*** (0.022)	−0.060*** (0.007)	−0.060*** (0.007)	−0.160*** (0.018)
Information Acquisition	−0.053** (0.021)	−0.055*** (0.021)	−0.070*** (0.007)	−0.069*** (0.007)	−0.120*** (0.014)
Age		0.004*** (0.001)		0.001* (0.001)	
Female		−0.019 (0.018)		−0.017 (0.014)	
Tertiary education		0.088*** (0.019)		0.058*** (0.015)	
East		0.057** (0.023)		0.017 (0.019)	
Employed		0.032 (0.021)		−0.018 (0.016)	
High income		−0.011 (0.019)		0.004 (0.014)	
Migration background		−0.020 (0.033)		0.022 (0.026)	
Control group mean	0.774		0.471		0.648
Observations	3,133	3,120	3,072	3,069	381

Note: Survey-weighted linear regression estimates in columns 1 to 4. Unweighted OLS regression estimates for the expert sample in column 5. Results in each column are based on a different data set. In columns 1 and 2, the dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". We regress the dependent variable on a treatment variable indicating whether participants were assigned to the control, *information provision* or *information acquisition* treatment group. In column 2, we add control variables. The set of control variables contains a "Female" dummy that is a binary variable with value one for female respondents, "Age (years)" that indicates the age of the respondent, a "Tertiary education degree" dummy taking value one if the respondent has completed tertiary education (German "Abitur"), a "2nd Generation migrant" taking value 1 if both parents of the respondent were not born in Germany, an "Employed" dummy taking the value one if the respondent is employed full-time, part-time, or self-employed, a "High income" dummy taking value one if the respondent reports a net household income at or above EUR 3,000 per month, and an "East" dummy taking value one if the respondent lives in the respective region (former GDR). In columns 3 to 5, we regress the population and expert predictions on average population support in treatment and control conditions on the treatment indicator variable. We have three prediction observations per individual. The number of non-experts in column 3 is 1,024. The number of experts in column 5 is 127. Standard errors in columns 3 to 5 are clustered at individual level. Significance levels: *p<0.1; **p<0.05; ***p<0.01. Data sources: ifo Education Survey, ifo Economists Panel 2020

Table A6 Treatment Effects on Non-Expert Predictions

	<i>Dep. var.: Predictions on Preference for Increased Education Spending</i>					
	<i>No Information</i>		<i>Information Provision</i>		<i>Information Acquisition</i>	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment :						
Incentive	0.027** (0.012)	0.024** (0.012)	0.022** (0.010)	0.021** (0.010)	0.024** (0.011)	0.023** (0.011)
Anchor	0.059*** (0.013)	0.058*** (0.012)	0.050*** (0.011)	0.050*** (0.011)	0.051*** (0.011)	0.051*** (0.011)
Experience	0.012 (0.010)	0.010 (0.010)	0.006 (0.009)	0.005 (0.009)	0.022** (0.009)	0.020** (0.009)
Age		-0.0001 (0.0003)		0.001*** (0.0002)		0.0004* (0.0003)
Female		0.002 (0.007)		-0.010 (0.006)		-0.004 (0.006)
Tertiary education		0.077*** (0.008)		0.027*** (0.007)		0.048*** (0.007)
East		0.017* (0.010)		0.015* (0.008)		0.021** (0.009)
Employed		-0.013 (0.008)		-0.014* (0.007)		-0.013 (0.008)
High income		0.024*** (0.008)		0.024*** (0.007)		0.010 (0.007)
Migration background		-0.022* (0.012)		-0.006 (0.011)		0.002 (0.011)
Control group mean	0.471		0.411		0.402	
Observations	6,175	6,156	6,175	6,156	6,175	6,156

Note: Survey-weighted linear regression estimates. We regress population predictions about average population support for higher education spending on the treatment indicator variable. Population support is a dummy variable assuming the value 1 if participants respond public education spending should "increase" or "strongly increase". Sample sizes for experimental groups: incentive treatment (1,040), anchor treatment (1,019), experience treatment (3,092), control group (1,024). Respondents in all treatment arms make predictions for the three conditions in the information provision experiment: control (no information), information provision and information acquisition. In even columns, we add control variables. The set of control variables contains a "Female" dummy that is a binary variable with value one for female respondents, "Age (years)" that indicates the age of the respondent, a "Tertiary education degree" dummy taking value one if the respondent has completed tertiary education (German "Abitur"), a "2nd Generation migrant" taking value 1 if both parents of the respondent were not born in Germany, an "Employed" dummy taking the value one if the respondent is employed full-time, part-time, or self-employed, a "High income" dummy taking value one if the respondent reports a net household income at or above EUR 3,000 per month, and an "East" dummy taking value one if the respondent lives in the respective region (former GDR). Significance levels: *p<0.1; **p<0.05; ***p<0.01. Data sources: ifo Education Survey, ifo Economists Panel 2020

Table A7 Non-Expert Predictions: Heterogeneity by Education, Certainty Level and Spending Beliefs

<i>Dependent Variable: Preference for Increased Education Spending</i>				
	<i>Population Share</i>	<i>Non-Expert Prediction</i>		<i>Expert Prediction</i>
		<i>D=0</i>	<i>D=1</i>	
	(1)	(2)	(3)	(4)
Tertiary Education				
Information Provision	-0.183*** (0.022)	-0.036*** (0.009)	-0.099*** (0.011)	-0.160*** (0.018)
Information Acquisition	-0.053** (0.021)	-0.054*** (0.009)	-0.095*** (0.011)	-0.120*** (0.014)
Constant	0.774*** (0.015)	0.439*** (0.011)	0.525*** (0.013)	0.648*** (0.017)
Observations	3,133	1,770	1,302	381
Certainty Level				
Information Provision	-0.183*** (0.022)	-0.055*** (0.009)	-0.069*** (0.011)	-0.160*** (0.018)
Information Acquisition	-0.053** (0.021)	-0.061*** (0.009)	-0.083*** (0.011)	-0.120*** (0.014)
Constant	0.774*** (0.015)	0.451*** (0.011)	0.504*** (0.014)	0.648*** (0.017)
Observations	3,133	1,857	1,215	381
Spending Beliefs				
Information Provision	-0.183*** (0.022)	-0.060*** (0.009)	-0.061*** (0.011)	-0.160*** (0.018)
Information Acquisition	-0.053** (0.021)	-0.075*** (0.009)	-0.062*** (0.011)	-0.120*** (0.014)
Constant	0.774*** (0.015)	0.495*** (0.011)	0.437*** (0.014)	0.648*** (0.017)
Observations	3,133	1,845	1,221	381

Note: Survey-weighted linear regression estimates in columns 1 to 3. Unweighted OLS regression estimates for the expert sample in column 4. Split of non-expert sample by tertiary education (upper panel), prediction certainty (middle panel) and spending beliefs (lower panel). Column 2 (3) shows regression results for the subsample dummy = 0 (= 1). The tertiary education dummy is 1 if an individual has a tertiary education degree and 0 otherwise. The certainty dummy is 1 if the individual has high prediction certainty, and 0 otherwise. In the lower panel, column 3 (2) summarizes regression results for respondents with an absolute difference between believed and actual education spending above (below) median. Results in columns 1, 2 and 3/4 are based on a different data set. In column 1, the dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". We regress the dependent variable on a treatment variable indicating whether participants were assigned to the control, *information provision* or *information acquisition* treatment group. In columns 2 to 4, we regress the population and expert predictions about average population support in treatment and control conditions on the treatment indicator variable. In columns 2 to 4, we have three observations per individual. The number of non-experts in column 2 (3) is 619 (405). The number of experts in column 3 is 127. The control group is the reference category. Standard errors in columns 2 to 4 are clustered at individual level. Significance levels: *p<0.1; **p<0.05; ***p<0.01. Data sources: ifo Education Survey, ifo Economists Panel 2020

Table A8 Expert Predictions: Heterogeneity by Field of Expertise, Certainty Level and Spending Beliefs

<i>Dependent Variable: Preference for Increased Education Spending</i>				
	<i>Population Share</i>	<i>Non-Expert Prediction</i>	<i>Expert Prediction</i>	
	(1)	(2)	<i>D=0</i>	<i>D=1</i>
	(1)	(2)	(3)	(4)
Education Focus				
Information Provision	-0.183*** (0.022)	-0.060*** (0.007)	-0.157*** (0.019)	-0.196*** (0.062)
Information Acquisition	-0.053** (0.021)	-0.070*** (0.007)	-0.117*** (0.015)	-0.154*** (0.052)
Constant	0.774*** (0.015)	0.471*** (0.009)	0.646*** (0.018)	0.673*** (0.054)
Observations	3,133	3,072	354	27
Certainty Level				
Information Provision	-0.183*** (0.022)	-0.060*** (0.007)	-0.169*** (0.021)	-0.145*** (0.033)
Information Acquisition	-0.053** (0.021)	-0.070*** (0.007)	-0.123*** (0.016)	-0.115*** (0.026)
Constant	0.774*** (0.015)	0.471*** (0.009)	0.651*** (0.020)	0.643*** (0.032)
Observations	3,133	3,072	237	144
Spending Beliefs				
Information Provision	-0.183*** (0.022)	-0.060*** (0.007)	-0.164*** (0.019)	-0.140*** (0.038)
Information Acquisition	-0.053** (0.021)	-0.070*** (0.007)	-0.121*** (0.016)	-0.107*** (0.028)
Constant	0.774*** (0.015)	0.471*** (0.009)	0.646*** (0.021)	0.643*** (0.032)
Observations	3,133	3,072	234	132

Note: Survey-weighted linear regression estimates in columns 1 and 2. Unweighted OLS regression estimates for the expert sample in columns 3 and 4. Split of expert sample by education focus (upper panel) and prediction certainty (lower panel). Column 3 (4) shows regression results for the subsample dummy = 0 (= 1). The education focus dummy is 1 if an expert has education economics of field of research and 0 otherwise. The certainty dummy is 1 if the expert has high prediction certainty, and 0 otherwise. In the lower panel, column 4(3) summarizes regression results for experts with an absolute difference between believed and actual education spending above (below) median. Results in columns 1, 2 and 3 to 4 are based on a different data set. In column 1, the dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". We regress the dependent variable on a treatment variable indicating whether participants were assigned to the control, *information provision* or *information acquisition* treatment group. In columns 2 to 4, we regress the population and expert predictions about average population support in treatment and control conditions on the treatment indicator variable. In columns 2 to 4, we have three observations per individual. The number of experts in column 3 (4) is 118 (9) in the upper panel and 79 (48) in the lower panel. The control group is the reference category. Standard errors in columns 2 to 4 are clustered at individual level. Significance levels: *p<0.1; **p<0.05; ***p<0.01. Data sources: ifo Education Survey, ifo Economists Panel 2020

Table A9 Examples Coding Scheme

Category	Examples
Knowledge/Deduction	
Deduction	'I thought that in the first case, people would assume that more should be spent on pupils, as there is widespread talk of problems in schools. In the second case, it would be less, as the expenditure probably seems high, and in the third case perhaps even less, as people may have been even more surprised at how high the expenditure is when they clicked on it.'
Knowledge	'Overall, spending on education in Germany is far too low. People recognise that conditions for good education in schools are getting worse and worse... lots of career changers, less qualified teachers, large classes, high proportion of foreigners, migrants, inclusion... lots of reforms in schools... Basically, there is a lack of money! Now, there are certainly people who believe that spending over €7,000 per pupil would be sufficient.'
Gut feeling	
Gut feeling	'I just listened to my gut feeling [...]', 'Purely a matter of feeling and thinking about how I myself would decide spontaneously', 'I looked at what I would estimate or how I would react myself, and then I just adjusted that a little bit, so it's rather inaccurate.', 'I didn't actually calculate anything.', 'Purely emotional assessment'
Other	
News	'Utilisation of information from digital and print media'
Personal situation	'From personal experience [...]', 'From conversations with my friends and family'
Other	'I think it is necessary to invest more in education [...]'
Don't Know	'I can't explain it'

Note: Table summarizes the coding strategy for the open-text question: 'Next, we would like to know how you arrived at your answers to the estimates in the last question. Please briefly describe as accurately as possible your thought process in arriving at your answer.'. Answers are manually assigned to one of the following codes: knowledge, deduction, gut feeling, news, personal situation, don't know and other. We cluster the different codes to three categories: 'Knowledge/Deduction', 'Gut Feeling' and 'Other' Data source: ifo Education Survey 2020.

Table A10 Non-Expert Predictions: Heterogeneity by Answer Strategy

<i>Dependent Variable: Preference for Increased Education Spending</i>					
	<i>Population Share</i>	<i>Non-Expert Prediction</i>			<i>Expert Prediction</i>
		<i>Gut Feeling</i>	<i>Other</i>	<i>Knowledge/Deduction</i>	
	(1)	(2)	(3)	(4)	(5)
Information Provision	-0.183*** (0.022)	-0.042*** (0.010)	-0.045*** (0.016)	-0.105*** (0.014)	-0.160*** (0.018)
Information Acquisition	-0.053** (0.021)	-0.054*** (0.010)	-0.057*** (0.017)	-0.108*** (0.014)	-0.120*** (0.014)
Constant	0.774*** (0.015)	0.444*** (0.012)	0.451*** (0.024)	0.566*** (0.015)	0.648*** (0.017)
Observations	3,133	1,395	537	888	381

Note: Survey-weighted linear regression estimates in columns 1 to 4. Unweighted OLS regression estimates for the expert sample in column 5. Split of non-expert sample by answer strategy. Column 2 - 4 show regression results for the three different answer strategies: gut feeling, other and knowledge/deduction. Results in columns 1, 2 - 4 and 5 are based on different data sets. In column 1, the dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". We regress the dependent variable on a treatment variable indicating whether participants were assigned to the control, *information provision* or *information acquisition* treatment group. In columns 2 to 5, we regress the population and expert predictions about average population support in treatment and control conditions on the treatment indicator variable. In columns 2 to 5, we have three observations per individual. Number of respondents in columns 2 to 4: 465, 179 and 296. The number of experts in column 3 is 127. The control group is the reference category. Standard errors in columns 2 to 5 are clustered at individual level. Significance levels: *p<0.1; **p<0.05; ***p<0.01. Data sources: ifo Education Survey, ifo Economists Panel 2020

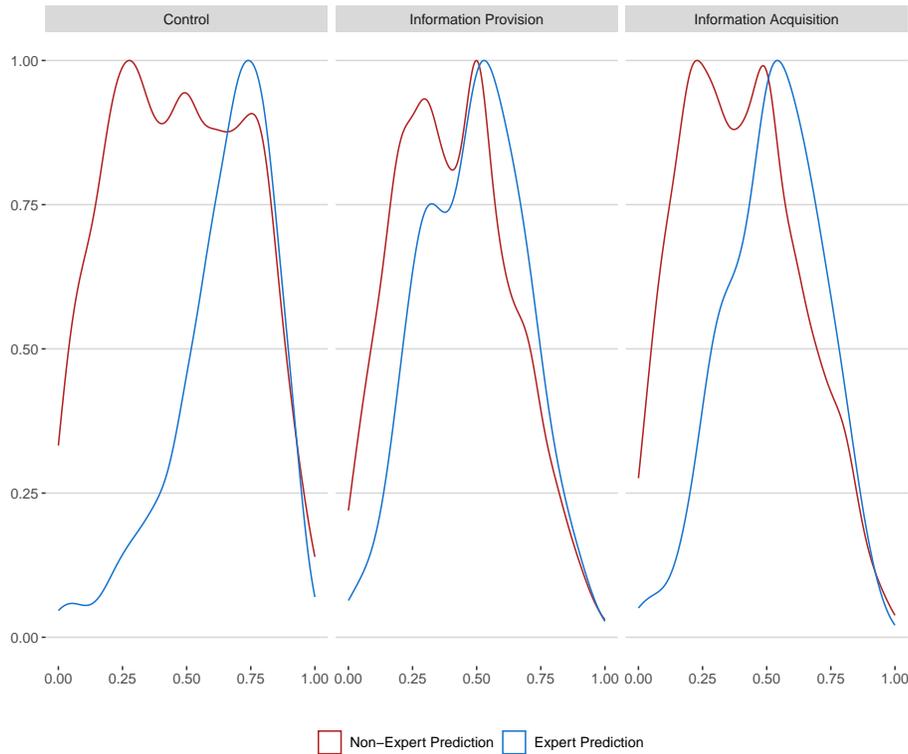
Table A11 Expert Predictions: Heterogeneity by Answer Strategy

	<i>Dependent Variable: Preference for Increased Education Spending</i>				
	<i>Population Share</i>	<i>Non-Expert Prediction</i>	<i>Expert Prediction</i>		
			<i>Other</i>	<i>Gut Feeling</i>	<i>Knowledge/Deduction</i>
	(1)	(2)	(3)	(4)	(5)
Information Provision	-0.183*** (0.022)	-0.060*** (0.007)	-0.176** (0.067)	-0.211*** (0.038)	-0.183*** (0.029)
Information Acquisition	-0.053** (0.021)	-0.070*** (0.007)	-0.129** (0.055)	-0.146*** (0.027)	-0.133*** (0.025)
Constant	0.774*** (0.015)	0.471*** (0.009)	0.690*** (0.087)	0.711*** (0.044)	0.664*** (0.027)
Observations	3,133	3,072	21	45	147

Note: Survey-weighted linear regression estimates in columns 1 and 2. Unweighted OLS regression estimates for the expert sample in columns 3 to 5. Split of expert sample by answer strategy. Column 3 - 5 show regression results for the three different answer strategies: other, gut feeling and knowledge/deduction. Results in columns 1, 2 and 3 - 5 are based on different data sets. In column 1, the dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". We regress the dependent variable on a treatment variable indicating whether participants were assigned to the control, *information provision* or *information acquisition* treatment group. In columns 2 to 5, we regress the population and expert predictions about average population support in treatment and control conditions on the treatment indicator variable. In columns 2 to 5, we have three observations per individual. Note that an expert answer can get assigned more than one answer strategy. Number of experts in columns 3 to 5: 7, 15 and 49. The control group is the reference category. Standard errors in columns 2 to 5 are clustered at individual level. Significance levels: *p<0.1; **p<0.05; ***p<0.01. Data sources: ifo Education Survey, ifo Economists Panel 2020

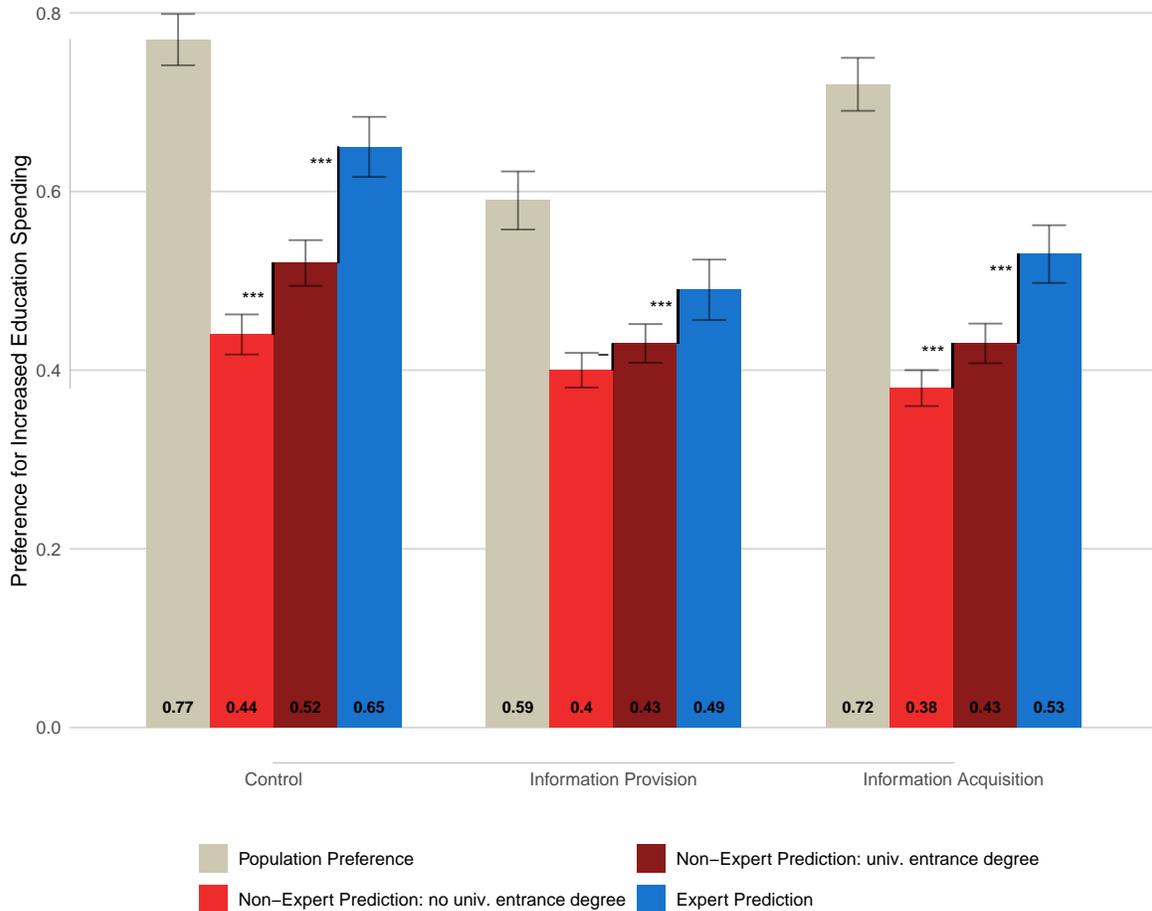
C Supplementary Figures

Figure B1 Population Preferences and (Non)-Expert Predictions



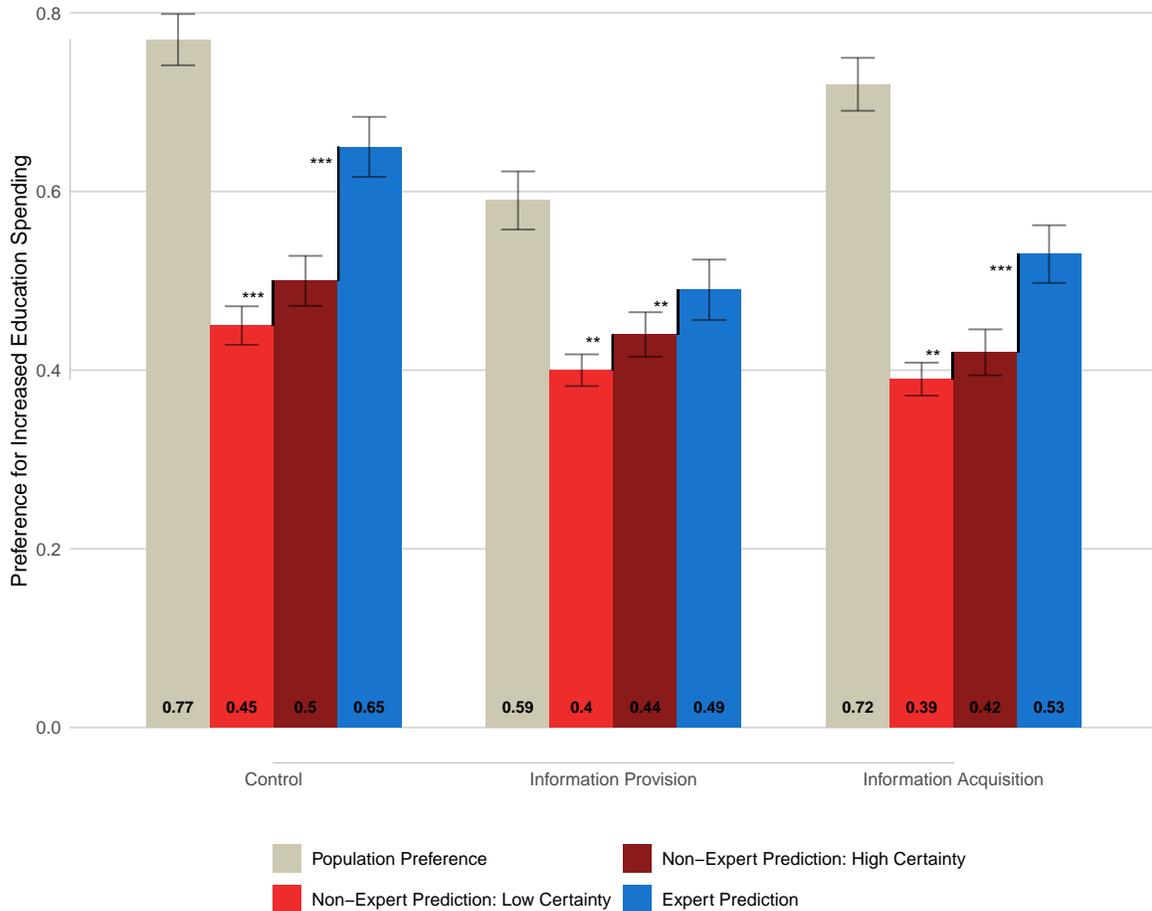
Note: The Figure shows densities for population preferences and non-expert and expert predictions. Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) are asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the information provision condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context: "The aim now is to estimate the answers of other participants in this survey as accurately as possible. Some other participants are randomly divided into three groups and asked about government spending on schools". Then, the three different questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The people are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?. Population and non-expert sample weighted by survey weights. Data sources: ifo Education Survey, ifo Economists Panel 2020

Figure B2 Non-Expert Predictions split by Education



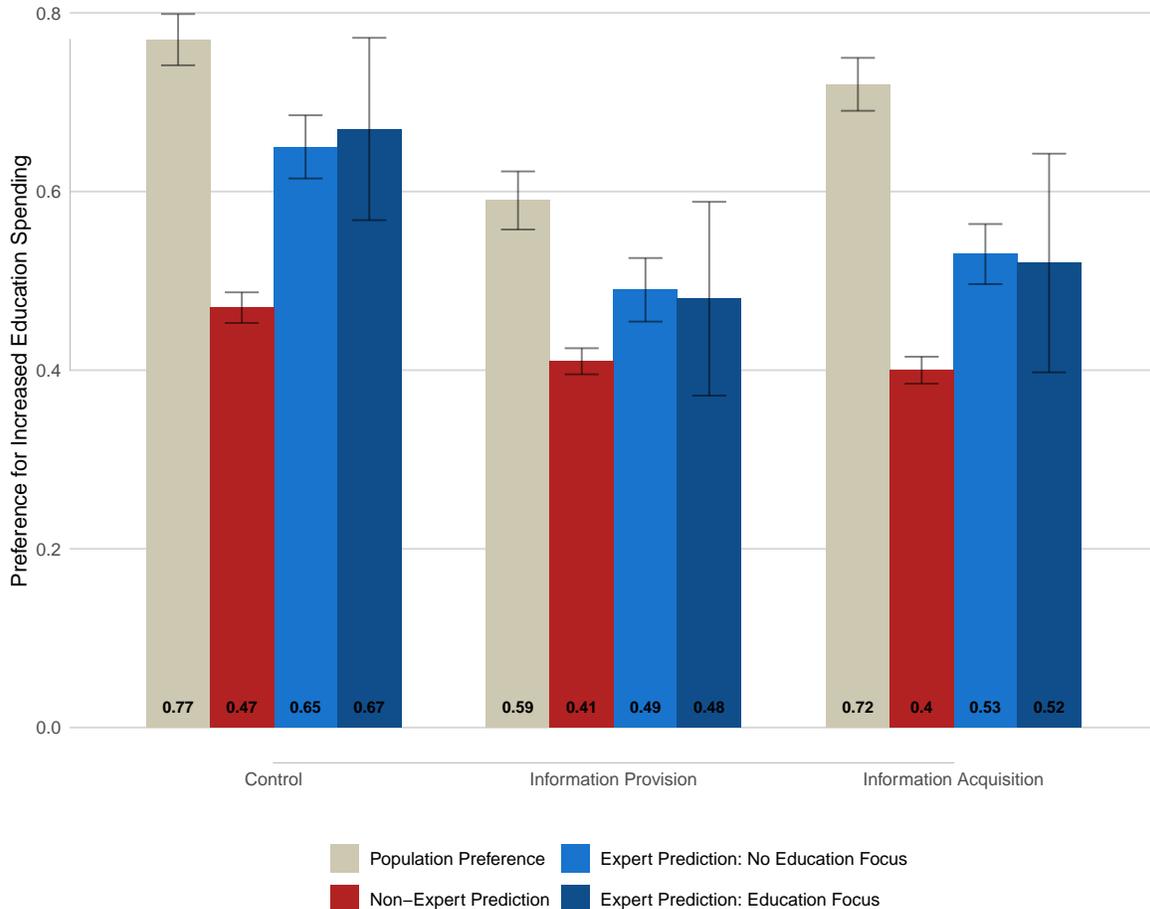
Note: The Figure shows means and confidence intervals for population preferences and non-expert and expert predictions. Population preferences are separately calculated for participants with and without university entrance degree. Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) are asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the information provision condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context: "The aim now is to estimate the answers of other participants in this survey as accurately as possible. Some other participants are randomly divided into three groups and asked about government spending on schools." Then, the three different questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The people are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?. Population and non-expert sample weighted by survey weights. Significance levels: *p<0.1; **p<0.05; ***p<0.01. Data sources: ifo Education Survey, ifo Economists Panel 2020

Figure B3 Non-Expert Predictions split by Prediction Certainty



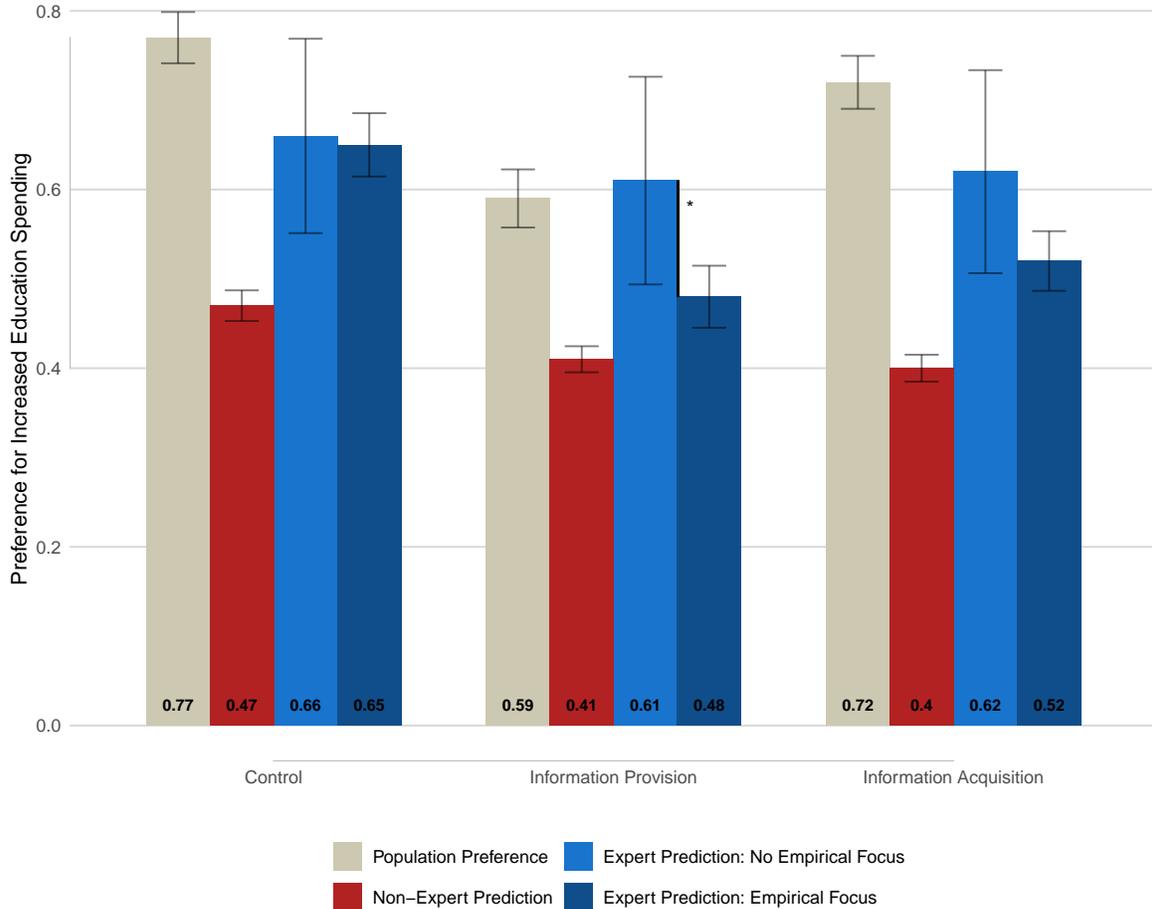
Note: The Figure shows means and confidence intervals for population preferences and non-expert and expert predictions. Population predictions are separately calculated for different prediction certainty levels. We form two categories: low (levels 1, 2, and 3) and high (levels 4, 5, 6 and 7). Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) are asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the information provision condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context: "The aim now is to estimate the answers of other participants in this survey as accurately as possible. Some other participants are randomly divided into three groups and asked about government spending on schools." Then, the three different questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The people are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?" Population and non-expert sample weighted by survey weights. Significance levels: *p<0.1; **p<0.05; ***p<0.01. Data sources: ifo Education Survey, ifo Economists Panel 2020

Figure B4 Expert Predictions split by Field of Expertise: Education Economics



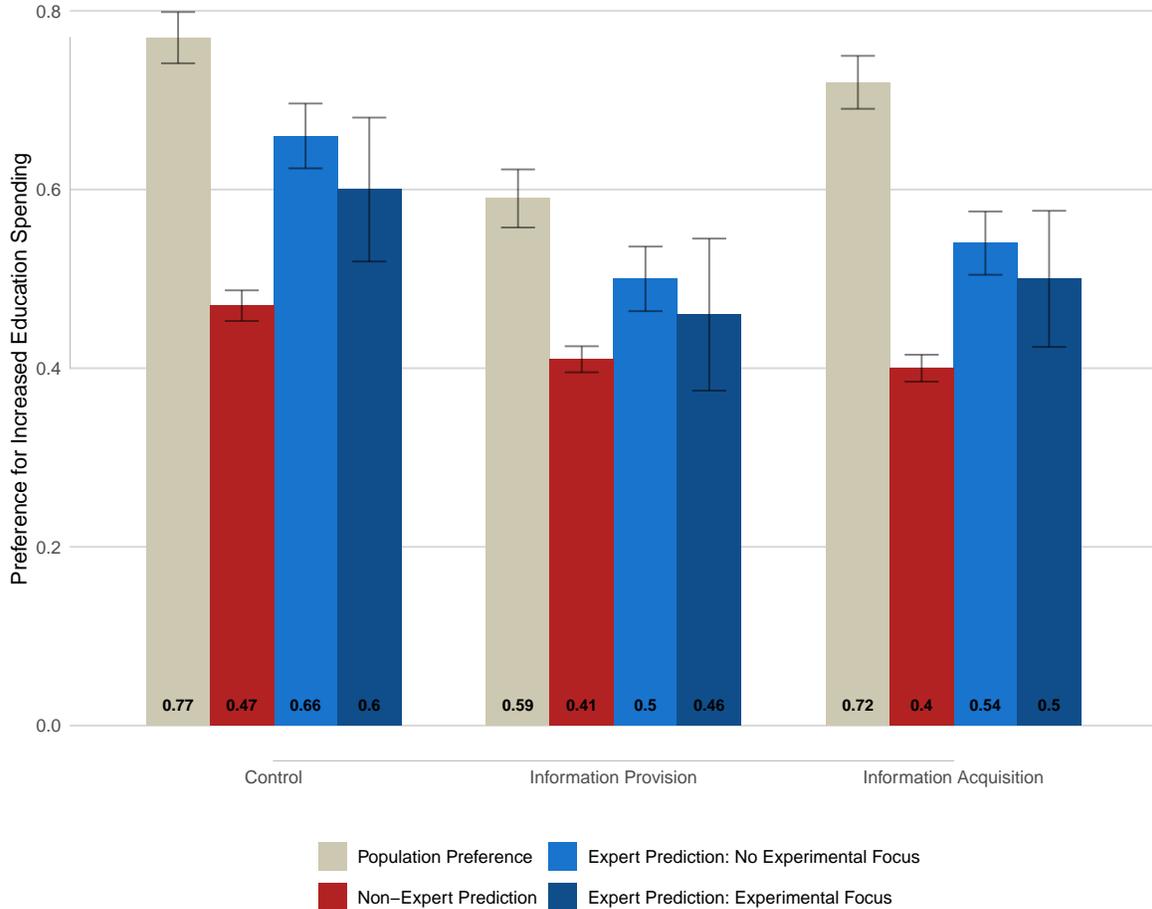
Note: The Figure shows means and confidence intervals for population preferences and non-expert and expert predictions. Expert predictions are separately calculated for experts having/not having a research focus in education economics. Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) are asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the information provision condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context: "The aim now is to estimate the answers of other participants in this survey as accurately as possible. Some other participants are randomly divided into three groups and asked about government spending on schools." Then, the three different questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The people are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?. Population and non-expert sample weighted by survey weights. Data sources: ifo Education Survey, ifo Economists Panel 2020

Figure B5 Expert Predictions split by Field of Expertise: Empirical Economics



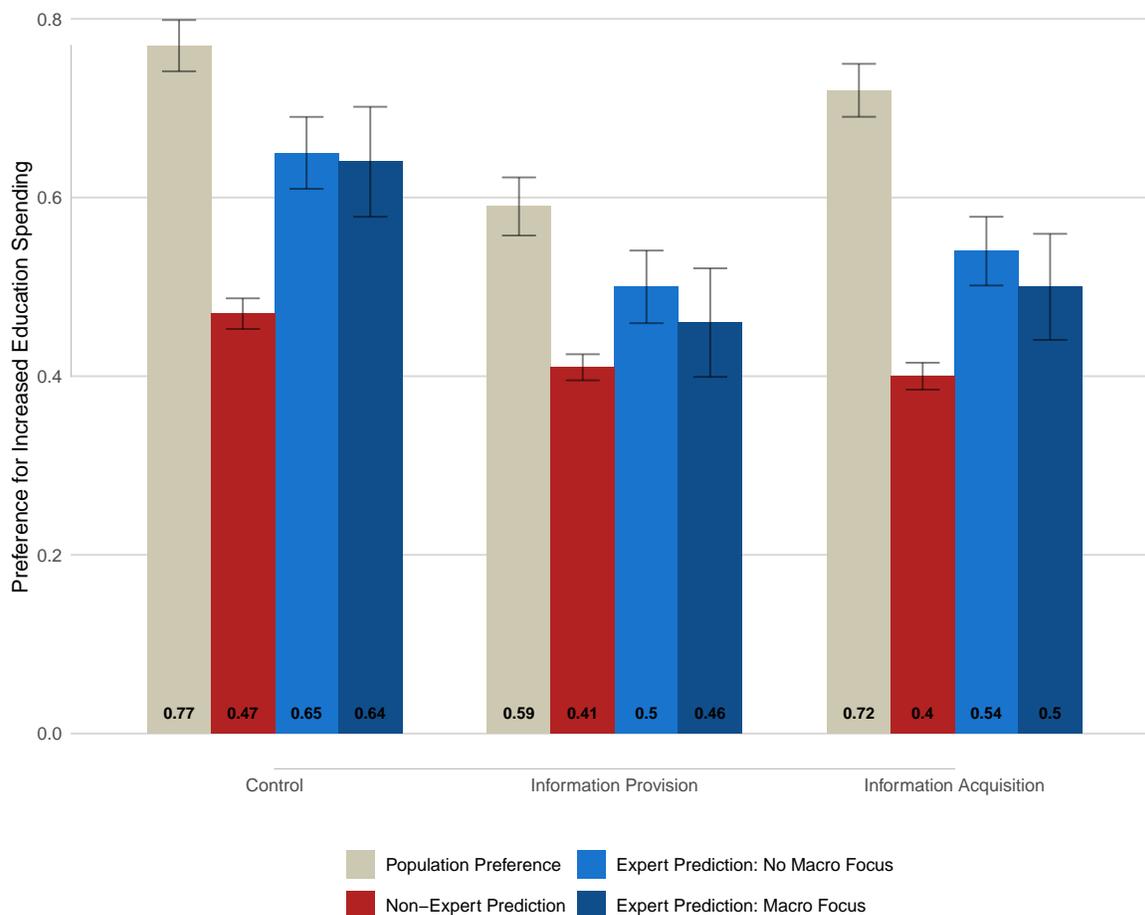
Note: The Figure shows means and confidence intervals for population preferences and non-expert and expert predictions. Expert predictions are separately calculated for experts having/not having a research focus in empirical economics. Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) are asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the information provision condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context: "The aim now is to estimate the answers of other participants in this survey as accurately as possible. Some other participants are randomly divided into three groups and asked about government spending on schools." Then, the three different questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The people are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?. Population and non-expert sample weighted by survey weights. Significance levels: *p<0.1; **p<0.05; ***p<0.01. Data sources: ifo Education Survey, ifo Economists Panel 2020

Figure B6 Expert Predictions split by Field of Expertise: Experimental Economics



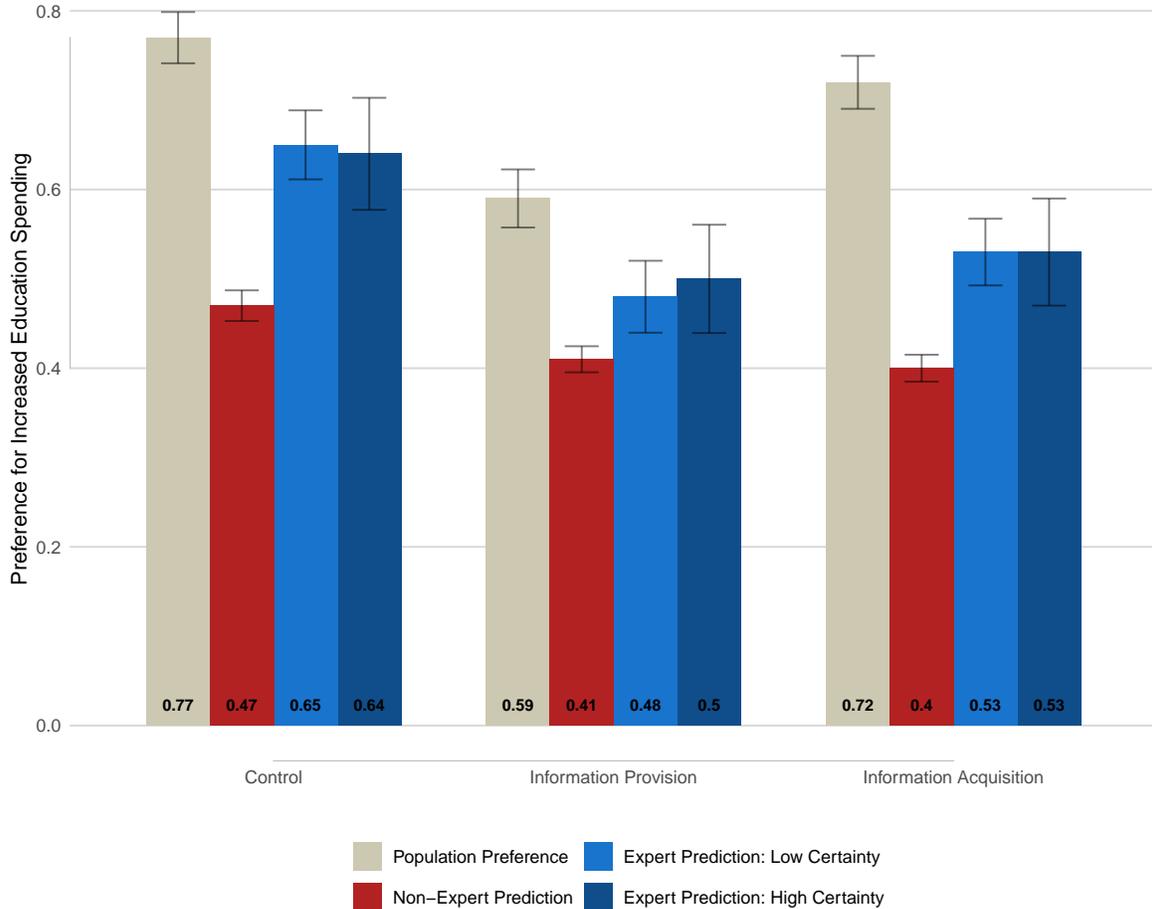
Note: The Figure shows means and confidence intervals for population preferences and non-expert and expert predictions. Expert predictions are separately calculated for experts having/not having a research focus in experimental economics. Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) are asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the *information provision* condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context: "The aim now is to estimate the answers of other participants in this survey as accurately as possible. Some other participants are randomly divided into three groups and asked about government spending on schools." Then, the three different questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The people are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?. Population and non-expert sample weighted by survey weights. Data sources: ifo Education Survey, ifo Economists Panel 2020

Figure B7 Expert Predictions split by Field of Expertise: Macroeconomics



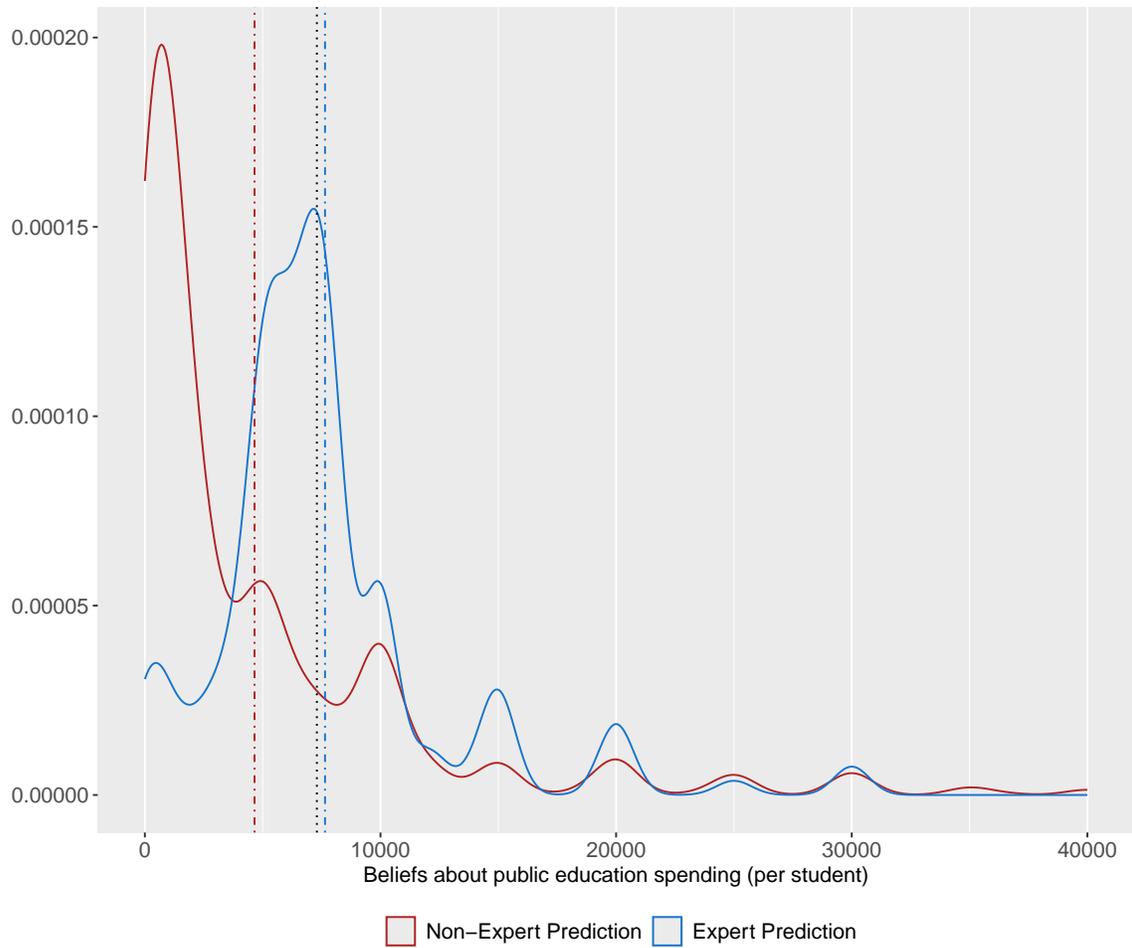
Note: The Figure shows means and confidence intervals for population preferences and non-expert and expert predictions. Expert predictions are separately calculated for experts having/not having a research focus in macroeconomics. Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) are asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the information provision condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context: "The aim now is to estimate the answers of other participants in this survey as accurately as possible. Some other participants are randomly divided into three groups and asked about government spending on schools." Then, the three different questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The people are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?. Population and non-expert sample weighted by survey weights. Data sources: ifo Education Survey, ifo Economists Panel 2020

Figure B8 Expert Predictions split by Prediction Certainty



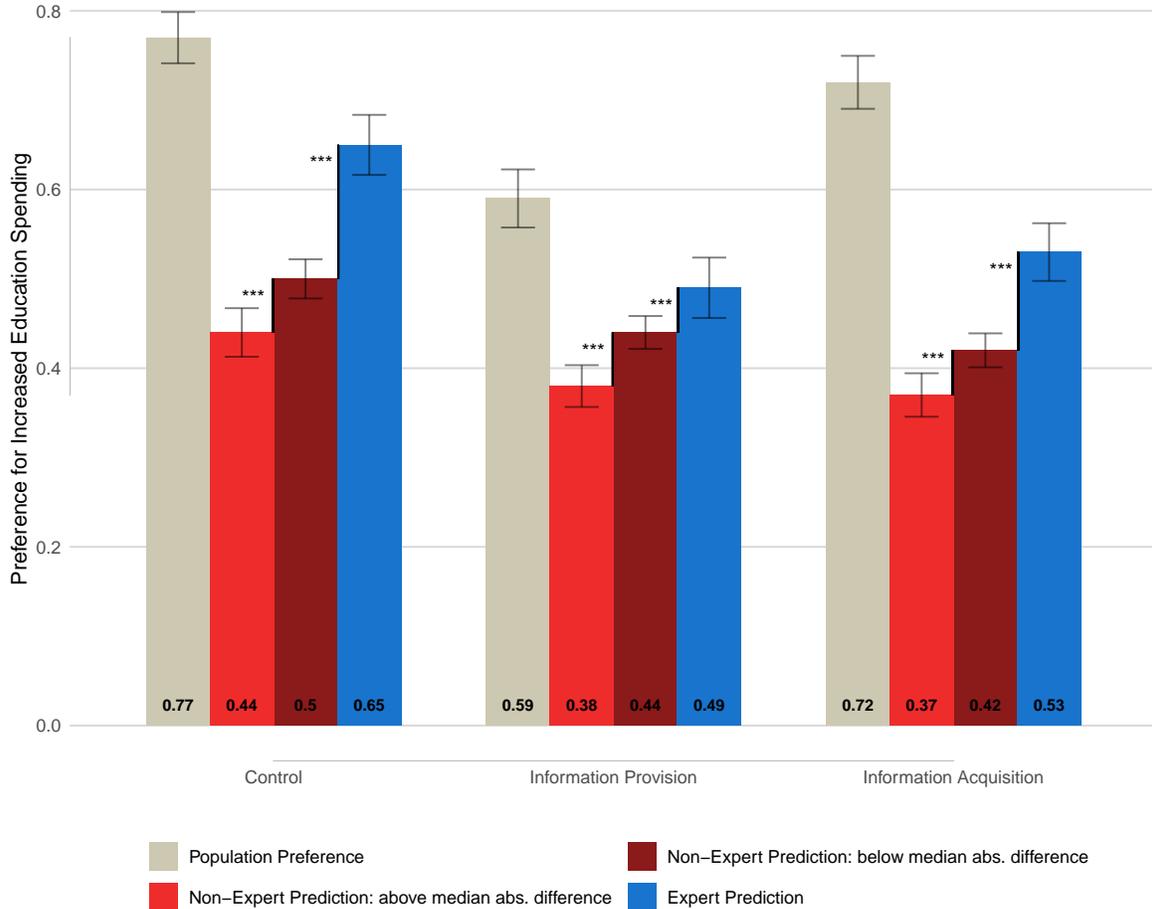
Note: The Figure shows means and confidence intervals for population preferences and non-expert and expert predictions. Expert predictions are separately calculated for different prediction certainty levels. We examine the two categories: low (levels 1, 2, and 3) and high (levels 4, 5, 6 and 7). Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) are asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the information provision condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context: "The aim now is to estimate the answers of other participants in this survey as accurately as possible. Some other participants are randomly divided into three groups and asked about government spending on schools." Then, the three different questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The people are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?" Population and non-expert sample weighted by survey weights. Significance levels: *p<0.1; **p<0.05; ***p<0.01. Data sources: ifo Education Survey, ifo Economists Panel 2020

Figure B9 Distribution of Spending Belief Differences



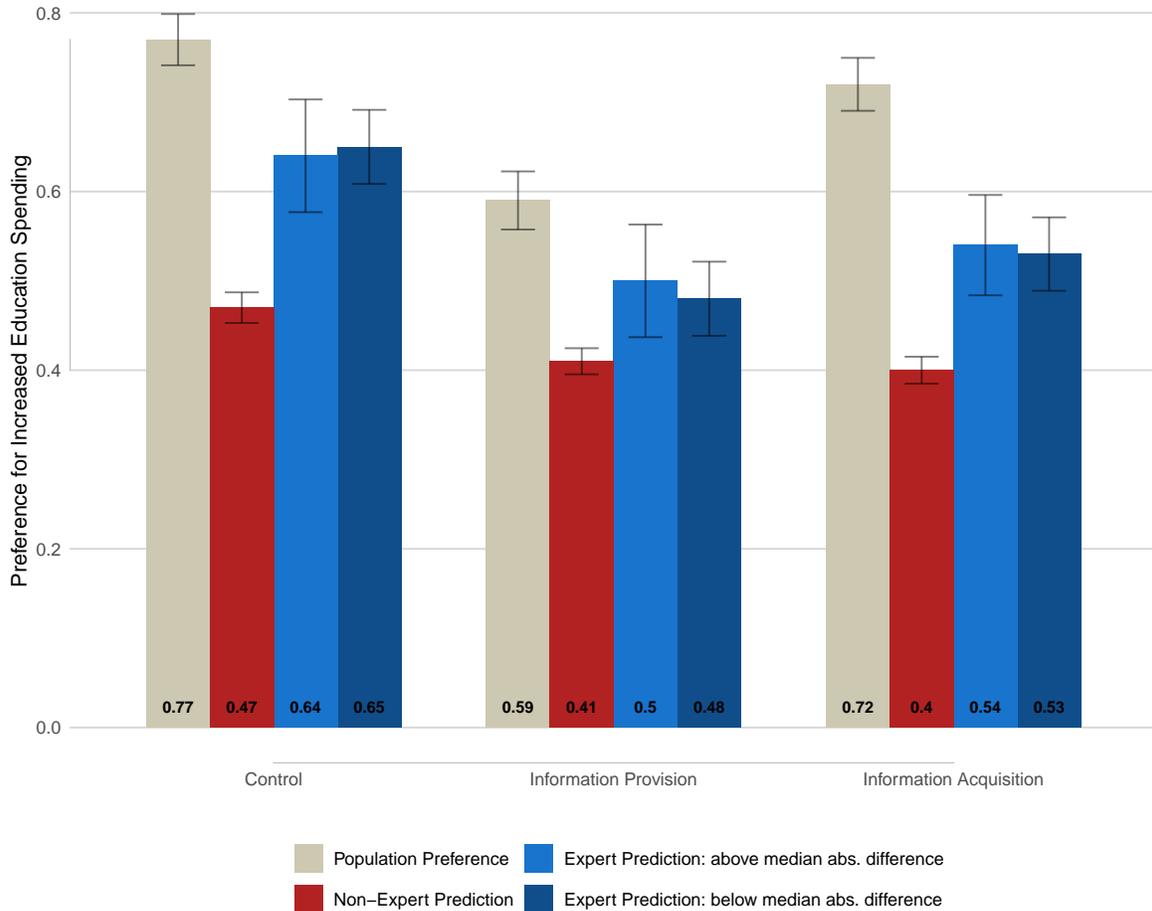
Note: Figure depicts the distribution spending beliefs for experts and non-experts, as well as actual education spending, depicted via the black dotted vertical line. We exclude observations above the 99th percentile as outliers. The true value of per student education spending is 7,300 Euro. In the expert sample, the mean is 7,646 Euro (standard deviation is 5,077 Euro and median is 7,000 Euro). In the non-expert sample, the mean is 4,685 Euros (standard deviation is 6,565 Euros and median is 2,000 Euros). Data sources: ifo Education Survey, ifo Economists Panel 2020.

Figure B10 Non-Expert Predictions split by Prior Beliefs



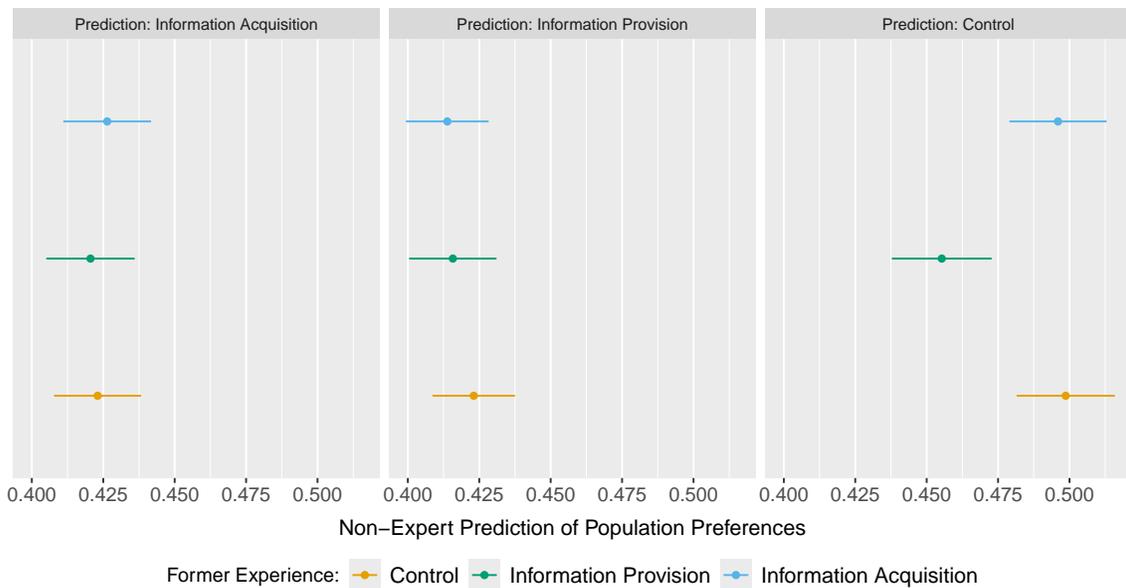
Note: The Figure shows means and confidence intervals for population preferences as well as non-expert and expert predictions. Population predictions are separately calculated for participants according to their education spending beliefs. We take the absolute difference between respondents' spending belief and actual education spending (EUR 7,300 per student) and create a dummy variable that takes value 1 if the absolute difference is above the median difference, and 0 otherwise. The median difference in education spending belief is 6,300 Euros for non-experts. Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) are asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the information provision condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context: "The aim now is to estimate the answers of other participants in this survey as accurately as possible. Some other participants are randomly divided into three groups and asked about government spending on schools.". Then, the three different questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The people are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?. Population and non-expert sample weighted by survey weights. Significance levels: *p<0.1; **p<0.05; ***p<0.01. Data sources: ifo Education Survey, ifo Economists Panel 2020

Figure B11 Expert Predictions split by Education Spending Beliefs



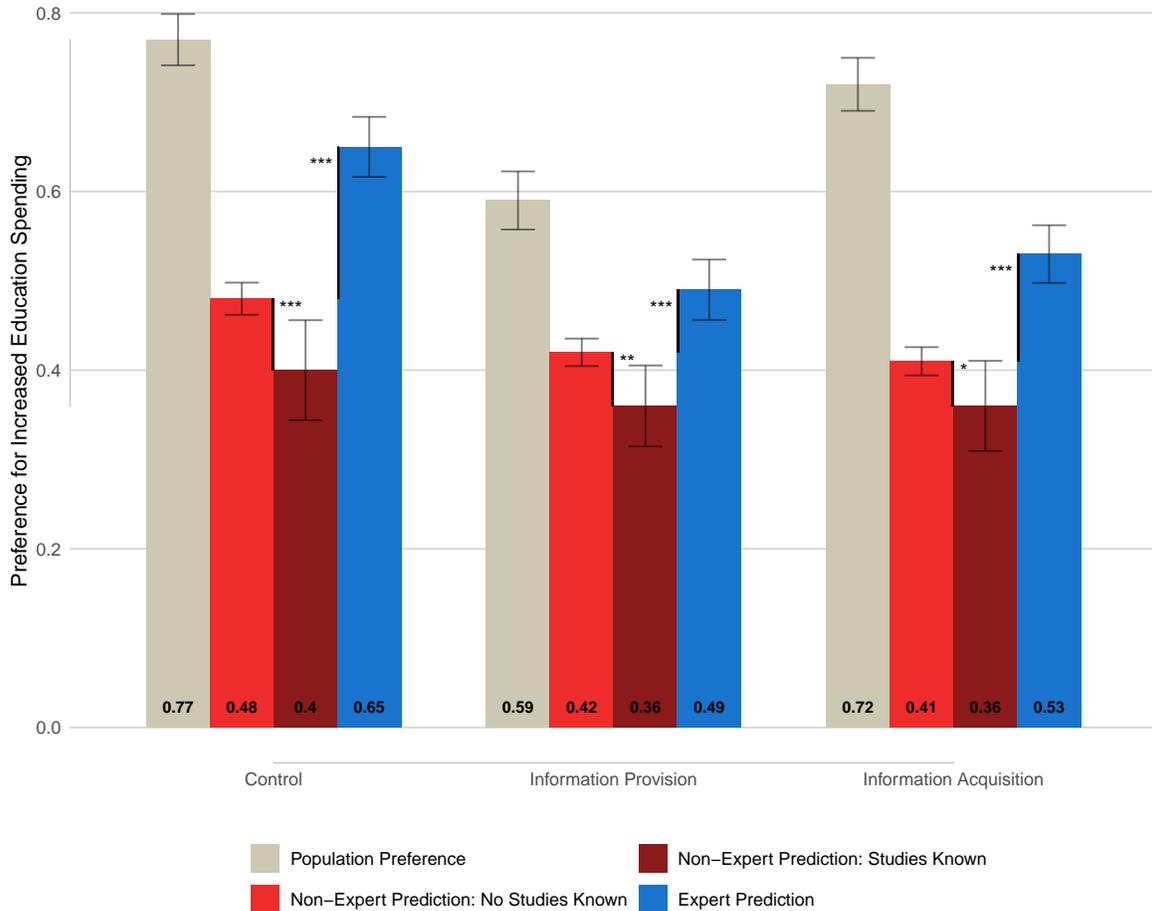
Note: The Figure shows means and confidence intervals for population preferences as well as non-expert and expert predictions. Expert predictions are separately calculated for experts being above/below the median of the absolute difference in education spending. The absolute spending difference is experts' spending beliefs minus actual education spending. Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) are asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the information provision condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context: "The aim now is to estimate the answers of other participants in this survey as accurately as possible. Some other participants are randomly divided into three groups and asked about government spending on schools.". Then, the three different questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The people are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?. Population and non-expert sample weighted by survey weights. Data sources: ifo Education Survey, ifo Economists Panel 2020

Figure B12 Predictions of Non-Experts from the Experience Group



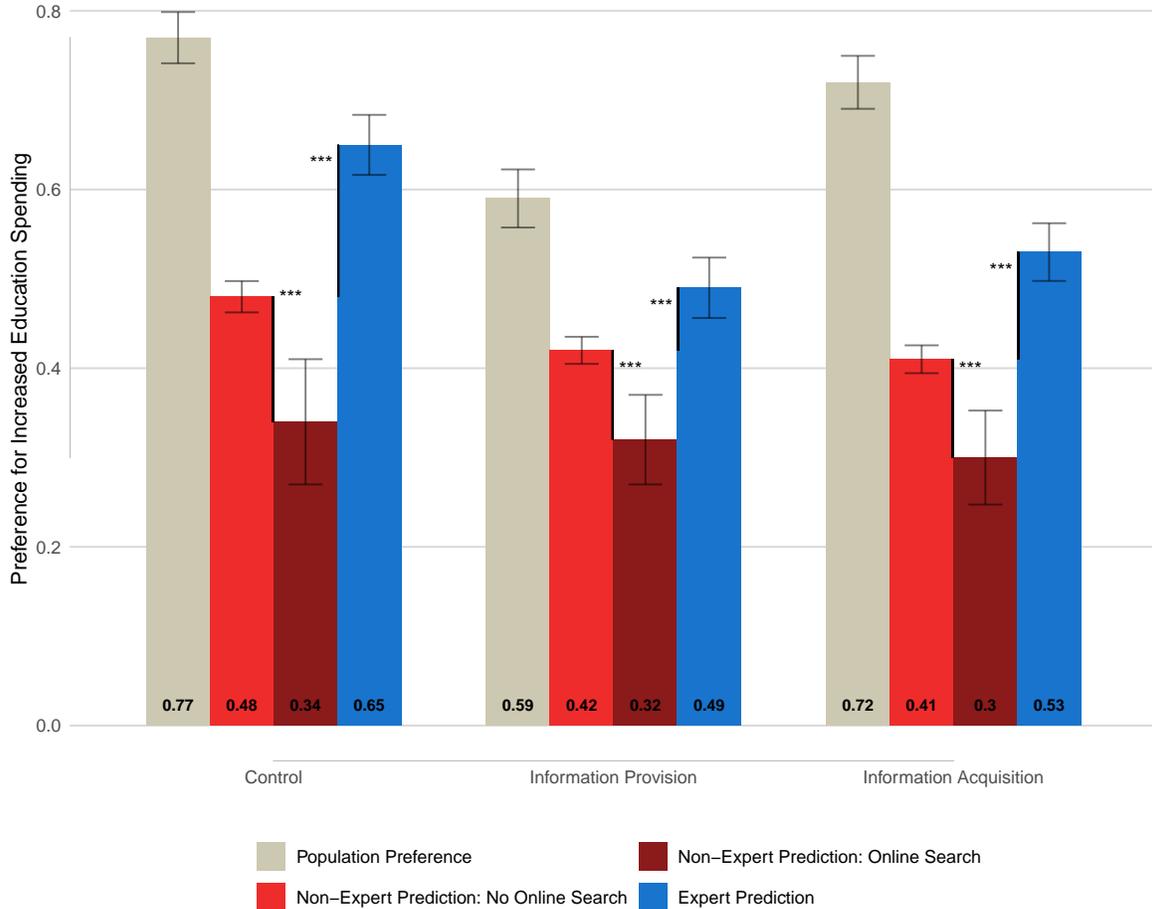
Note: Figure shows treatment coefficients and confidence intervals for experience group participants. In the experience treatment, participants take part in the information provision experiment and indicate their preference for or against increased education spending in either one of the three experimental groups: control, information provision, information acquisition. Participants are then asked to predict preferences for the other survey participants for the three different conditions: control, information provision, information acquisition. Each panel in the figure depicts treatment effects for one prediction. Survey weights are applied. Data sources: ifo Education Survey

Figure B13 Non-expert Predictions split by Knowledge of Relevant Studies



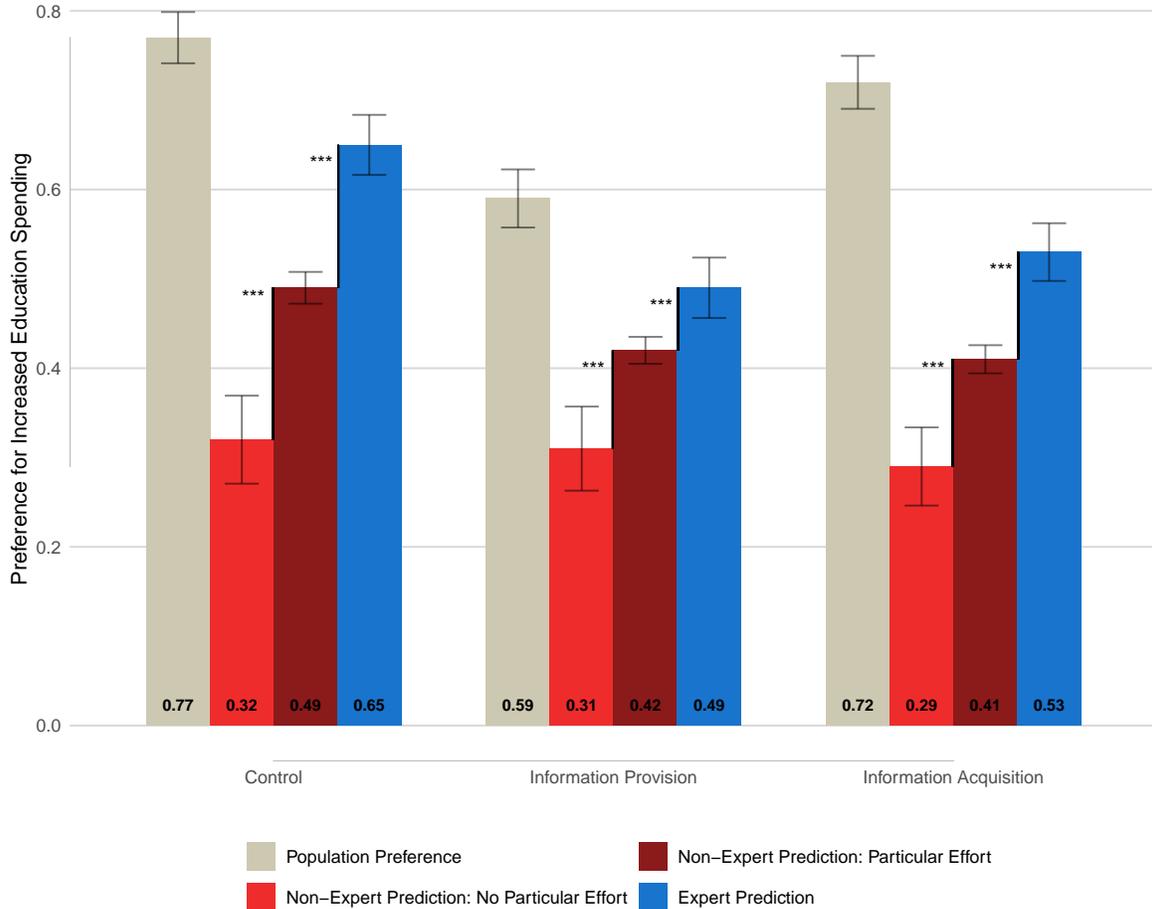
Note: The Figure shows means and confidence intervals for population preferences and non-expert and expert predictions. Non-expert predictions are separately calculated for experts rather or strongly agreeing to know (10%)/not know (90%) scientific studies focusing on support of education spending. The question asks: "Now we would like to revisit how you approached the previous estimation questions. To what extent do the following statements apply to you? I am familiar with scientific studies on support for education spending and based my answers on them." Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) are asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the information provision condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context. Then, the three different questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The people are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?. Population and non-expert sample weighted by survey weights. Data sources: ifo Education Survey, ifo Economists Panel 2020

Figure B14 Non-expert Predictions split by Search for Information



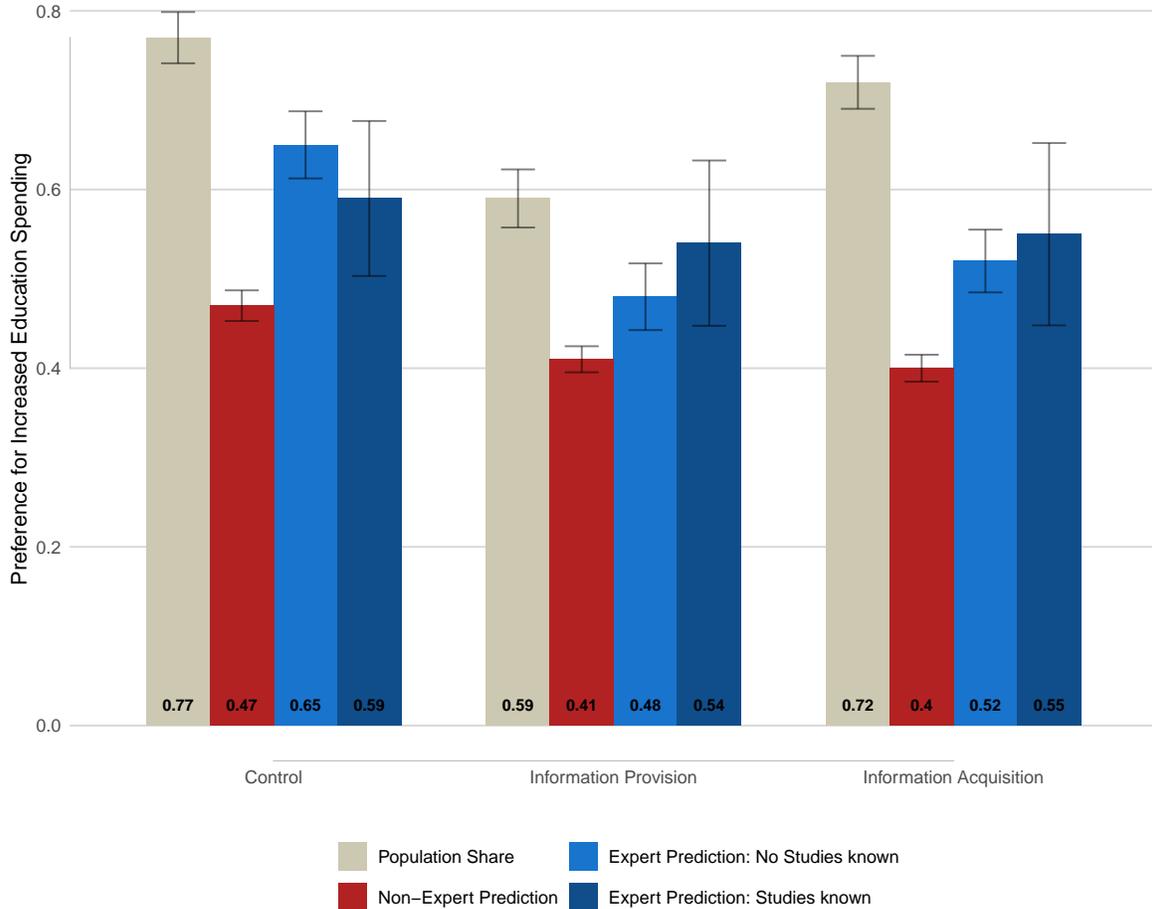
Note: The Figure shows means and confidence intervals for population preferences and non-expert and expert predictions. Non-expert predictions are separately calculated for respondents rather or strongly agreeing to search (7%)/not search (93%) for information online. The question asks: "Now we will look again at how you approached the previous estimation questions. To what extent do the following statements apply to you? I searched for information on the Internet in order to make the best possible estimate." Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) are asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the information provision condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context. Then, the three different questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The people are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?" Population and non-expert sample weighted by survey weights. Significance levels: *p<0.1; **p<0.05; ***p<0.01. Data sources: ifo Education Survey, ifo Economists Panel 2020

Figure B15 Non-Expert Predictions split by Particular Effort



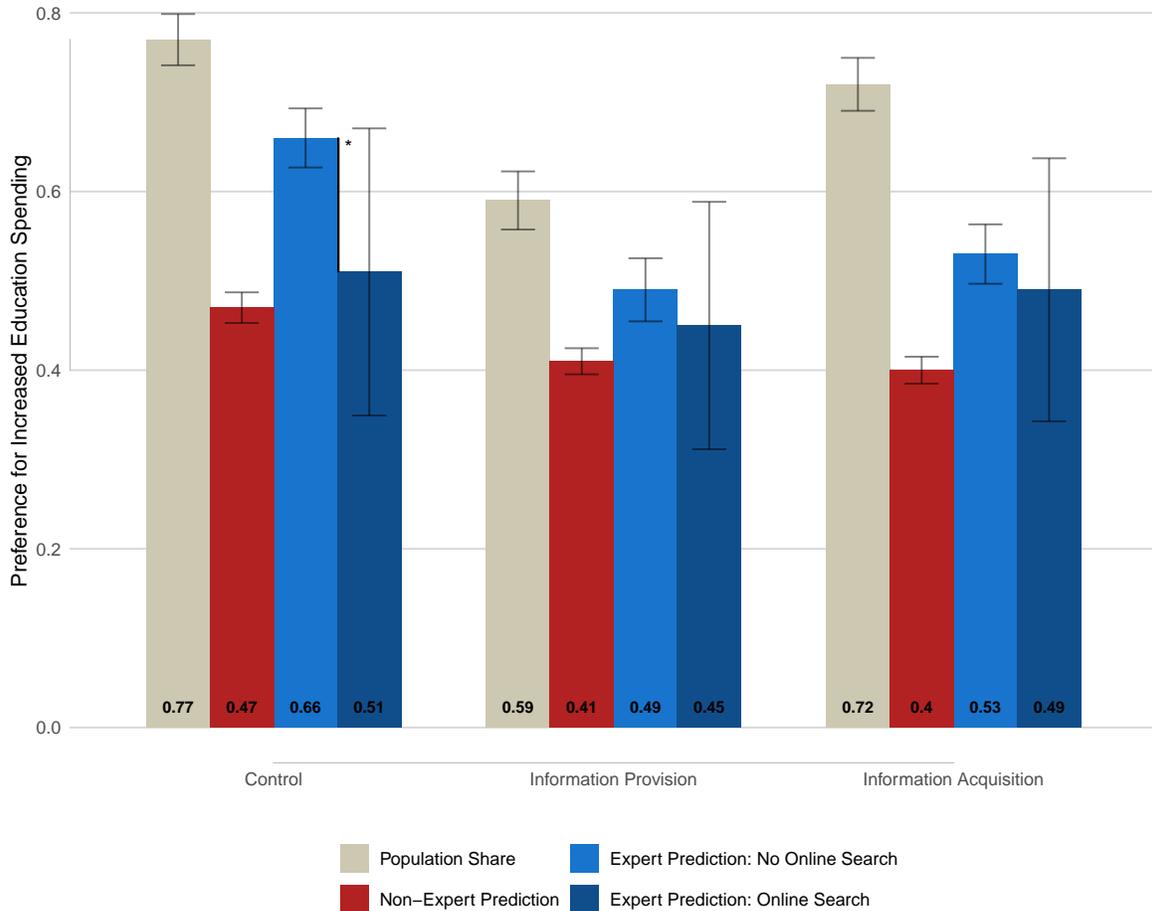
Note: The Figure shows means and confidence intervals for population preferences and non-expert and expert predictions. Non-expert predictions are separately calculated for respondents rather or strongly agreeing to try (91.5%)/not try (8.5%) their best to estimate as accurately as possible. The questions asks: "Now we will look again at how you approached the previous estimation questions. To what extent do the following statements apply to you? I tried my best to estimate as accurately as possible." Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) are asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the information provision condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context. Then, the three different questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The people are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?. Population and non-expert sample weighted by survey weights. Significance levels: *p<0.1; **p<0.05; ***p<0.01. Data sources: ifo Education Survey, ifo Economists Panel 2020

Figure B16 Expert Predictions split by Knowledge of Relevant Studies



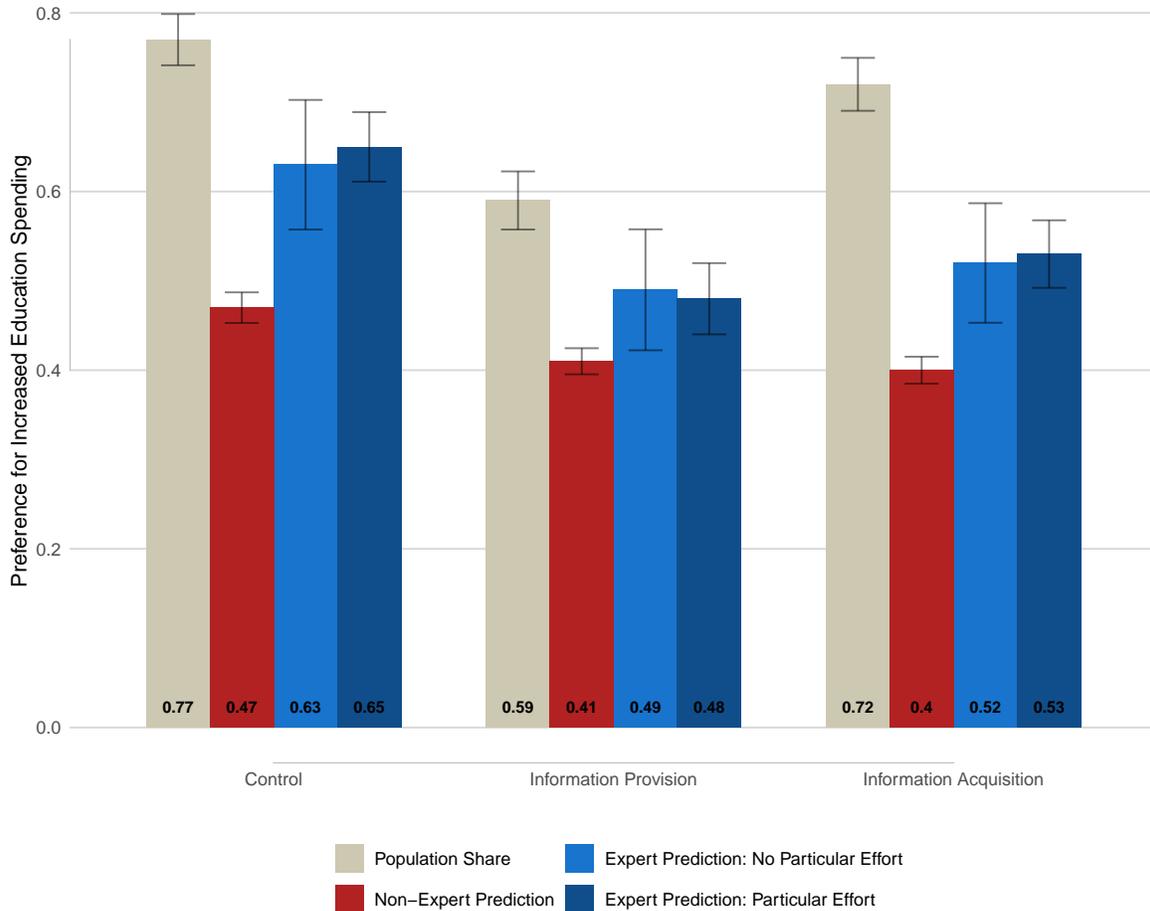
Note: The Figure shows means and confidence intervals for population preferences and non-expert and expert predictions. Expert predictions are separately calculated for experts rather or strongly agreeing to know (N = 17)/not know (N = 105) scientific studies focusing on support of education spending. The question asks: "Now we would like to revisit how you approached the previous estimation questions. To what extent do the following statements apply to you? I am familiar with scientific studies on support for education spending and based my answers on them." Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) are asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the information provision condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context. Then, the three different questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The people are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?. Population and non-expert sample weighted by survey weights. Data sources: ifo Education Survey, ifo Economists Panel 2020

Figure B17 Expert Predictions split by Search for Information



Note: The Figure shows means and confidence intervals for population preferences and non-expert and expert predictions. Expert predictions are separately calculated for experts rather or strongly agreeing to search ($N = 12$)/not search ($N = 111$) for information online. The question asks: "Now we will look again at how you approached the previous estimation questions. To what extent do the following statements apply to you? I searched for information on the Internet in order to make the best possible estimate." Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) are asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the information provision condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context. Then, the three different questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The people are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?. Population and non-expert sample weighted by survey weights. Significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Data sources: ifo Education Survey, ifo Economists Panel 2020

Figure B18 Expert Predictions split by Particular Effort



Note: The Figure shows means and confidence intervals for population preferences and non-expert and expert predictions. Expert predictions are separately calculated for experts rather or strongly agreeing to try (N = 96)/not try (N = 28) their best to estimate as accurately as possible. The questions asks: "Now we will look again at how you approached the previous estimation questions. To what extent do the following statements apply to you? I tried my best to estimate as accurately as possible." Treatment arms: *information provision* and *information acquisition*. Participants in the control group (1,045) are asked: "In your opinion, should public spending on schools in Germany increase, decrease or remain unchanged?" Participants in the information provision condition (1,037) receive the information that average public spending on education in Germany amounts to 7,300 Euros per student per year. Instead of seeing this information right away, participants in the *information acquisition* treatment (1,051) could click on a button to find out about public spending per student. The dependent variable is a dummy variable assuming the value 1 if participants respond spending should "increase" or "strongly increase". Participants in the prediction sample (1,024) are first informed about the context. Then, the three different questions are given to them explicitly marking the difference between conditions. In a third step, participants are asked to predict a number for each condition: "What proportion of respondents do you think will answer "strongly increase" or "increase"?" Calculations for the population are based on survey weights. Data source for population preferences and predictions: ifo Education Survey 2020. For the expert predictions, 127 economists answered in the course of the Economists' Panel: "The aim now is for you to estimate the answers of the ifo Education Barometer 2020 survey participants as well as possible. The ifo Education Barometer is currently surveying a representative sample of people between the ages of 18 and 69. The people are randomly divided into three groups and asked about state spending on schools. Group 1 ("No information") answers the following question: (...). Group 2 ("Information") answered the following question: (...). All three groups can choose between five possible answers when answering the question: "strongly increase", "increase", "remain roughly unchanged", "decrease", "strongly decrease". What proportion of respondents do you think will answer "strongly increase" or "increase"?. Population and non-expert sample weighted by survey weights. Data sources: ifo Education Survey, ifo Economists Panel 2020