Model Uncertainty

Robin Musolff and Florian Zimmermann*

This version: July 1, 2026

Abstract

Mental models help people navigate complex environments. This paper studies how people deal with model uncertainty. In an experiment, participants estimate a company's value, facing uncertainty about which one of two models correctly determines its true value. Using a between-subjects design, we vary the computational complexity of implementing the models. Results show that in high-complexity conditions people fully neglect model uncertainty in their actions. However, their beliefs continue to reflect model uncertainty. Furthermore, we show that complexity, via full uncertainty neglect, leads to higher confidence in the optimality of own actions. We present a simple framework that can generate this set of results and provide evidence on underlying mechanisms that corroborates key model intuitions.[1]

1 Introduction

People rely on mental models to interpret and navigate the complexities of external reality. These models are mental frameworks that shape how individuals process information, form expectations, and make decisions. In contexts of economic decision-making, mental models play a critical role, as individuals attempt to simplify and make sense of complex environments. However, in many situations, economic agents face not only

*Robin Musolff: University of Cologne (robin.musolff@gmail.com); Florian Zimmermann: University of Bonn and Max Planck Institute for Behavioral Economics (florian.zimmermann@uni-bonn.de). We are grateful to Chiara Aina, Kai Barron, John Conlon, Benjamin Enke, Tilman Fries, Nicola Gennaioli, Thomas Graeber, Alex Imas, Botond Koszegi, Alexander Laubel, Chris Roth and Sevgi Yuksel for very helpful comments and suggestions. Funding by the Deutsche Forschungsgemeinschaft (DFG) through CRC TR 224 (Project A01) and under Germany's Excellence Strategy - EXC 2126/1 - 390838866 is gratefully acknowledged. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC Starting Grant: 948424 - MEMEB).

[1]Accompanying materials are available online: Online Appendix (Sections B and C); Supplementary Material D (Additional Results); Supplementary Material E (Model Framework); and Supplementary Material F (Experimental Instructions).

uncertainty about future outcomes, but also model uncertainty—uncertainty regarding which mental model best captures the true dynamics of the environment.

Model uncertainty is prevalent in many areas of economics. For instance, when forming expectations about future inflation or returns on investments, individuals may be uncertain about which underlying model of the economy is most appropriate to use. Similarly, when assessing the effectiveness of government policies, people may lack clarity on which mental model best explains how government actions translate into economic outcomes. In these cases, individuals face a dual problem: not only must they process new information, but they must also decide how to map this information into a decision, given uncertainty about competing models of the world. This problem is especially severe in environments where working with models is complex. Arguably, models of the world differ substantially in how complex they are to implement and work with. Consider, for example, a simple valuation rule of thumb such as comparing a stock's price-to-earnings ratio to a familiar benchmark — applying this model is computationally trivial. By contrast, assessing the same stock using a multi-stage discounted cash flow model requires a sequence of non-trivial calculations. The question hence arises how people function under model uncertainty when the computational complexity of working with those models is high. The uncertainty about which model is correct can then have profound implications for decision-making, beliefs, and ultimately, economic outcomes.

Existing research on mental models has largely focused on understanding the determinants of model selection, how mental models can be used to persuade others, and the consequences of relying on misspecified models. For example, Schwartzstein and Sunderam (2021) and Barron and Fries (2024) examine how individuals might be persuaded to adopt specific mental models under different contexts, while Heidhues, Kőszegi, and Strack (2018) explore how biased models can lead to systematic mistakes in decision-making. Similarly, recent work by Frick, Iijima, and Ishii (2022) investigates the strategic use of mental models in shaping economic beliefs. However, less is known about how individuals cope with model uncertainty in the first place.

In this paper, we aim to fill this gap by studying how individuals operate under model uncertainty. We hypothesize that the computational complexity of mental models determines whether people neglect model uncertainty. Our experimental results indeed reveal that, when complexity is high, people simplify the world by operating as if one model of the world is correct, hence fully neglecting model uncertainty. This is echoed by hovering data which suggest that when complexity is high, people have a tendency to predominantly attend to one specific model of the world. These results are robust to variations in the signal space as well as variations in the specific task people need to perform. Turning to implications of this complexity-induced simplification, we document two key results: (i) neglect of model uncertainty does not translate into a simplified view of the world. When we directly elicit beliefs about which model of the world is correct, these beliefs do reflect model uncertainty despite model neglect in actions, creating a wedge between actions and beliefs; (ii) model uncertainty neglect creates an illusion of certainty. The complexity-induced simplification leads to higher levels of confidence in the optimality of own actions such that, perhaps counter to intuitions, computational complexity increases rather than decreases decision confidence in our setting. In the final part of the paper we present a simple framework that features top-down as well as bottom-up attention and can help explain this pattern of results as well as insights from additional mechanism experiments that corroborate key model intuitions.

We design an experiment that allows us to infer from their actions how people deal with model uncertainty. We implement our experiment in a financial decision-making context where participants' task is to provide valuations for fictitious companies. There is a set of variables that are potentially relevant to determine company values. Participants face uncertainty about which of two models of the world makes correct use of these variables to estimate the value of a company. Participants are endowed with a 50-50 prior and then receive a noisy signal about which model is correct. In a between-subjects design, we manipulate the computational complexity of these models. Our main design goal here was to isolate the role of computational complexity, keeping other layers of complexity as well as all other aspects of the decision environment constant. We achieve this by fixing the underlying models themselves and varying only the difficulty of implementing or working with them. In the low-complexity condition, each model provides a direct estimate of the company's value using the variables as input, requiring no computations from participants to get at each model's estimate. In the high-complexity condition, we add a layer of computational complexity, as participants need to compute the value estimates of the models themselves using the variables as inputs. Participants provide value estimates for eight different companies. Afterwards, we also measure beliefs about which model participants think is correct. This allows us to examine the extent to which complexity influences both actions (i.e., the valuation) and beliefs (i.e., the perception of which model is correct).

Our findings suggest that computational complexity significantly alters how people deal with model uncertainty. Compared to the low-complexity condition, participants in the high-complexity condition have a more pronounced tendency to simplify the world by neglecting model uncertainty in their actions, behaving as if the more likely model is definitely correct. This relates to the concept of coarse categorization, where decision-makers collapse distinct states of the world into a single category and act as if finer distinctions do not exist (Steiner and Stewart (2015)), as well as to the use of coarse mental categories more broadly (Mullainathan, Schwartzstein, and Shleifer (2008)). The tendency to act as if there is no model uncertainty implies an overreaction to the signal about the correct model. In other words, when complexity is high, participants place too much weight on the value estimate provided by the more likely model. Data on hovering times indicate that complexity affects decision-making through an attention channel (Bordalo, Gennaioli, and Shleifer (2012), Bordalo, Conlon, Gennaioli, Kwon, and Shleifer (2023b), Bordalo, Conlon, Gennaioli, Kwon, and Shleifer (2023a), Ba, Bohren, and Imas (2022), Steiner, Stewart, and Matějka (2017)). In the high-complexity condition, participants spend significantly more time focusing on the signal-congruent model, whereas in the low-complexity condition, their attention is more evenly distributed between the two models.

In another experiment, we verify that simplification does not require one model to be more likely than the other. When both models of the world are equally likely, computational complexity nonetheless leads to a neglect of model uncertainty. The robustness of our results is further established across a series of additional experiments. Specifically, the effect persists when the order of value estimation and model uncertainty elicitation is counterbalanced, when signal accuracy is varied, and when the hovering component is removed from the experimental design. Finally, substituting the value estimation task with an investment task leaves the central result unchanged — complexity consistently and significantly undermines the incorporation of model uncertainty — confirming that the phenomenon generalizes beyond the specific task parameters of the primary study.

Next, we turn to implications of this complexity-induced simplification. We first document that neglect of model uncertainty in actions does not translate into beliefs. Participants' stated beliefs continue to reflect model uncertainty. This disconnect between beliefs and actions creates a systematic wedge between what individuals believe and how they behave under model uncertainty. This echoes earlier work by Giglio, Maggiori, Stroebel, and Utkus (2021), Ameriks, Kézdi, Lee, and Shapiro (2020), Beutel and M. Weber (2023), Laudenbach, A. Weber, R. Weber, and Wohlfart (forthcoming) who have identified a gap between subjective beliefs and economic behavior in different contexts.[2] Notably, this gap arises directly after participants stated their valuations, and persists even after a 1 day delay, as we verify in a separate experiment that spans over 2 days.

Second, using additional experiments where we measure participants' confidence in the optimality of their actions, we document that confidence in the optimality of own actions is higher in the high-complexity conditions compared to the low-complexity conditions. This holds for an unincentivized confidence measure, where participants state the probability that their guesses were optimal, as well as an incentivized measure, where participants place a bet on the optimality of their actions. Prior literature in contexts different from ours has shown that complexity tends to increase cognitive uncertainty (Enke and Graeber (2023), Enke, Graeber, and Oprea (forthcoming), Enke, Graeber, and Oprea (2023)). In contrast, it seems that in the context of model uncertainty, the possibility to respond to complexity by simplifying the world through full neglect of model uncertainty leads to an illusion of certainty and hence increased confidence in action optimality.

In the final part of the paper we present a simple framework that formalizes this intuition, generates the key results from our experiments and yields further testable hypotheses. In the model, an agent forms a mental representation of a decision problem through the interplay of bottom-up and top-down attentional processes. Upon encountering a decision problem, a bottom-up, cue-dependent memory retrieval process determines which stored mental representation becomes salient (Bordalo, Gennaioli, Lanzani, and Shleifer (2025)). In a subsequent top-down process (Kőszegi and Matějka (2020), Steiner, Stewart, and Matějka (2017), and Gabaix (2014)), the agent delibera-

[2] Yang (2023) provides an explanation for the attenuated relation between beliefs and actions based on a specific form of cognitive uncertainty.

tively decides whether to further simplify this representation. All actions, belief elicitations, and confidence assessments are then executed within the mental representation so formed. We provide direct empirical support for basic model mechanisms using open-text elicitations. Furthermore, the model yields a sharp testable prediction: specific cues should activate distinct mental representations, with corresponding and predictable behavioral consequences. We show experimentally that a cue designed to trigger the broad, non-simplified representation of the problem significantly reduces confidence in action optimality under high complexity, in line with our model's predictions.

Our work directly relates to a growing literature on (misspecified) mental models (Schwartzstein and Sunderam (2021), Montiel Olea, Ortoleva, Pai, and Prat (2022), Gagnon-Bartsch, Rabin, and Schwartzstein (2023), Mailath and Samuelson (2019), Heidhues, Kőszegi, and Strack (2018), Heidhues, Kőszegi, and Strack (2023), Frick, Iijima, and Ishii (2022), Aina (2024), Barron and Fries (2024), Ambuehl and Thysen (2024), Charles and Kendall (2025), Esponda, Vespa, and Yuksel (2024), Fréchette, Vespa, and Yuksel (2024)).[3] Until now, this literature has largely focused on an analysis of the determinants of model selection, how mental models can be used to persuade others, and the consequences of relying on misspecified models. In contrast, our focus is on how people deal with model uncertainty. A common assumption in the literature is that people work with a single (possibly misspecified) model when making decisions, rather than entertaining multiple weighted models simultaneously (e.g. Schwartzstein (2014), Schwartzstein and Sunderam (2021), Montiel Olea, Ortoleva, Pai, and Prat (2022)). We empirically document this kind of simplification, show that it increases with decision complexity and study its implications. Aina and Schneider (2025) study how people update their beliefs in the presence of competing models that could explain observed data. They provide evidence that the majority of people select the model that explains the observed data best. While our focus is on the role of computational complexity and the implications of model uncertainty neglect, consistent with our finding of full neglect of model uncertainty, they also find that many participants in their experiments build their belief formation process on only one model of the world.

Our work also ties to a literature that studies how limited attention shapes how peo-

[3] Relatedly a growing theoretical and empirical literature studies the role of stories and narratives in economics (e.g., Shiller (2017), Eliaz and Spiegler (2022), Andre, Haaland, Roth, and Wohlfahrt (2024), Graeber, Roth, and Zimmermann (2024), Graeber, Roth, and Schesch (2024)).

ple react to information.[4] Bordalo, Conlon, Gennaioli, Kwon, and Shleifer (2023b) and Bordalo, Conlon, Gennaioli, Kwon, and Shleifer (2023a) provide formal frameworks as well as experimental evidence of how attention (and memory) patterns shape belief formation and information processing. Ba, Bohren, and Imas (2022) show that attention processes can lead to overreaction to information. Esponda, Oprea, and Yuksel (2023) show that a form of representativeness heuristic has important implications in contexts of statistical discrimination. Enke and Zimmermann (2019), Enke (2020), and Graeber (2023) provide evidence that people systematically fail to attend to key aspects of the information environment when processing new information.[5] We document a related phenomenon in the context of model uncertainty where people selectively attend to only one model of the world.

Work on rational inattention (Sims (2003), Kőszegi and Matějka (2020), Caplin and Dean (2015), Steiner, Stewart, and Matějka (2017), and Gabaix (2014)) formalizes top-down attention processes where agents decide to limit attention to reduce the complexity of situations. Relatedly, a literature on coarse categorization studies how agents simplify the state space itself, grouping distinct states together and treating them as equivalent (Steiner and Stewart (2015) and Mullainathan, Schwartzstein, and Shleifer (2008)). Steiner and Stewart (2015), for instance, show that frequent trading opportunities amplify the price distortions that arise when traders categorize states coarsely. Our finding that high complexity leads people to fully collapse two models of the world into one can be viewed as a behavioral instance of this kind of categorization, here driven by computational complexity rather than by the trading frequency or persuasion incentives studied in this literature. In general, we show that such top-down simplification can lead to distortions of reality in the sense that they influence which types of mental presentations are top of mind in subsequent decisions, for instance when assessing decision confidence.

Our work also relates to a literature that studies the effect of complexity on atten-

[4] More broadly, Bordalo, Gennaioli, and Shleifer (2012), Bushong, Rabin, and Schwartzstein (2021), Kőszegi and Szeidl (2013) formalize how contextual features steer attention and focus and hence influence behavior. (See Dertwinkel-Kalt, Gerhardt, Riener, Schwerter, and Strang (2022) as well as Somerville (2022) for experimental tests.)

[5] Hartzmark, Hirshman, and Imas (2021) show that ownership leads to overreaction to information, an effect that is driven by channeled attention on information. Augenblick, Lazarus, and Thaler (forthcoming) and Fan, Liang, and Peng (2024) instead focus on the role of cognitive noise and similarity patterns, respectively, in explaining over- and underreaction to information.

tion. Recent research on complexity has made substantial progress in defining and quantifying complexity and studying implications in different decision contexts (e.g., Oprea (2020), Kendall and Oprea (2024), Enke and Shubatt (2024), Shubatt and Yang (2024), Arrieta and Nielsen (2024)). A common finding is that limited attention on a subset of relevant decision parameters is a simplification response to complexity (Enke (2024), Ba, Bohren, and Imas (2022), Enke (2020), Enke and Zimmermann (2019), Graeber (2023)). We show that in the presence of model uncertainty, computational complexity induces people to fully neglect model uncertainty. Furthermore, we document that due to the simplification response of uncertainty neglect, confidence in the optimality of own actions is higher in the high-complexity conditions compared to the low complexity conditions.

The rest of the paper is structured as follows. In Section 2, we describe the baseline experimental design. Section 3 presents the results on how model complexity leads to the simplification of model uncertainty in actions. Section 4 then delves into implications of this simplification, studying beliefs and cognitive uncertainty. In Section 5 we present a short model that can generate the observed pattern of results, and Section 6 provides empirical tests of its key predictions. We conclude in Section 7.

2 Baseline Experimental Design

We designed our experiment with the following goals in mind: (i) implement an economically meaningful and somewhat natural decision environment that allows us to study how people deal with model uncertainty in the face of computational complexity; (ii) achieve a well-defined notion of model uncertainty that allows us to exogenously vary computational complexity in a straightforward way and (iii) being able to infer the weighting of competing models through the measurement of actions.

The Task and Model Uncertainty. We chose a financial decision-making task. In the experiment, respondents had to estimate the value of 8 fictitious companies. The correct company value was determined by one of two models, "The CEO is key" or "Products are crucial". Following existing work on mental models, we conceptualize a model as a functional mapping from observable variables to outcomes or predictions.[6] Specifically, the two models used different variables as inputs and consist of functional forms that translate inputs into value estimates. One of the models was correct, meaning that it provided the correct company values for all 8 companies, while the other model produced uninformative values.

Each of the two models consists of a formula to calculate the proposed values. "The CEO is key" had the variables CEO competence C and Supporting Staff S as inputs, which could be used to calculate the company value as $C \times S - C - S + 10$. "Products are crucial" had the inputs number of products P and research cost R, and the company value was given by $P \times (10 - R) + R - P$.

To implement model uncertainty, the correct rule was determined in secret by the computer with a simulated coin flip. Therefore, without any additional information, the probability that either rule produced the correct company values was 50%. Respondents then received a noisy but informative signal about the correct model. This signal corresponded to the truth with a probability of 65%.[7] This was visualized using a ball drawn from an urn containing 65 balls with the correct model and 35 balls with the incorrect model. Afterwards, to measure actions, respondents were asked to provide value estimates for the 8 companies, each featuring different sets of variable realizations.[8]

Company value estimates were incentivized through a binarized scoring rule.[9] In this way, the chance of respondents to win a bonus was maximized by stating their best guess. Danz, Vesterlund, and Wilson (2022) document empirically that the binarized scoring rule can lead to systematic bias. Notice that such bias (if present in our setting) would not compromise our identification which relies on the comparison of value estimates between conditions of high and low complexity. Furthermore, our results are robust to using investment behavior rather than value estimates as an outcome, which features a

[6] Schwartzstein and Sunderam (2021), for instance, define a mental model as a likelihood function that maps a history of outcomes to posterior beliefs that then potentially lead to actions. Similarly, the literature on misspecified mental models typically formalizes a mental model as a mapping from input variables to predictions or beliefs (see Heidhues, Kőszegi, and Strack (2018), Heidhues, Kőszegi, and Strack (2023)).

[7] We did not provide any feedback between rounds. Hence, the only information respondents receive about which model is correct is the signal described above.

[8] The variable realizations were integers between 0 and 10 and all implemented configurations were selected to yield company values between 10 and 100.

[9] Every tenth participant was eligible to receive a bonus payment, in which case a random decision from the survey was incentivized. If a company value guess was incentivized, respondents received \$10 with a probability (in percent) of $100 - 100 \times (\text{Truth}/100 - \text{Guess}/100)^2$, where Truth is the true company value and Guess the company value guess.

different incentive scheme (see Section 3.3).

Complexity Manipulation. We argue that computational complexity of mental models is an important source of variation in reality. Some models can be very simple to implement, while others are computationally demanding. We varied the computational complexity of the two models between-subjects.[10] A key design challenge here is to manipulate computational complexity, leaving other aspects of the model unchanged. We solved this by fixing the actual underlying models, only manipulating the computational difficulty of implementing or working with the models. In treatment condition *HighComplexity*, working with the models was complex. Specifically, respondents had to calculate the company values under both models themselves. To obtain the Bayesian company value guess, they then needed to weight both values by the respective probabilities of 65% and 35%. In treatment condition *LowComplexity*, they were provided with the calculated company values for both models. Hence, respondents only needed to combine and weight the two values to make their guess, which significantly reduced the complexity of working with the models.[11]

Cognitive Uncertainty and Beliefs about Model Uncertainty. We measured cognitive uncertainty, i.e., people's confidence in the optimality of their value estimates similar to Enke and Graeber (2023). Specifically, after the series of 8 value estimation tasks, we asked people: *'How certain are you that, on average, your guesses were no more than 10 points away from the best possible guess given the information you received?'* Respondents indicated their answer on a scale from 0 to 100 percent. The cognitive uncertainty measure was only implemented in a subset of experiments (see Table A.1 and Section 4.2) and was not incentivized. We conducted an additional experiment that features an incentivized confidence measure where participants place a bet on the optimality of their actions (see Section 4.2).

In the last part of the experiment, respondents had to guess the probabilities that either of the two models generated the correct company values. Since respondents received a noisy signal during the company valuation task, this corresponds to the stan-

[10] In essence, our manipulation raises the cognitive cost of evaluating each model, and the simplification response that follows is the mechanism we study. We think of notions such as calculation fatigue or math aversion as part of what it means for a task to be computationally complex.

[11] Indeed, average guessing times in the Baseline experiment were significantly shorter at 17.09 seconds in *LowComplexity*, compared to 49.11 seconds in *HighComplexity* ($p<0.001$).

dard bookbag and chips belief updating task with a signal precision of 65% (cf. Benjamin (2019)). The stated belief was incentivized using a binarized scoring rule.[12] Participants were also asked an un-incentivized direct recall question where they were asked to state which rule was indicated to be more likely by the signal.

Design Details and Procedures. Respondents in both treatment conditions initially received the same set of instructions, explaining both treatments. Afterwards, they went through compulsory comprehension checks and two test runs, where they could practice applying each of the model formulas.[13] After they completed the test runs, they were randomly assigned their treatment, received the noisy signal, and subsequently completed the 8 rounds of the task.

On the decision screens, respondents had to hover their mouse over the name of the respective model to uncover the variables and formula or calculated company value (cf. Ba, Bohren, and Imas (2022)). This allows us to study the attention paid to both of the models and signal realizations.

The experiments in this paper were preregistered on the platform AsPredicted. The preregistrations include the experimental design, hypotheses, analyses, sample sizes, and exclusion criteria. Table A.1 provides links to each preregistration.

We conducted the experiments online using the survey provider Prolific. Respondents were recruited from the United States and restricted to be fluent in English. To qualify for the survey, participants had to pass comprehension checks after reading the instructions. The experimental instructions can be found in Supplementary Material F[14]. The median completion time in the Baseline Experiment was 22 minutes. Respondents received a fixed payment of $4 for the initial study. Every tenth respondent had the chance of winning an additional bonus of up to $10.

As preregistered, we focus our analysis on two different samples. After reading the instructions, but before treatment assignment, we presented respondents with two test runs for how to calculate the estimates of the models under high complexity (as de-

[12]If a probability guess was incentivized, respondents received $10 with a probability of $100 - 100 \times (\text{Truth}/100 - \text{Guess}/100)^2$ %, where Truth is either 100 or 0, and Guess is the stated probability between 0 and 100.

[13]We kept instructions and comprehension checks identical across conditions to avoid confounding computational complexity with instruction complexity. Uniform instructions and comprehension checks also ensure that there cannot be differential selection out of the experiment due to different failing rates for the comprehension checks.

[14]Available online: Supplementary Material F (Experimental Instructions).

scribed above). There, they could familiarize themselves with how to calculate the estimates under both decision models. In order to ensure that we have a respondent pool that is in principle able to solve the formulas in the high-complexity condition, we preregistered to restrict our sample to respondents who correctly answered both of the two example decisions. The restricted sample of the Baseline Experiment includes 230 respondents. All figures and tables in the main text and Online Appendix B are based on the restricted sample. Supplementary Material D reproduces all exhibits using the also preregistered more lenient sample that only requires one of the two questions to be answered correctly, featuring 319 respondents for the Baseline Experiment.

3 Complexity and Simplification of Model Uncertainty

3.1 Framework and Hypothesis

Take a decision-maker whose task is to state a value estimate g_{actual}. Recall that in each decision scenario, the correct company value corresponds to one of the two values given by the models "The CEO is key" and "Products are crucial". The noisy signal received by the participant then indicates the model that is more likely to deliver the correct value. We call the signal-consistent value $v_{consistent}$, and the signal-inconsistent value proposed by the other model $v_{inconsistent}$. Since the signal reveals the correct rule with a probability of 65%, the rational guess for the company value is given by

$$g_{rational} = 0.65 \times v_{consistent} + 0.35 \times v_{inconsistent}.$$

Now take a decision-maker who seeks to simplify the world by neglecting model uncertainty. Such a decision-maker will base their value estimates exclusively on the more likely model. The naive benchmark that fully neglects model uncertainty and takes the signal at face value is hence given by

$$g_{naive} = v_{consistent}.$$

As preregistered, the main statistical measure we employ is the naive weight λ implicitly defined by

$$g_{actual} = \lambda \times g_{naive} + (1-\lambda) \times g_{rational}, \tag{1}$$

which can be rearranged to obtain

$$\lambda = \frac{g_{actual} - g_{rational}}{g_{naive} - g_{rational}}. \tag{2}$$

The naive weight λ equals 1 if a participant states the naive guess (full neglect of model uncertainty) and 0 if they state the rational guess.

We expect that decision-makers will be more prone to simplify the world if complexity is high. Therefore we state the following, preregistered, hypothesis:

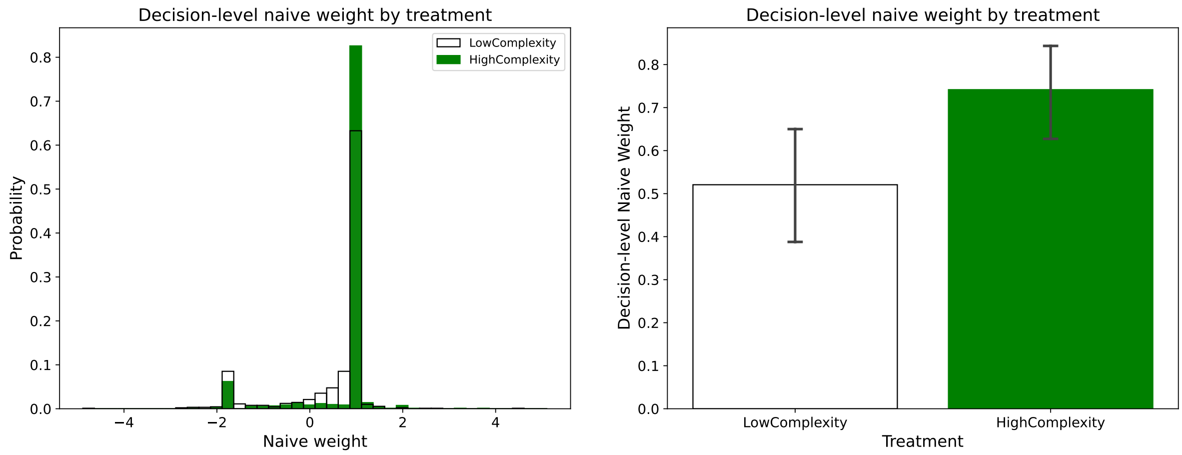
Hypothesis. *Decision-makers are more likely to neglect model uncertainty when complexity is high, compared to when complexity is low.*

3.2 Results

Figure 1a plots histograms of decision-level naive weights in our main sample for both treatments. The distribution in *HighComplexity* features more mass at 1 than the one in *LowComplexity*, indicating a larger amount of fully naive guesses and overreaction to information in the former condition. Conversely, the distribution for *LowComplexity* has more mass around 0 than for *HighComplexity*, implying more guesses in close proximity to the rational benchmark. The prevalence of fully naive guesses is 80% in *HighComplexity*, compared to 59% in *LowComplexity* ($p < 0.01$). Hence, the vast majority of guesses is fully naive in *HighComplexity*.[15] Figure 1b confirms this pattern by comparing the average naive weight across treatments, finding a significantly larger average naive weight under complexity ($p < 0.01$).

Table 1 complements this analysis by regressing the company value guesses on the rational and naive benchmarks. This can be interpreted as estimating Equation 1 without the restriction that the weights on the rational and naive benchmarks add up to 1. We can see that the fully naive benchmark is relatively more predictive of participants' guesses in treatment *HighComplexity* compared to *LowComplexity*.

[15] A large fraction of guesses in *LowComplexity* also reveal full naivete, which may be caused by the residual complexity of the general experimental set-up or more specifically of the need to combine the signal-consistent and signal-inconsistent values into a value estimate.



(a) Distribution of decision-level naive weights

(b) Mean decision-level naive weights

Figure 1: Decision-level naive weights in LowComplexity and HighComplexity conditions. *Panel (a) plots the distribution of naive weights λ calculated as specified in Equation 2, using the restricted sample of the Baseline Experiment with 230 participants. Panel (b) plots average naive weights.*

Result 1. *There is substantially more neglect of model uncertainty when complexity is high, compared to when it is low.*

We note that the naive weight defined in Equation (2) is quite sensitive to outliers. To ensure that our results are not driven by outliers, we look at the 8 decision scenarios for each subject and compute the median naive weight. In an analysis that was not preregistered, we plot the histogram of median naive weights for each subject in Online Appendix Figure B.1. We again observe a higher mass around naive weights of 1 in *HighComplexity*, and more mass between 0 and 1 in *LowComplexity*, confirming the earlier results that full neglect of model uncertainty is more prevalent under complexity.

Supplementary Material D.1 replicates all exhibits using the also preregistered more lenient sample that only requires one of the two questions to be answered correctly. In Online Appendix B.6 we show that we replicate all the findings for the Baseline Confidence Experiment (see Section 4.2 and Table A.1).

Learning and Consistency Across Rounds. In an analysis that was not preregistered, Figure A.1 plots the evolution of the decision-level naive weight across the eight decision rounds, pooling observations from the Baseline Experiment and two robustness studies featuring an identical action stage. The treatment difference is present in every individual round and there is no systematic trend in either direction, suggesting that the

Table 1: Company Value Guesses

Dependent variable:	Company Value Guess		
Sample:	LowComplexity (1)	HighComplexity (2)	Pooled (3)
Rational Benchmark	0.593*** (0.079)	0.343*** (0.065)	0.593*** (0.079)
Naive Benchmark	0.481*** (0.070)	0.703*** (0.057)	0.481*** (0.070)
Rational B. × HighComplexity			-0.250** (0.102)
Naive B. × HighComplexity			0.222** (0.090)
R^2	0.881	0.918	0.900
Observations	904	936	1840

The table presents OLS regressions of respondents' company value guesses on the rational and naive benchmarks as detailed in Section 3.1. The table is based on the restricted sample of the Baseline Experiment. Column (1) uses observations from the LowComplexity treatment, column (2) from the HighComplexity treatment, and column (3) from both. Stars highlight significant differences from 0 with * for $p < 0.10$, ** for $p < 0.05$, *** for $p < 0.01$. Standard errors are clustered on the subject level.

treatment effect does not diminish with experience.

Role of Attention. The above results indicate that people simplify complexity by neglecting model uncertainty. To further corroborate this finding, we study the role of attention. Recall that respondents had to hover their mouse over either "The CEO is key" or "Products are crucial" to observe the parameters needed to make company value guesses. In *LowComplexity*, the company value was displayed only when hovering over the respective rule. In *HighComplexity*, the respective variable realizations needed to calculate the company value were displayed. The resulting data on hover times allows us to study how much attention participants paid to either rule.

Figure 2 shows the distribution of the median consistent hovering share per subject — defined as the median across the eight tasks of the ratio of hover time spent on the signal-consistent rule to total hover time on both rules — separately for both treatments. In treatment *HighComplexity*, an overwhelming majority of participants only considers the variables of the signal-consistent model. In treatment *LowComplexity*, most participants

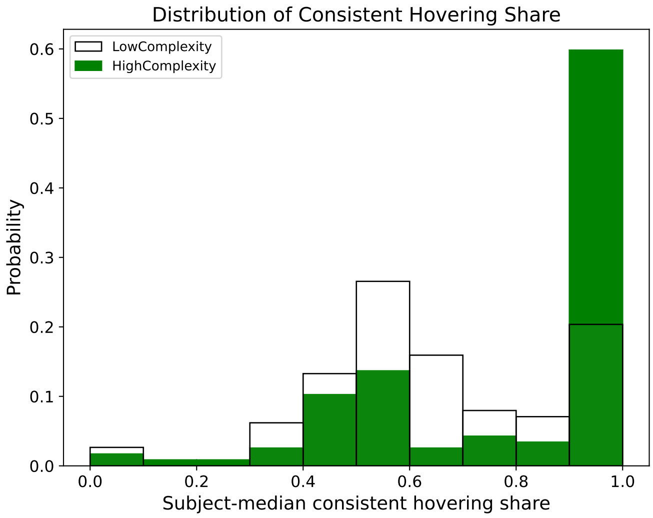


Figure 2: Distribution of subject-medians of the consistent hovering shares. *The figure plots the distribution of the median consistent hovering share per participant — where, for each participant, we take the median across their eight tasks of the ratio of hover time spent on the signal-consistent rule to total hover time on both rules. The figure uses the restricted sample of the Baseline Experiment with 230 participants.*

consider the proposed values by both models, with the mode being at equal hovering times for both the signal-consistent and inconsistent model.

Table 2 analyzes more formally how attention differs between the two treatments. The first column shows that the average hover time for the signal-consistent rule more than triples with higher complexity. The second column shows that the hover time for the signal-inconsistent rule also increases, but by much less. The third column confirms that the share of consistent hover time increases in *HighComplexity*, meaning that participants in this condition pay relatively more attention to the signal-consistent rule.

In an analysis that was not preregistered, Table A.2 examines whether the treatment difference in naive weight persists when comparing decisions with similar hovering shares. We bin decisions by their consistent hovering share and report mean naive weights separately for each treatment. The treatment difference is small and statistically insignificant in four out of five bins — the only bin where *HighComplexity* exhibits significantly higher naive weights is the 0.8–1.0 bin, which concentrates the vast majority of *HighComplexity* decisions (2124 out of 3136). This pattern suggests that selective attention is one of the key mechanisms through which complexity-induced simplification manifests: computational complexity shifts participants toward more selective attention, which in turn drives the higher naive weight.

Table 2: Hover Times

Dependent variable:	Consistent Hover Time	Inconsistent Hover Time	Consistent Share
Sample:	Pooled (1)	Pooled (2)	Pooled (3)
Constant	2.641***	1.533***	0.635***
	(0.232)	(0.153)	(0.018)
HighComplexity	11.611***	1.588***	0.153***
	(0.954)	(0.392)	(0.027)
R^2	0.262	0.033	0.079
Observations	1840	1840	1840

The table presents OLS regressions using the restricted sample of the Baseline Experiment. Hover times were winsorized at the top at the 97.5% quantile. All columns use observations from both the High-Complexity and LowComplexity conditions. In column (1), the time that respondents spent looking at the values for the signal-consistent rule is regressed on a constant and a treatment dummy for the High-Complexity condition. In column (2), the dependent variable is the time spent looking at the signal-inconsistent rule. In column (3), it is the share of time that was spent looking at the signal-consistent rule. Stars highlight significant differences from 0 with * for $p < 0.10$, ** for $p < 0.05$, *** for $p < 0.01$. Standard errors are clustered on the subject level.

3.3 Robustness

Investment Behavior. In the baseline experiment, our main outcome measure is respondents' value estimates for the hypothetical companies. We ran two additional preregistered versions of the Baseline Experiment that replace the company value estimate with an investment decision. The design of the experiments was exactly as described in Section 2, with only one change. Instead of providing a guess for the value of the company, participants were asked to submit an investment bid for each company. For each decision, they received a budget of 100 cents and could subsequently decide how much to bid for the company. The bids were incentivized through a random auction mechanism: a random price between 10 and 100 cents was drawn for the company. If the bid was greater than or equal to the price, the company was bought by the participant, paying the drawn price, and receiving the true company value. If the bid was lower than the price, there was no transaction.

Both experiments were preregistered (see Table A.1). To ensure comparability with the other experiments, for both experiments we focus on the results from the restricted sample. For experiment 1 this yields 193 participants, for experiment 2 this sample features 323 participants.[16]

Taken together, the results go in the same direction as for the Baseline Experiment. The overall prevalence of fully naive bids in experiment 1 is 73% in *HighComplexity* and 47% in *LowComplexity* ($p<0.001$). Online Appendix Figure B.5b shows that the mean naive weight is significantly higher under complexity. Hover times show the same patterns as in the Baseline Experiment, with there being significantly more consistent hovering in *HighComplexity*. In experiment 2, the share of fully naive guesses is 70% in *HighComplexity* and 41% in *LowComplexity*, and the mean naive weight is also significantly higher under complexity as can be seen in Online Appendix Figure B.9b, confirming the results from the first experiment. Again, hover times also show the same pattern as before (see Online Appendix B.4).[17]

Result 2. *Our results are robust to using a different outcome, namely investment behavior. People's investment bids reflect substantially more full neglect of model uncertainty when complexity is high, compared to when it is low.*

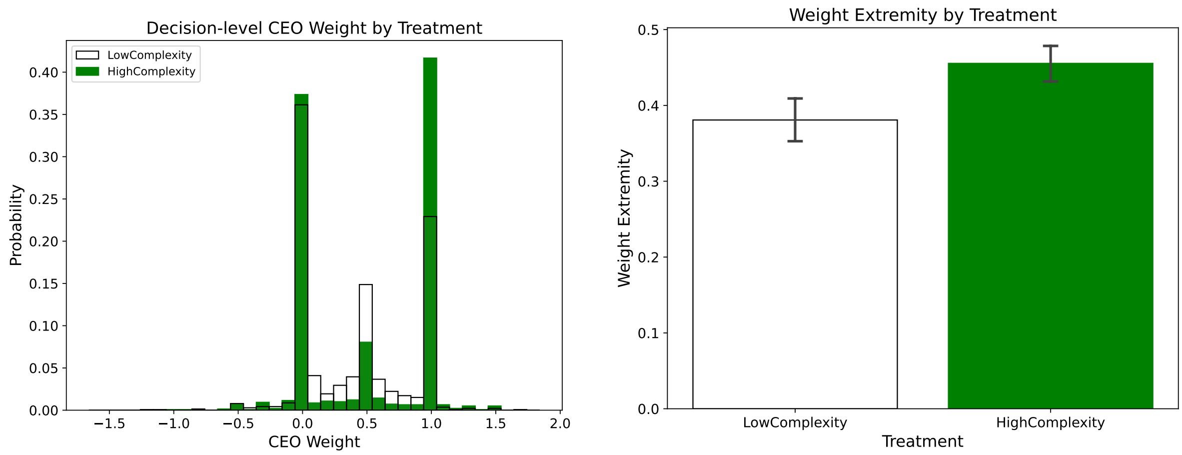
Equally Likely Models. In the baseline experiment, we endow respondents with a natural candidate for simplification, namely the objectively more likely model. Hence, the question arises whether simplification also occurs if both models are equally likely. To address this, we conducted an additional preregistered study, where participants completed a version of the Baseline Experiment that excluded the informative signal. Instead, they were only endowed with a 50-50 prior when estimating company values. The link to the preregistration can be found in Table A.1.

The experiment followed the design described in Section 2 but omitted both the signal and the belief elicitation regarding the rule determining company values. To maintain the experiment's length and incentive structure, we replaced the belief elicitation with an unrelated belief-updating task.

In the absence of a noisy signal, there is no clear reason to expect participants to simplify model uncertainty in a specific direction when faced with complexity. There-

[16]Notice that for experiment 1, we deviate from the specification in the preregistration where we preregistered the use of the lenient sample and the full sample. Supplementary Material D.3 and Supplementary Material D.4 produce the corresponding results. Also notice that for experiment 1 we preregistered a smaller sample size than for the other experiments.

[17]Supplementary Material D.3 and D.5 produce results for the more lenient sample. Overall, results are similar to the restricted sample, although the treatment difference in mean naive weights in experiment 1, while directionally present, fails to be significant in the preregistered lenient sample ($p=0.207$).



(a) Distribution of decision-level CEO weights

(b) Mean decision-level weight extremity

Figure 3: Decision-level CEO weights and weight extremity in LowComplexity and High-Complexity conditions of the Equally Likely Models Experiment. *Panel (a) plots the distribution of CEO weights γ calculated as specified in Equation 3, using the restricted sample of the Equally Likely Models study with 348 participants. Panel (b) plots the average weight extremity $|\gamma - \frac{1}{2}|$ calculated as specified in Equation 4.*

fore, our hypothesis for this study was that participants in the *HighComplexity* condition would more frequently state guesses equal to the values proposed by either rule, whereas more intermediate guesses would be observed in the *LowComplexity* condition.

To test this, we examine the implicit decision weight, γ, assigned to the company value proposed by the "The CEO is key" rule, defined as

$$g_{actual} = \gamma \times v_{CEO} + (1-\gamma) \times v_{Products}. \tag{3}$$

Additionally, we define weight extremity as

$$\text{Weight extremity } = |\gamma - \frac{1}{2}|. \tag{4}$$

Our preregistered hypothesis then is that weight extremity is higher in the *HighComplexity* condition compared to the *LowComplexity* condition.

The preregistered restricted sample consists of 348 participants who correctly solved both HighComplexity example decision screens.

Figure 3a displays the distribution of decision-level CEO weights for both treatments. In the *LowComplexity* condition, the distribution has greater mass at intermediate values, particularly near the rational CEO weight of 0.5. In contrast, the *HighComplexity* distribution shows more mass at weights corresponding to simplified guesses, especially at a CEO weight of 1.[18] Figure 3b further confirms that the average weight extremity in *HighComplexity* is significantly higher than in *LowComplexity*. A more direct measure yields the same conclusion: participants stated a guess equal to the value of one of the two rules in 76.7% of decisions under *HighComplexity*, compared to 56.3% under *LowComplexity* — a difference of 20.5 percentage points ($p < 0.001$) that is of a similar magnitude to the Baseline Experiment.

Online Appendix Figure B.2 displays the distribution of the share of time participants spent hovering over the "The CEO is key" model in both treatments. Participants in *HighComplexity* are more likely to focus on a single model, whereas those in *LowComplexity* typically attend to both models.

In Supplementary Material D.2, we replicate the analysis using the more lenient sample of 452 participants who answered at least one of the two HighComplexity example screens correctly. All results continue to hold in this sample.[19]

Result 3. *People also neglect model uncertainty as a response to complexity when both models are equally likely.*

Baseline Confidence 75%, No Hover, and Beliefs First. We preregistered three additional experiments, each replicating the Baseline design with a single modification. In the *Baseline Confidence 75% Experiment,* the signal precision is raised from 65% to 75%, providing a stronger signal about the true model. In the *No Hover Experiment*, all company value information is permanently visible without requiring participants to hover, removing the attention channel. In the *Beliefs First Experiment*, participants state

[18]The asymmetry between 0 and 1 in the *HighComplexity* distribution may partially be explained by the order in which the two rules were displayed. "The CEO is key" was displayed first for 55% of *HighComplexity* participants versus 47% in *LowComplexity*, and regressing CEO weight on an indicator for whether "The CEO is key" was displayed first yields a coefficient of 0.11 ($p = 0.02$) in *HighComplexity* versus -0.01 ($p = 0.83$) in *LowComplexity*. Interestingly, the significant order effect in *HighComplexity* is itself consistent with simplification: participants in *LowComplexity* consider both rules and are unaffected by presentation order, whereas those in *HighComplexity* appear to anchor on whichever rule is shown first.

[19]A perhaps interesting follow-up question is whether people that simplify model uncertainty as a response to complexity do so in a consistent fashion, always focusing on the same model. To shed some light on this, in an analysis that was not preregistered, Online Appendix Figure B.3 restricts the sample to participants in the *HighComplexity* condition who made only fully naive guesses—that is, who consistently guessed a value corresponding to either the "The CEO is key" or "Products are crucial" rule for every decision. This applies to 91 out of all 174 participants in *HighComplexity*. The figure suggests a large fraction of respondents (about half) consistently simplify in the same direction, always choosing values from the same model, while a sizeable fraction alternates between models.

their probability beliefs about which model determines company values *before* submitting their value guesses, rather than after. The preregistered restricted samples consist of 352, 429, and 364 participants, respectively.

In all three experiments, the main result replicates: the mean naive weight is significantly higher in *HighComplexity* than in *LowComplexity*, as can be seen in Figures B.23b, B.28b, and B.32b in Online Appendix B.8, B.9, and B.10. Moreover, the results on hovering and the role of attention replicate in the *Baseline Confidence 75%* and *Beliefs First Experiments*. Supplementary Material D.8, D.9, and D.10 replicate the results in the more lenient sample that only requires one of the two example questions to be answered correctly.

4 Results: Implications of Simplification

4.1 Beliefs about Models

We have shown that model complexity leads to a simplification response: people tend to fully neglect model uncertainty in their actions. As preregistered, we now ask whether this simplification also causes a distorted view of reality, namely that people misperceive model uncertainty when directly asked.[20]

For this purpose, beliefs about which model is correct were elicited immediately after all company value guesses had been submitted.

Figure 4a shows histograms of probability guesses (signed in the direction of the received signal), separately for treatments *LowComplexity* and *HighComplexity* in the Baseline experiment. There are no visible differences in beliefs. This is confirmed by Figure 4b which plots average guesses for the probability that "The CEO is key" is the correct model, stratified by the signal participants received and their assigned treatment. There is no significant treatment difference in average beliefs for either signal.

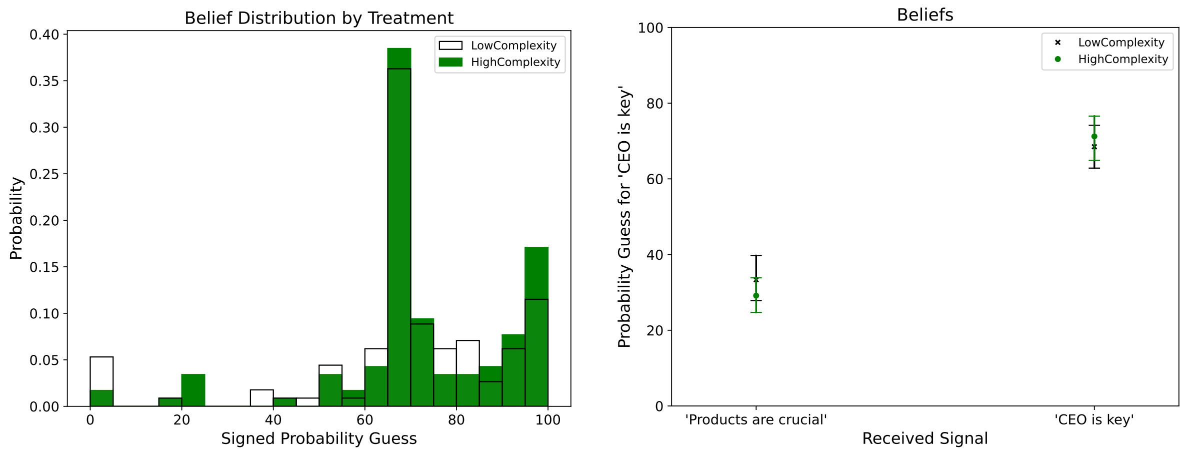
Table 3 presents regressions for signed probability guesses and correct recall of the signal. The regressions confirm that there are no significant treatment differences in beliefs. Similarly, there are no differences in the accuracy of recall of the received signal, as column (2) reveals.

[20]Notice that, while we preregistered this analysis, we did not preregister a specific hypothesis.

Table 3: Beliefs and Recall in the Baseline Experiment

Dependent variable:	Probability Guess	Correct Recall
Sample:	Pooled (1)	Pooled (2)
Constant	67.679***	0.965***
	(2.078)	(0.018)
HighComplexity	3.351	-0.007
	(2.819)	(0.026)
R^2	0.006	0.000
Observations	230	230

The table presents OLS regressions using the restricted sample of the Baseline Experiment. In column (1), the dependent variable is the probability guess for the likely state. Column (2) uses a dummy for whether respondents correctly recall the more likely model. All columns use observations from both the HighComplexity and LowComplexity conditions. Stars highlight significant differences from 0 with * for $p < 0.10$, ** for $p < 0.05$, *** for $p < 0.01$. Standard errors are clustered on the subject level.



(a) Histogram of signed probability guesses.

(b) Mean probability guesses for "The CEO is key".

Figure 4: Beliefs in the Baseline Experiment. *In Panel (a) beliefs are converted into the direction of the more likely model, so that 65 corresponds to the Bayesian probability. Panel (b) plots mean probability guesses for the model "The CEO is key". The figure is based on the restricted sample of the Baseline Experiment with 230 participants.*

We also investigate belief patterns in the Investment 1, Investment 2, Baseline Confidence, Baseline Confidence 75% and No Hover experiments, and find no significant treatment differences in beliefs and recall in any of these studies (see the respective sections of Online Appendix B). In the Beliefs First Experiment, we find a marginally significant positive effect of complexity on probability guesses (Online Appendix Table B.15).[21] Taken together, these results suggest that the complexity-induced neglect of model uncertainty in actions does not fully translate into corresponding belief patterns.

Result 4. *The neglect of model uncertainty in actions does not translate into beliefs about model uncertainty. These beliefs appear to reflect model uncertainty and are largely unaffected by complexity.*

Robustness: Delayed Belief Elicitation. In an additional pilot study that was not preregistered (see Table A.1), we investigated whether adding a delay between actions (value guesses) and belief elicitation about model uncertainty could potentially induce treatment differences in beliefs. It is conceivable that it might take time for the simplification in the action space to also carry over to beliefs. The rationale for this study was hence to test whether, even though simplification-induced misperceptions do not form immediately, they might consolidate over time as participants reflect on their choices. The design was exactly as described in Section 2, with an additional second part that took place one day after the initial survey. Respondents in the *Immediate* condition stated their beliefs immediately after the company value guesses as in the Baseline Survey, and completed unrelated tasks during the second survey. Respondents in the *Delay* condition completed the company value guesses and unrelated tasks during the first part of the survey, and the belief elicitation during the second survey one day later.

We first note that also in this experiment, we see that complexity induces people to neglect model uncertainty in actions. However, as Online Appendix B.5 reveals, both with and without delay, simplification does not induce biased beliefs about model uncertainty.

[21]In Supplementary Material D we present the results for the lenient samples. Here, we find a (marginally) stronger belief response under complexity in the Baseline and Baseline Confidence studies, and no significant treatment differences in Investment Experiment 1, Investment Experiment 2 and the No Hover and Baseline Confidence 75% Experiments (Supplementary Material D.9 and D.8).We find a significant positive effect in the Beliefs First Experiment (Supplementary Material D.10).

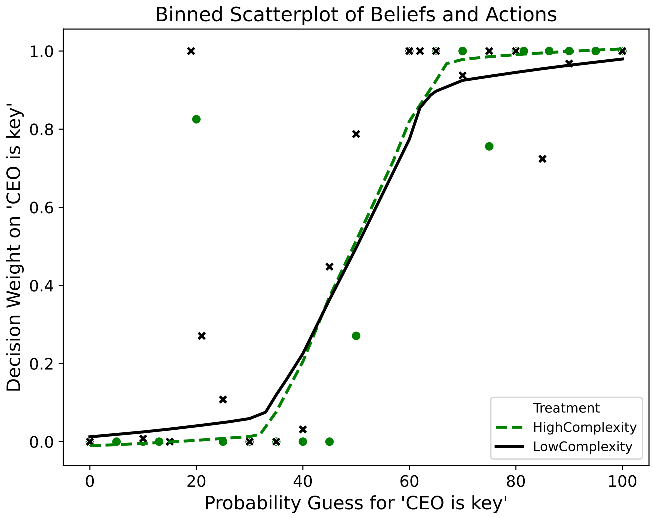


Figure 5: The relationship between decisions and beliefs. *The figure shows a binned scatterplot using the restricted sample of the Baseline Experiment with 230 participants. It has the probability guess that "The CEO is key" is the more likely model on the horizontal axis, and the decision weight γ as defined in Equation 3 on the vertical axis. The lines show LOWESS regressions based on all datapoints.*

To summarize, adding a one day delay does not induce treatment differences both in stated beliefs and correct recall of the received signal.

Results on Belief-Action Link. Notice that our results imply a complexity-induced wedge between actions and beliefs. A recent literature has investigated the link between beliefs and actions (Giglio, Maggiori, Stroebel, and Utkus (2021), Ameriks, Kézdi, Lee, and Shapiro (2020), Beutel and M. Weber (2023), Laudenbach, A. Weber, R. Weber, and Wohlfart (forthcoming)), and the determinants of how strongly beliefs translate into actions. Charles, Frydman, and Kilic (2024) find that increased complexity of forming a belief weakens the transmission of beliefs into actions. Similarly, Yang (2023) as well as Enke, Graeber, Oprea, and Yang (2024) highlight a link between information processing constraints and a weak elasticity of decisions with respect to economic fundamentals.

To investigate this wedge in our experiment more formally, Figure 5 presents a binned scatterplot with beliefs on the horizontal axis, and corresponding actions on the vertical axis, using data from the Baseline Experiment. Beliefs are given by the probability guess for the "The CEO is key" model. To ensure comparability, actions are represented by the implicit decision weight γ on the company value proposed by the "The CEO is key" model, as defined in Equation 3.

The plot illustrates the overreaction in actions caused by the complexity of calculating the optimal guess: when moving away from a probability guess of 50%, the decision weights in *HighComplexity* quickly move to extreme values, while the response is more muted in *LowComplexity*. This implies that the effect of complexity on the belief-action link is not straightforward in our setting. In a neighborhood around 50%, complexity increases responsiveness of actions to beliefs. However, when moving to more extreme beliefs, complexity renders the action response rather flat, since naive guessing is already triggered starting from fairly moderate beliefs.

We show analogous plots for the Investment 1, Investment 2, Baseline Confidence, Baseline Confidence 75%, No Hover and Beliefs First studies in Online Appendix B, as well as results for the lenient samples in Supplementary Material D, each showing qualitatively identical results.

Result 5. *Complexity induces a wedge between beliefs and actions. When complexity is high, beliefs continue to reflect model uncertainty, actions tend to be based on neglect of model uncertainty.*

4.2 Confidence

Here we investigate whether the complexity-induced simplification of neglect of model uncertainty in actions affects how people view the optimality of their actions. This type of confidence has been shown to explain a broad range of behavioral anomalies and also mediates to what extent individual biases matter for aggregate outcomes (Enke and Graeber (2023), Enke, Graeber, and Oprea (forthcoming), Enke, Graeber, and Oprea (2023)). A typical and highly intuitive finding in the literature is that complexity reduces confidence in action optimality.

Baseline Confidence. In a replication of the Baseline Experiment, we elicited cognitive uncertainty, i.e. people's confidence in the optimality of their value estimates. The experiment was preregistered, including the analysis of the cognitive uncertainty measure (see Table A.1). As preregistered, we use the same sample restriction as for the Baseline study, resulting in a restricted sample with 336 participants.[22]

Figure 6 shows the average confidence levels by treatment condition. We find that

[22] Similar to the analysis of beliefs, while we preregistered this analysis, we did not preregister a specific hypothesis.

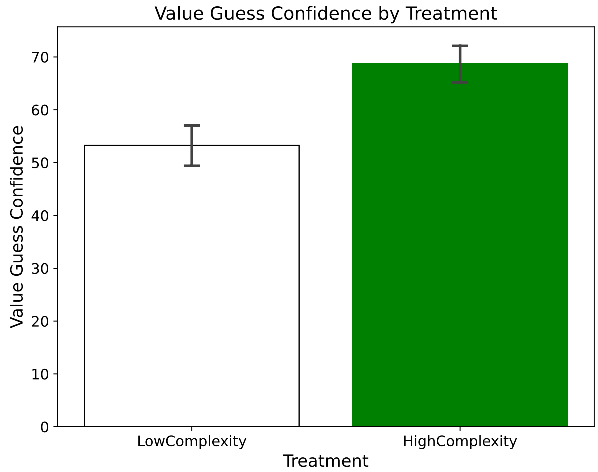
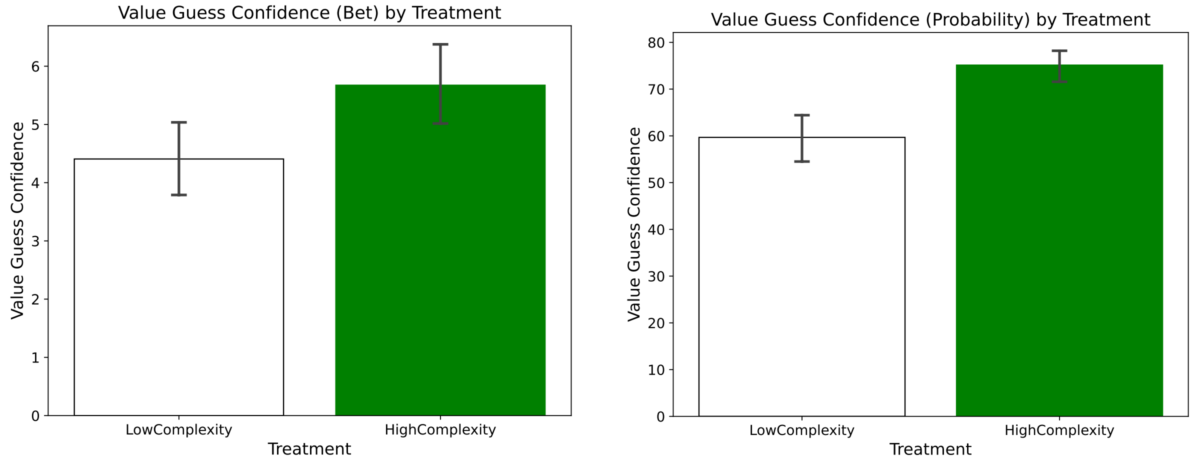


Figure 6: Average value guess confidence in the Baseline Confidence Experiment. *The figure plots the average confidence that respondents had in their company value guesses, using the restricted sample of the Baseline Confidence study with 336 participants.*

participants in the *HighComplexity* condition are significantly *more* confident in their guesses than those in the *LowComplexity* condition—despite exhibiting greater simplification and a lower rate of rational guesses under complexity, a seemingly contradictory pattern.

We also elicited the same measure in the Equally Likely Models, Baseline Confidence 75%, No Hover, and Beliefs First studies and featured the analysis in the preregistration. Online Appendix Figure B.4 confirms that the finding replicates in the Equally Likely Models study, and Figures B.27, B.31, and B.36 show that it replicates in the three additional experiments as well. Additional results presented in Supplementary Material D show that the results also hold true when using the more lenient sample restrictions.

Incentivized Confidence. In another replication of the Baseline Confidence experiment, we further added an incentivized version of the confidence elicitation. Here, participants could bet on the optimality of their guesses. They received an endowment of $10 and could choose how much to bet (between $0 and $10) on the event that their guesses had been, on average, no more than 10 points away from the best possible guess. The bet was multiplied by 3 if their guesses had indeed been accurate and was lost otherwise. After completing all company value guesses, participants answered the non-incentivized (probability) and incentivized (bet) versions on the same screen. The experiment was preregistered, including the analysis of the incentivized and non-incentivized cognitive uncertainty measures (see Table A.1). As preregistered, we use



(a) Incentivized (bet) confidence

(b) Non-incentivized (probability) confidence

Figure 7: Average value guess confidence in the Incentivized Confidence Experiment. *Panel (a) plots the average incentivized confidence measure, while panel (b) plots the non-incentivized measure, both using the restricted sample of the Incentivized Confidence study with 203 participants.*

the same sample restriction as for the Baseline study, resulting in a restricted sample with 203 participants. We report the results on confidence for this experiment in the main text and refer the reader to Online Appendix B.7 for the remaining analyses.

Figure 7 plots the average confidence measures. Panel (b) replicates the previous result that participants in the high-complexity condition stated a higher confidence level in the optimality of their company value guesses than their counterparts in the low-complexity condition. Panel (a) presents results from the incentivized betting task. While these data are naturally more noisy due to the betting context, the Figure again confirms our result that higher complexity yields higher confidence. Supplementary Material D.7 produces these results under the more lenient sample restrictions, yielding qualitatively identical results.

Taken together, our results indicate that confidence in the optimality of own actions is higher in the high-complexity conditions compared to the low complexity conditions. It seems that in contrast to prior results from different contexts, in the context of model uncertainty, the possibility to respond to complexity by greatly simplifying the world through full neglect of model uncertainty leads to an illusion of certainty and hence increased confidence in action optimality.

Result 6. *Complexity leads to higher confidence in the optimality of own actions.*

5 A Simple Framework

We have shown that respondents simplify model uncertainty when computational complexity is high. This, however, does not carry over to beliefs about model uncertainty. Finally and perhaps most surprisingly, higher computational complexity increases confidence in the optimality of own actions rather than decreasing it. In this part we present a simple framework and show how it can generate this pattern of results. We view the framework as a vehicle to make sense of the data we see and to generate new hypotheses, rather than a contribution by itself.

In the framework, when faced with a decision problem, an agent first forms a mental representation of the problem. This process is shaped both by bottom-up and top-down attention. Upon being presented a decision problem, a bottom-up process of cue-dependent memory determines which of the currently stored mental representations is top of mind (Bordalo, Gennaioli, Lanzani, and Shleifer (2025)). Then, in a top-down process (Sims (2003), Kőszegi and Matějka (2020), Caplin and Dean (2015), Steiner, Stewart, and Matějka (2017), and Gabaix (2014)), the agent decides whether they want to further simplify this representation. This section contains the basic intuitions, while Supplementary Material E presents the formal framework.

This framework formalizes economic decision making as a cognitively constrained process operating over a structured internal database of mental representations. Initially, this database comprises only a broad representation, which encodes a probabilistic assessment of the decision environment (e.g., a 0.65 likelihood for the more probable state) along with objects of the contextual environment (e.g., action, confidence, beliefs about model uncertainty).

Each decision - whether it concerns an action, confidence judgment, or belief report - invokes a two-step cognitive process to form a mental representation of the decision problem:

1. Bottom-Up Retrieval: Agents hold a database of mental representations. When faced with a decision, similarity-based recall (Bordalo, Gennaioli, and Shleifer (2020), Bordalo, Conlon, Gennaioli, Kwon, and Shleifer (2023b), Enke, Schwerter, and Zimmermann (2024), Graeber, Roth, and Zimmermann (2024), Jiang, Liu, Peng, and Yan (forthcoming)) determines which representation is top of mind.

In other words, the representation most similar to the current decision cue is retrieved from memory.

2. Top-Down Simplification: Once a representation is top of mind, the agent, in a top-down process, evaluates whether the cognitive costs of computing a response using the retrieved representation exceed the benefits. If so, the agent further simplifies the representation to reduce processing demands.

The key insight from our framework is that people make decisions within the mental representation they formed for this specific decision problem. Hence, specific decisions (actions, confidence judgments, or belief reports) in a given underlying environment may be based on very different mental representations of that environment, allowing us to explain our seemingly contradictory pattern of results.

Estimation of Company Values. Upon observing a signal, the agent must compute an optimal guess. Since this is the first decision agents take, the database of representations only contains the broad representation of the problem. Under low complexity, cognitive costs are manageable, so participants compute the optimal guess using the full broad representation of the problem, i.e., fully taking model uncertainty into account. Under high complexity, cognitive costs may exceed the benefits, prompting simplification to a mental representation in which model uncertainty is fully neglected. Once a decision has been taken, the corresponding representation is added to the database, including their contextual features.

Confidence Elicitation. In the low complexity treatment, when faced with the confidence assessment, agents will retrieve the broad representation (their database only consists of broad representations). We assume confidence judgments to be cognitively light, based on a cognitive science literature that suggests that metacognition of this type occurs rather effortlessly (Koriat (1993), Koriat (2000), and Koriat and Levy-Sadot (1999)). Hence, no further simplification occurs. In the high complexity treatment, similarity-based retrieval favors the recall of the simplified representation. This is because the confidence elicitation cues the estimation tasks (it explicitly asks for confidence about the estimation task), which in the high-complexity condition led to simplification. As described above, confidence in own action optimality is then assessed by agents within this mental representation. Consequently, confidence can be elevated by complexity if confidence within the simplified high complexity environment is higher than in the full representation of the low complexity environment.[23]

Belief Elicitation. In the low complexity treatment, the broad representation is again retrieved and used without simplification, yielding probabilistically grounded belief reports about model uncertainty. In the high complexity treatment, retrieval of mental representations favors the broad representation over the simple one, since the belief task directly asks about model uncertainty and is hence more similar to the broad representation. We assume the cognitive costs of belief elicitation to be negligible (arguably belief formation already occurred). Hence, no further simplification occurs. People then state the beliefs of their mental representation and therefore beliefs in both treatments will tend to reflect model uncertainty.

A Testable Prediction. The framework yields a sharp testable prediction regarding the confidence illusion in *HighComplexity*. Confidence is inflated because, when the confidence question is asked, similarity-based retrieval favors the simplified representation: the confidence question cues the prior estimation task, in which *HighComplexity* subjects simplified to a single model. Confidence is then assessed within this simplified representation, which encodes no residual model uncertainty, producing the certainty illusion. This logic implies that any cue presented immediately before the confidence question that makes the broad representation more salient should override the default retrieval of the simplified one. If the broad representation is instead retrieved, agents assess their confidence within a mental model that still encodes genuine uncertainty about which model is correct, thereby reducing confidence relative to agents who receive no such cue. We test this prediction in Section 6.

[23]While this seems plausible, our model does not formalize why confidence within the simplified high complexity environment may be higher than in the full representation of the low complexity environment. The key insight from our framework is that confidence is assessed within the mental representation of the problem.

6 Mechanism Evidence

Motivated by the framework in the previous section, in this section we seek to provide further empirical evidence that sheds some light on underlying mechanisms. We focus on our confidence result as it is perhaps the most striking and since our framework yields sharp testable predictions here. Specifically, we explore, using different empirical approaches, whether the higher confidence under *HighComplexity* is due to respondents assessing confidence within the simplified representation of reality (full uncertainty neglect).

Existing Data. First, using existing data, we explore whether the higher confidence under *HighComplexity* is driven by the act of simplification itself—that is, by behaving as if one model is definitively correct. Since the naive guess removes all ambiguity about which model to use, it may create an illusion of certainty that inflates confidence. Confidence is always elicited after value guesses, so reverse causality is unlikely. The analysis is at the subject level and pools all four studies with non-incentivized confidence measures; it was not preregistered.

The first evidence comes from within-condition comparisons, reported in columns (4) and (5) of Table 4. In *HighComplexity*, subjects who simplified more—acting as if the signal-consistent model were definitely correct—reported significantly higher confidence (Share Naive coefficient: 17.32, column 4). Importantly, within *HighComplexity* the naive subjects are precisely those who exerted *less* effort: by focusing only on one model they avoided the harder task of computing and weighing both. An effort-justification account would predict that subjects who worked more should be more confident; the data show the opposite. The same positive relationship between naive guessing and confidence holds within *LowComplexity* (18.46, column 5), where computation was minimal for all subjects, further suggesting that it is the act of simplification itself rather than the effort exerted that drives confidence. The near-identical magnitude across conditions, confirmed by the insignificant interaction term in column (3), reinforces this conclusion.

Column (1) shows the raw treatment gap: *HighComplexity* subjects report confidence 12.96 points higher on average. Column (2) shows that adding Share Naive—which is substantially higher under complexity—reduces this gap by 25%, consistent with simpli-

Table 4: Naive Guessing and the Confidence Gap

Dependent variable:	Confidence				
Sample:	Pooled (1)	Pooled (2)	Pooled (3)	HC (4)	LC (5)
HighComplexity	12.96***	9.70***	10.49***		
	(1.22)	(1.22)	(3.42)		
Share Naive		18.02***	18.34***	17.32***	18.46***
		(1.77)	(2.07)	(3.34)	(2.08)
HighComplexity × Share Naive			-1.02		
			(3.89)		
Study FEs	Yes	Yes	Yes	Yes	Yes
R^2	0.076	0.146	0.146	0.055	0.108
Observations	1481	1481	1481	743	738

Notes: OLS regressions at the subject level (strict sample). The dependent variable is value guess confidence (0–100). The data pool four studies with non-incentivized confidence measures (Baseline, Baseline Confidence 75%, No Hover, Beliefs First). Share Naive is the subject-level share of decisions where the value guess equals the naive benchmark. Columns (4)–(5) restrict to HighComplexity and LowComplexity subjects, respectively. All specifications include study fixed effects. Heteroskedasticity-robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

fication being one channel through which complexity raises confidence. Together, these results suggest that naive guessing creates an illusion of certainty that inflates confidence at the individual level, and that the greater prevalence of naive guessing under high complexity accounts for a meaningful share of the treatment gap.

Open Text Evidence. In two studies (NoHover and BeliefsFirst, $N = 793$), participants were asked to explain their confidence immediately after the confidence question in an open-text format:

> *"You just stated your confidence in your guesses for the values of 8 companies. On this screen, we are interested in how you approached this question. Please briefly describe why you are/are not confident in the accuracy of your guesses."*

We use an LLM (OpenAI GPT-4o, temperature = 0) to code each response on binary dimensions. All regressions pool both studies with study fixed effects. Online Appendix C reports details on the LLM-coding approach. The confidence responses directly corroborate the model's mechanism. Our model predicts that *HighComplexity* subjects

Table 5: LLM-Coded Open Text Responses: Confidence Elicitation

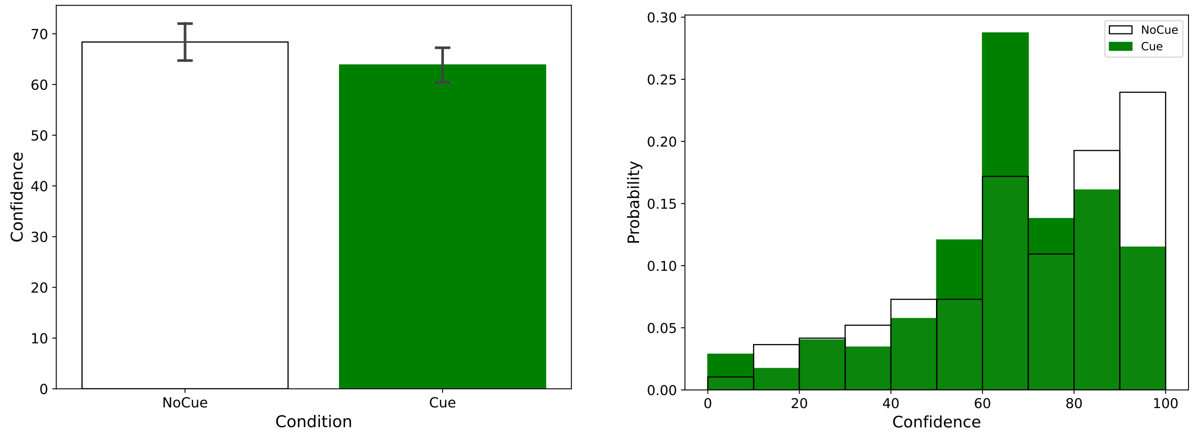
Dimension	HighComplexity	LowComplexity	Difference (SE)	p-Value
Model Uncertainty	0.449	0.528	−0.081 (0.051)	0.113
High Calculation Quality	0.251	0.015	+0.235 (0.033)	<0.001***

Notes: Share of subjects coded 1 on each binary dimension, by treatment condition. Difference (SE) is the OLS coefficient on *HighComplexity* with study fixed effects and robust standard errors, pooling NoHover and BeliefsFirst ($N = 384$). Coding prompt and dimension definitions are reported in Online Appendix C. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

assess confidence within the simplified representation—in which one model is taken as definitively correct—while *LowComplexity* subjects assess confidence within the broad representation, which retains residual model uncertainty. Consistent with this, *LowComplexity* subjects more often acknowledge residual model uncertainty (53% vs. 45%, $p = 0.11$), consistent with operating within the broad representation. *HighComplexity* subjects, by contrast, are far more likely to explain their confidence by referring to how carefully they performed their calculations (25% vs. 2%, $p < 0.001$): having collapsed to the simplified representation in which only one model's formula is computed, they treat the quality of that calculation as sufficient grounds for confidence, without awareness that they ignored the alternative. Table 5 summarizes the results.

We also elicited open-ended descriptions of how participants arrived at their value guesses. Consistent with the main results, *HighComplexity* subjects are more likely to describe using only the signal-consistent model (79% vs. 72%, $p = 0.083$), while *LowComplexity* subjects more often describe combining both model estimates (12% vs. 4%, $p = 0.003$). Results are reported in Online Appendix Table C.1.

Visual Cue Experiment. To further probe the mechanism and directly test the prediction developed in Section 5, we conducted an additional preregistered experiment. After completing their value guesses, but immediately before the confidence elicitation, participants in the *Cue* condition were shown the urn from which their signal had been drawn—65 signal-consistent balls and 35 signal-inconsistent balls—reminding them of the residual uncertainty they had faced throughout the task. Participants in the *NoCue* condition proceeded directly to the confidence question. Both conditions featured the high-complexity version of the company valuation tasks where participants had to calculate estimates themselves. If the confidence illusion is driven by the retrieval of the simplified representation at the time of confidence elicitation, a cue that makes the broad representation more salient should at least partially disrupt it and reduce confidence. Since the cue is shown only after all value guesses are complete, it cannot affect actions.



(a) Mean confidence by condition

(b) Distribution of confidence by condition

Figure 8: Confidence in the Visual Cue Experiment (strict sample, 366 participants). *NoCue* participants proceeded directly to the confidence question after completing their value guesses; *Cue* participants were first shown the urn composition underlying their signal. Confidence is measured on a 0–100 scale.

Figure 8 presents the results from the strict sample. The left panel plots average confidence by condition. The *Cue* mean (63.9) is lower than the *NoCue* mean (68.4), and the difference of -4.52 points (SE $= 2.50$) is marginally significant ($p = 0.070$), consistent with the predicted direction. The point estimate accounts for roughly one third of the overall confidence gap documented in the main results.

The right panel reveals a distributional pattern that complements the mean effect. In the *NoCue* condition, roughly 24% of subjects report confidence in the 90–100 range, reflecting the highly polarised nature of the simplification-induced certainty illusion. In the *Cue* condition, this share falls to approximately 12%, with the displaced mass accumulating in the 60–70 range, which rises from 17% to 29%. The cue thus appears to temper the most extreme confidence responses—nudging subjects who would otherwise report near-certainty down to intermediate confidence—while leaving the lower end of the distribution largely unchanged. This pattern is consistent with the cue targeting subjects with initially high confidence: those who fully embraced the certainty illusion are nudged downward, while subjects who were already reporting intermediate confidence are largely unaffected. Supplementary Material D.11 presents the results under the more lenient sample restrictions, yielding qualitatively identical results.

7 Conclusion

This paper explores how individuals navigate model uncertainty in economic decision-making and demonstrates that complexity significantly influences the way people deal with model uncertainty. Through a controlled experimental framework, we investigate whether individuals account for model uncertainty when making decisions or instead simplify the world by implicitly assuming one model is correct. Our findings reveal that when model complexity is high, individuals tend to neglect model uncertainty in their actions, behaving as if one model is definitively correct. However, this neglect of uncertainty does not translate into beliefs, as participants' stated beliefs continue to reflect model uncertainty. This creates a systematic wedge between actions and beliefs. Furthermore, our results show that complexity-induced simplification leads to increased confidence in decision optimality, contradicting prior findings that suggest complexity typically raises cognitive uncertainty.

Our results provide a direct test of the widely held assumption in the theoretical literature on misspecified models that people attend to a single model when making decisions, rather than entertaining multiple weighted models simultaneously.

By systematically manipulating the complexity of models, we provide robust evidence that individuals simplify complex problems by focusing on a single mental model. Data on attention allocation further support this conclusion, showing that participants in high-complexity conditions spend more time attending to one model, rather than considering both models equally. This attention-based mechanism helps explain why complexity amplifies the tendency to neglect model uncertainty in actions. The robustness of these findings is confirmed across different tasks, including an investment decision context, demonstrating that the observed pattern extends beyond the specific valuation task used in our primary experiment.

Our findings contribute to a deeper understanding of how complexity shapes economic cognition. Existing research on mental models has largely focused on selection and persuasion, whereas our study highlights a novel dimension: the impact of complexity on how people handle competing models. Our findings reveal that, when complexity is high, in order to be able to operate and make decisions, people need to simplify the world by acting as if only one model exists. Interestingly, this complexity-induced simplification can lead to an illusion of certainty, where people are confident about the optimality of their actions.

While we do not study persuasion directly, an intuitive implication of our results may be that, when complexity is high, the presence of model uncertainty might make decision-makers more susceptible to persuasive narratives that present a single, seemingly definitive interpretation of economic realities.

References

Aina, Chiara (2024). “Tailored Stories”. In: *Working Paper*.

Aina, Chiara and Florian Schneider (2025). “Weighting Competing Models”. In: *Working Paper*.

Ambuehl, Sandro and Heidi C. Thysen (2024). “Choosing Between Causal Interpretations: An Experimental Study”. Working Paper.

Ameriks, John, Gábor Kézdi, Minjoon Lee, and Matthew D Shapiro (2020). “Heterogeneity in expectations, risk tolerance, and household stock shares: The attenuation puzzle”. In: *Journal of Business & Economic Statistics* 38.3, pp. 633–646.

Andre, Peter, Ingar Haaland, Christopher Roth, and Johannes Wohlfahrt (2024). “Narratives about the Macroeconomy”. In: *Working Paper*.

Arrieta, Gonzalo and Kirby Nielsen (2024). “Procedural Decision-Making In The Face Of Complexity”. In: *Working Paper*.

Augenblick, Ned, Eben Lazarus, and Michael Thaler (forthcoming). “Overinference from Weak Signals and Underinference from Strong Signals”. In: *Quarterly Journal of Economics*.

Ba, Cuimin, J Aislinn Bohren, and Alex Imas (2022). “Over-and underreaction to information”. In: *Available at SSRN 4274617*.

Barron, Kai and Tilman Fries (2024). “Narrative Persuasion”. In: *Working Paper*.

Benjamin, Daniel J (2019). “Errors in probabilistic reasoning and judgment biases”. In: *Handbook of Behavioral Economics: Applications and Foundations 1* 2, pp. 69–186.

Beutel, Johannes and Michael Weber (2023). “Beliefs and portfolios:Causal evidence”. In: *Working Paper*.

Bordalo, Pedro, John Conlon, Nicola Gennaioli, Spencer Kwon, and Andrei Shleifer (2023a). “How People Use Statistics”. In: *Working Paper*.

— (2023b). “Memory and Probability”. In: *Quarterly Journal of Economics* 138.1, pp. 265–311.

Bordalo, Pedro, Nicola Gennaioli, Giacomo Lanzani, and Andrei Shleifer (2025). “A Cognitive Theory of Reasoning and Choice.” In: *Working Paper*.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2012). “Salience theory of choice under risk”. In: *The Quarterly journal of economics* 127.3, pp. 1243–1285.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2020). “Memory, Attention, and Choice”. In: *The Quarterly Journal of Economics* 135.3, pp. 1399–1442. DOI: 10.1093/qje/qjaa012.

Bushong, Benjamin, Matthew Rabin, and Joshua Schwartzstein (2021). “A model of relative thinking”. In: *The Review of Economic Studies* 88.1, pp. 162–191.

Caplin, Andrew and Mark Dean (2015). “Revealed Preference, Rational Inattention, and Costly Information Acquisition”. In: *American Economic Review* 105.7, pp. 2183–2203. DOI: 10.1257/aer.20140117.

Charles, Constantin, Cary Frydman, and Mete Kilic (2024). “Insensitive investors”. In: *The Journal of Finance* 79.4, pp. 2473–2503.

Charles, Constantin and Chad Kendall (2025). “Causal Narratives”. Working Paper.

Danz, David, Lise Vesterlund, and Alistair Wilson (2022). “Belief Elicitation and Behavioral Incentive Compatibility”. In: *American Economic Review* 112.9, pp. 2851–2883.

Dertwinkel-Kalt, Markus, Holger Gerhardt, Gerhard Riener, Frederik Schwerter, and Louis Strang (2022). “Concentration bias in intertemporal choice”. In: *The Review of Economic Studies* 89.3, pp. 1314–1334.

Eliaz, Kfir and Rani Spiegler (2022). “A Model of Competing Narratives”. In: *American Economic Review* 110.12, pp. 3786–3816.

Enke, Benjamin (2020). “What you see is all there is”. In: *The Quarterly Journal of Economics* 135.3, pp. 1363–1398.

— (2024). “The Cognitive Turn in Behavioral Economics”. In: *Working Paper*.

Enke, Benjamin and Thomas Graeber (2023). “Cognitive Uncertainty”. In: *Quarterly Journal of Economics* 138.4, pp. 2021–2067.

Enke, Benjamin, Thomas Graeber, and Ryan Oprea (2023). “Confidence, Self-selection and Bias in the Aggregate”. In: *American Economic Review* 113.7, pp. 1933–1966.

— (forthcoming). “Complexity and Time”. In: *Journal of the European Economic Association*.

Enke, Benjamin, Thomas Graeber, Ryan Oprea, and Jeffrey Yang (2024). *Behavioral Attenuation*. Tech. rep. National Bureau of Economic Research.

Enke, Benjamin, Frederik Schwerter, and Florian Zimmermann (2024). “Associative Memory, Beliefs and Market Interactions”. In: *Journal of Financial Economics* 157, p. 103853. DOI: 10.1016/j.jfineco.2024.103853.

Enke, Benjamin and Cassidy Shubatt (2024). “Quantifying Lottery Choice Complexity”. In: *Working Paper*.

Enke, Benjamin and Florian Zimmermann (2019). “Correlation neglect in belief formation”. In: *The review of economic studies* 86.1, pp. 313–332.

Esponda, Ignacio, Ryan Oprea, and Sevgi Yuksel (2023). “Seeing What is Representative”. In: *Quarterly Journal of Economics* 138.4, pp. 2607–2657.

Esponda, Ignacio, Emanuel Vespa, and Sevgi Yuksel (2024). “Mental Models and Learning: The Case of Base-Rate Neglect”. In: *American Economic Review* 114.3, pp. 752–782. DOI: 10.1257/aer.20201004.

Fan, Tony, Yucheng Liang, and Cameron Peng (2024). “The Inference-Forecast Gap in Belief Updating”. In: *Working Paper*.

Fréchette, Guillaume R., Emanuel Vespa, and Sevgi Yuksel (2024). “Extracting Models From Data Sets: An Experiment”. Working Paper.

Frick, Mira, Ryota Iijima, and Yuhta Ishii (2022). “Dispersed behavior and perceptions in assortative societies”. In: *American Economic Review* 112.9, pp. 3063–3105.

Gabaix, Xavier (2014). “A Sparsity-Based Model of Bounded Rationality”. In: *The Quarterly Journal of Economics* 129.4, pp. 1661–1710. DOI: 10.1093/qje/qju024.

Gagnon-Bartsch, Tristan, Matthew Rabin, and Joshua Schwartzstein (2023). “Channeled attention and stable errors”. In: *Working Paper*.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus (2021). “Five facts about beliefs and portfolios”. In: *American Economic Review* 111.5, pp. 1481–1522.

Graeber, Thomas (2023). “Inattentive inference”. In: *Journal of the European Economic Association* 21.2, pp. 560–592.

Graeber, Thomas, Christopher Roth, and Constantin Schesch (2024). “Explanations”. In: *Working Paper*.

Graeber, Thomas, Christopher Roth, and Florian Zimmermann (2024). “Stories, statistics, and memory”. In: *Quarterly Journal of Economics*.

Hartzmark, Samuel, Samuel Hirshman, and Alex Imas (2021). “Ownership, Learning, and Beliefs”. In: *Quarterly Journal of Economics* 136.3, pp. 1665–1717.

Heidhues, Paul, Botond Kőszegi, and Philipp Strack (2018). “Unrealistic Expectations and Misguided Learning”. In: *Econometrica* 86.4, pp. 1159–1214.

Heidhues, Paul, Botond Kőszegi, and Philipp Strack (2023). “Misinterpreting Yourself”. In: *Working Paper*.

Jiang, Zhengyang, Hongqi Liu, Cameron Peng, and Hongjun Yan (forthcoming). “Investor Memory and Biased Beliefs: Evidence from the Field”. In: *The Quarterly Journal of Economics*.

Kendall, Chad and Ryan Oprea (2024). “On the Complexity of Forming Mental Models”. In: *Quantitative Economics* 15, pp. 175–211.

Koriat, Asher (1993). “How do we know that we know? The accessibility model of the feeling of knowing”. In: *Psychological Review* 100.4, pp. 609–639. DOI: 10.1037/0033-295X.100.4.609.

— (2000). “The feeling of knowing: Some metatheoretical implications for consciousness and control”. In: *Consciousness and Cognition* 9.2, pp. 149–171. DOI: 10.1006/ccog.2000.0433.

Koriat, Asher and Ravit Levy-Sadot (1999). “Processes underlying metacognitive judgments: Information-based and experience-based monitoring of one’s own knowledge”. In: *Dual-Process Theories in Social Psychology*. Ed. by Shelly Chaiken and Yaacov Trope. New York: Guilford Press, pp. 483–502.

Kőszegi, Botond and Filip Matějka (May 2020). “Choice Simplification: A Theory of Mental Budgeting and Naive Diversification”. In: *The Quarterly Journal of Economics* 135.2, pp. 1153–1207. DOI: 10.1093/qje/qjz043.

Kőszegi, Botond and Adam Szeidl (2013). “A model of focusing in economic choice”. In: *The Quarterly journal of economics* 128.1, pp. 53–104.

Laudenbach, Christina, Annika Weber, Rüdiger Weber, and Johannes Wohlfart (forthcoming). “Beliefs About the Stock Market and Investment Choices: Evidence from a Field Experiment”. In: *Review of Financial Studies*.

Mailath, George J and Larry Samuelson (2019). “The Wisdom of a Confused Crowd: Model-Based Inference”. In: *Working Paper*.

Montiel Olea, José Luis, Pietro Ortoleva, Mallesh M Pai, and Andrea Prat (2022). “Competing models”. In: *The Quarterly Journal of Economics* 137.4, pp. 2419–2457.

Mullainathan, Sendhil, Joshua Schwartzstein, and Andrei Shleifer (2008). “Coarse Thinking and Persuasion”. In: *The Quarterly Journal of Economics* 123.2, pp. 577–619. DOI: 10.1162/qjec.2008.123.2.577.

Oprea, Ryan (2020). “What Makes a Rule Complex?” In: *American Economic Review* 110.12, pp. 3913–3951.

Schwartzstein, Joshua (2014). “Selective attention and learning”. In: *Journal of the European Economic Association* 12.6, pp. 1423–1452.

Schwartzstein, Joshua and Adi Sunderam (2021). “Using models to persuade”. In: *American Economic Review* 111.1, pp. 276–323.

Shiller, Robert (2017). “Narrative Economics”. In: *American Economic Review* 107.4, pp. 967–1004.

Shubatt, Cassidy and Jeffrey Yang (2024). *Tradeoffs and Comparison Complexity*. Tech. rep. Working Paper.

Sims, Christopher A. (Apr. 2003). “Implications of Rational Inattention”. In: *Journal of Monetary Economics* 50.3, pp. 665–690. DOI: 10.1016/S0304-3932(03)00029-1.

Somerville, Jason (2022). “Range-Dependent Attribute Weighting in Consumer Choice: An Experimental Test”. In: *Econometrica* 90.2, pp. 799–830.

Steiner, Jakub and Colin Stewart (2015). “Price Distortions under Coarse Reasoning with Frequent Trade”. In: *Journal of Economic Theory* 159, pp. 574–595. DOI: 10.1016/j.jet.2015.07.011.

Steiner, Jakub, Colin Stewart, and Filip Matějka (2017). “Rational Inattention Dynamics: Inertia and Delay in Decision-Making”. In: *Econometrica* 85.2, pp. 521–553. DOI: 10.3982/ECTA13636.

Yang, Jeffrey (2023). “On the Decision-Relevance of Subjective Beliefs”. In: *Available at SSRN 4425080*.

Appendix

Overview of Data Collections

Table A.1: Overview of Data Collections

Collection	Participants	Description	Link to pre-analysis plan
Baseline	600	As described in Section 2. Treatments: HighComplexity and LowComplexity. Outcomes: Company value guesses, hover times, beliefs.	https://aspredicted.org/t79h-2mkj.pdf
Equally Likely Models	600	As Baseline, but without noisy indication, hence 50-50 belief about more likely model, and with additional guess confidence measure. Treatments: HighComplexity and LowComplexity. Outcomes: Company value guesses, guess confidence, hover times.	https://aspredicted.org/92cr-9kwd.pdf
Investment 1	400	As Baseline, but company value guesses replaced by investment decisions. Treatments: HighComplexity and LowComplexity. Outcomes: Investment decisions, hover times, beliefs.	https://aspredicted.org/4c38-7psb.pdf
Investment 2	600	As Investment, but larger sample size. Treatments: HighComplexity and LowComplexity. Outcomes: Investment decisions, hover times, beliefs.	https://aspredicted.org/7vcv-gq8z.pdf
Delayed Belief Elicitation	588	As Baseline, but additional variation in the timing of belief elicitation: For half the respondents, beliefs are elicited immediately as in Baseline, for the other half they are elicited with a one-day delay. Treatments: (HighComplexity, LowComplexity) × (Immediate, Recall). Outcomes: Company value guesses, hover times, beliefs.	Not preregistered
Baseline Confidence	600	As Baseline, but with additional guess confidence measure. Treatments: HighComplexity and LowComplexity. Outcomes: Company value guesses, guess confidence, hover times, beliefs.	https://aspredicted.org/v6cv-y9yh.pdf
Incentivized Confidence	600	As Baseline Confidence, but with additional incentivized guess confidence measure on the same page as the non-incentivized measure. Treatments: HighComplexity and LowComplexity. Outcomes: Company value guesses, guess confidence, hover times, beliefs.	https://aspredicted.org/txsk-3hky.pdf
Baseline Confidence 75%	600	As Baseline Confidence, but the signal is correct with 75% probability instead of 65%. Treatments: HighComplexity and LowComplexity. Outcomes: Company value guesses, guess confidence, hover times, beliefs, open text responses on either guesses or guess confidence.	https://aspredicted.org/ud3y9p.pdf
No Hover	600	As Baseline Confidence, but variables and company value estimates are directly displayed without the need to hover. Treatments: HighComplexity and LowComplexity. Outcomes: Company value guesses, guess confidence, beliefs.	https://aspredicted.org/ig6jk5.pdf
Beliefs First	600	As Baseline Confidence, but beliefs about the correct rule are elicited before participants make their company value guesses. Treatments: HighComplexity and LowComplexity. Outcomes: Company value guesses, guess confidence, hover times, beliefs, open text responses on either guesses or guess confidence.	https://aspredicted.org/b5mm45.pdf
Visual Cue	600	As Baseline Confidence, but after completing all value guesses and immediately before the confidence elicitation, participants in the Cue condition were shown the urn composition underlying their signal (65 signal-consistent, 35 signal-inconsistent balls). Both conditions use HighComplexity only. Treatments: Cue and NoCue. Outcomes: Company value guesses, guess confidence, hover times, beliefs.	https://aspredicted.org/4hn23x.pdf

This Table provides an overview of the different data collections. The sample sizes refer to the size of the original data collection, prior to applying exclusion restrictions.

Accompanying materials are available online: Online Appendix (Sections B and C); Supplementary Material D (Additional Results); Supplementary Material E (Model Framework); and Supplementary Material F (Experimental Instructions).

Additional Pooled Results

Here, we present additional non-preregistered analyses, pooling the three studies with an identical action stage: the Baseline Experiment, the Baseline Confidence Experiment, and the Incentivized Confidence Experiment (769 participants in the restricted sample).

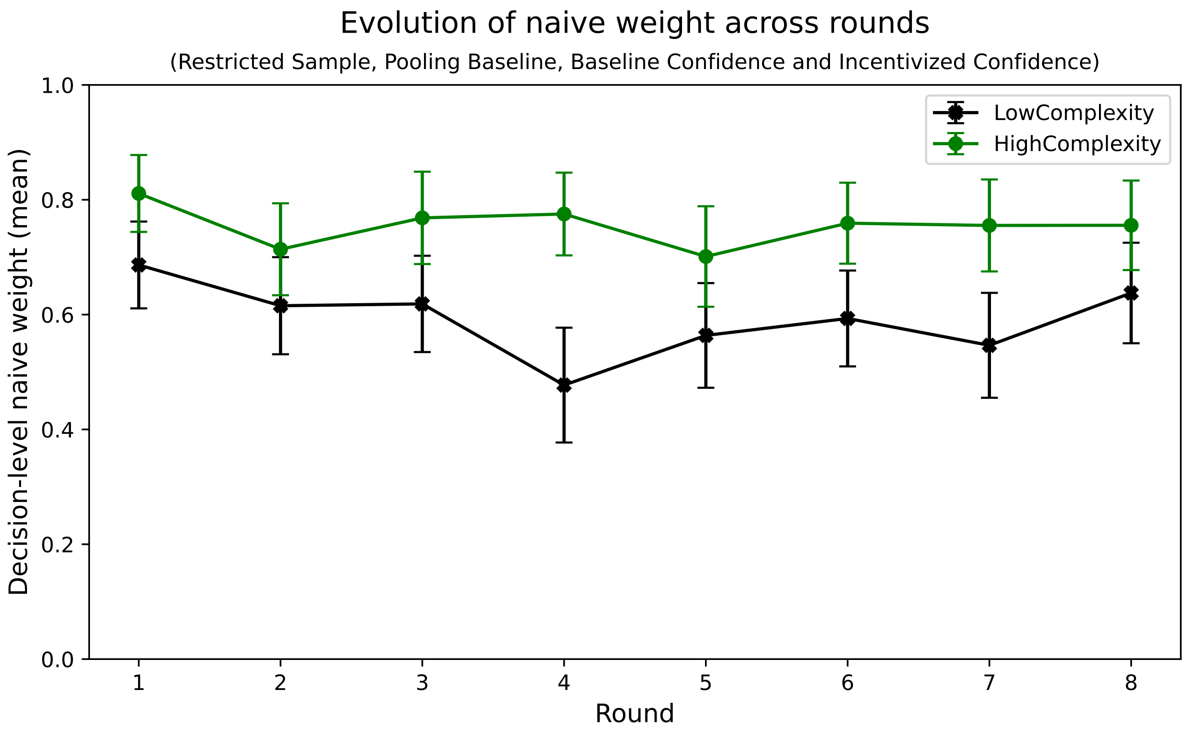


Figure A.1: Evolution of the decision-level naive weight across rounds. *The figure plots the mean decision-level naive weight λ by treatment and decision round, using the restricted samples of the Baseline Experiment (230 participants), the Baseline Confidence Experiment (336 participants), and the Incentivized Confidence Experiment (203 participants). Error bars indicate 95% confidence intervals.*

Table A.2: Naive Weight by Consistent Hovering Share Bin

Consistent Hover Share	λ_{Low}	λ_{High}	Difference	p-value	N_{Low}	N_{High}
0.0–0.2	−0.412	−0.796	−0.384	0.095	129	146
0.2–0.4	0.260	0.218	−0.042	0.754	298	165
0.4–0.6	0.534	0.494	−0.040	0.637	1025	456
0.6–0.8	0.634	0.590	−0.044	0.554	726	245
0.8–1.0	0.901	0.978	0.078	0.002	838	2124
Full Sample	0.592	0.755	0.163	0.000	3016	3136

The table reports mean naive weights (λ) by bin of decision-level consistent hovering share, pooling the restricted samples of the Baseline, Baseline Confidence, and Incentivized Confidence Experiments (6152 decisions, 769 subjects total). The consistent hovering share is defined as the ratio of hover time spent on the signal-consistent rule to total hover time on both rules. The p-value refers to the difference in λ between *HighComplexity* and *LowComplexity* within each bin, with standard errors clustered by participant.