Model Uncertainty

Robin Musolff Florian Zimmermann

Supplementary Material: Model Framework

E Model Framework

We present a simple framework that can generate the key results of our experiments. The framework formalizes the construction of a mental representation of the decision problem. This process is shaped both by bottom-up (Bordalo, Gennaioli, Lanzani, and Shleifer (2025)) and top-down attention (Kőszegi and Matějka (2020), Steiner, Stewart, and Matějka (2017), and Gabaix (2014)). Upon presentation of a decision-problem, a bottom-up process of cue-dependent memory determines which of the currently stored mental representation is top of mind. Then, in a top-down process, the agent decides whether they want to further simplify this representation.

E.1 Setup

The framework closely follows our experimental environment. There are two models of the world that can explain how company values are determined, i.e. $m \in \{A, B\}$. While only one model is correct, ex-ante there is a 50-50 chance of each model being correct. After observing a noisy but informative signal, the Bayesian posterior is given by $\pi = \Pr(m = A) = 0.65$, i.e. we assume without loss of generality that A is the model that is more likely to be correct.

E.2 Representations

When faced with a decision (action, confidence or belief elicitation), the agent forms a mental representation of the decision problem. We assume that at any moment t the agent holds a database (a finite set) of representations

$$\mathbf{R}_t = \left\{ r_i = (\hat{\pi}(r_i), V_{r_i}^{\text{context}}) \right\}$$

The scalar $\hat{\pi}(r_i)$ stores the probability attached to model A, while $V_{r_i}^{\text{context}}$ is a vector of length M that collects the contextual features of the environment in which r_i was formed, including the objects involved (action, confidence, belief about model uncertainty). The latter are stored as binary (Yes, No) features, where Yes implies that the object is present.

We model the formation of a mental representation as a two stage process. The first stage is based on bottom-up attention and similarity based recall (Bordalo, Gennaioli, Lanzani, and Shleifer (2025)), while the second stage is a top-down attention decision of whether or not to simplify the representation that was recalled in the first stage (Sims (2003), Kőszegi and Matějka (2020), Caplin and Dean (2015), Steiner, Stewart, and Matějka (2017), and Gabaix (2014)).

First Stage: Bottom-Up Similarity Based Recall. When a decision screen appears, we model this as a cue vector ξ. The agent then retrieves the stored representations based on the similarity of the contextual features V_r^{context} and ξ. The cue vector is also of length M and binary features of which objects are part of the vector (action, confidence, belief about model uncertainty). The recall probability $p(r)$ for a given representation is given by the relative similarity of its contextual features to the cue:

$$p(r) = \frac{S\left(\xi,\, V_r^{\text{context}}\right)}{\sum_{r' \in \mathbf{R}_t} S\left(\xi,\, V_{r'}^{\text{context}}\right)}, \tag{E.1}$$

where $S(\cdot,\cdot)$ is the similarity function and $\mathbf{R}_t$ is the current set of stored representations. We define $S(\cdot,\cdot)$ as follows,

$$S\left(\xi,\, V_r^{\text{context}}\right) : \prod_{i=1}^{M} \xi_i \times \prod_{i=1}^{M} V_i \to [0,1],$$

and require that it is symmetric, increasing in the number of features that share the same value, equals 1 if and only if the two vectors are identical and equals 0 if and only if no feature is shared by the two vectors.

Second Stage: Top-Down Simplification Decision. After retrieving a non-trivial representation r with $0.5 \leq \hat{\pi}(r) < 1$ the agent may keep it as it is or collapse it into a simplified version that sets $\hat{\pi} = 1$.

Optimal Action Given a Posterior π. For the case of actions (value estimates) let us first characterize the optimal estimate given some fixed representation. Under the binarised scoring rule with prize P the agent's utility is

$$u_m(a) = (1 - (\tfrac{a_m}{100} - \tfrac{a}{100})^2) \times u(P),$$

when the true model is $m \in \{A, B\}$ and a_m is the company value guess provided by model m. Hence the expected utility from guess a is

$$\mathbb{E}_{m\sim\pi}[u_m(a)] = (1 - \frac{1}{10000}\big[\pi\,(a_A - a)^2 + (1-\pi)(a_B - a)^2\big]) \times u(P).$$

Because P and the constant are irrelevant for the maximization, the agent chooses the a that minimizes $\pi(a_A - a)^2 + (1-\pi)(a_B - a)^2$, yielding

$$a^*(\pi) = \pi\, a_A + (1-\pi)\, a_B.$$

Complexity and Cognitive Cost. We examine cognitive costs through the lens of computational complexity, focusing on how the computational demands of a task impose burdens on decision-makers. These burdens then may be reduced via simplification, specifically, by neglecting model uncertainty.

We posit that computational costs vary systematically across decision contexts in a hierarchical fashion, structured by the computational operations each context demands (Table E.1 summarizes our reasoning.)

To illustrate this hierarchy, consider a decision-maker who has already formed beliefs about model uncertainty prior to elicitation (in our experimental design, information about model uncertainty is presented early, before any decisions are elicited). In the high-complexity condition, participants are asked to provide a valuation for a firm. This requires two sequential operations: (i) computing individual value estimates under each candidate model using the provided formulas, and (ii) integrating these estimates into a single value by weighting them according to model uncertainty. In the low-complexity condition, step (i) is bypassed — value estimates are provided directly — and participants need only perform the integration in step (ii).

For belief elicitations, computational demands are negligible, as the relevant beliefs

have already been formed prior to elicitation. For confidence elicitations, we draw on a cognitive science literature on metacognition which characterizes "feeling of knowing" judgments as largely effortless, arising from automatic, non-deliberative processes (Koriat (1993), Koriat (2000), and Koriat and Levy-Sadot (1999)).

Table E.1: Cognitive cost hierarchy across decision contexts

Decision context	**Cognitive cost**	**Required operations**
Valuation (high complexity)	High	(i) Compute model-specific value estimates; (ii) integrate estimates under model uncertainty
Valuation (low complexity)	Low	(i) — (estimates provided); (ii) integrate given estimates under model uncertainty
Belief elicitation	Negligible	Retrieve already-formed beliefs; no computation required
Confidence elicitation	Negligible	Effortless metacognitive access ("feeling of knowing"; Koriat (1993), Koriat (2000), and Koriat and Levy-Sadot (1999))

We hence assume that decisions differ in their level of computational complexity $c \in \{\emptyset, \text{low}, \text{high}\}$. Here, $\emptyset$ means negligible complexity. We set $c = \text{high}$ for company-value guesses in our high-complexity treatment, $c = \text{low}$ for company-value guesses in the low-complexity treatment, and $c = \emptyset$ for the confidence and belief elicitations. Recall that we label the models so that the relevant posterior always satisfies $\pi \geq 0.5$ (i.e. model A is the more likely one), so that simplifying means collapsing to full certainty $\pi = 1$. Accordingly, we specify the cognitive cost of acting on belief π under complexity level c as

$$K(c, \pi) = (1 - \pi)\kappa_c$$

where we assume

$$\kappa_\emptyset = 0 < \kappa_{\text{low}} < \kappa_{\text{high}} \quad \text{and} \quad \kappa_{\text{low}} < \frac{\Delta U(0.65)}{1 - 0.65} < \kappa_{\text{high}},$$

where

$$\Delta U(\pi) = \mathbb{E}_{m \sim \pi}\big[u_m(a^*(\pi))\big] - \mathbb{E}_{m \sim \pi}\big[u_m(a^*(1))\big]$$

is the utility cost of simplifying the representation due to loss of precision in the company value guess if the belief for model A is π. Thus, for company-value guesses, a broad

representation is more costly than a simplified one, with a larger gap under high than low complexity, while for confidence and belief screens with $c = \emptyset$ no cost difference arises.

Utility Comparison for Simplification. Given a stored representation r with implied belief $\hat{\pi}(r)$ and complexity level c, maintaining the broad representation yields

$$\mathbb{E}_{m\sim\hat{\pi}(r)}\big[u_m\big(a^*(\hat{\pi}(r))\big)\big] - K\big(c, \hat{\pi}(r)\big),$$

whereas collapsing to certainty ($\pi = 1$) yields

$$\mathbb{E}_{m\sim\hat{\pi}(r)}\big[u_m\big(a^*(1)\big)\big] - K(c, 1).$$

Hence, the agent simplifies if and only if

$$K(c, \hat{\pi}(r)) - K(c, 1) > \mathbb{E}_{m\sim\hat{\pi}(r)}\big[u_m\big(a^*(\hat{\pi}(r))\big)\big] - \mathbb{E}_{m\sim\hat{\pi}(r)}\big[u_m\big(a^*(1)\big)\big] = \Delta U\big(\hat{\pi}(r)\big),$$

i.e. the cognitive-cost saving $K(c, \hat{\pi}(r)) - K(c, 1)$ exceeds the expected-utility loss $\Delta U\big(\hat{\pi}(r)\big)$. Because $K(\emptyset, \hat{\pi}(r)) - K(\emptyset, 1) = 0$, simplifying on the confidence and belief screens ($c = \emptyset$) would reduce expected utility without any cost-saving. Hence the agent never simplifies under $c = \emptyset$.

Using and Storing Representations. Once a representation has been formed and used to guide a decision, it is added to the database of representations. $V_r^{context}$ of such a representation contains the contextual features of the decision problem in which it was used.

E.3 Timeline of the Experiment

Before stating our predictions, let us briefly recap the timeline of the experiment. Respondents first read the instructions and observe the signal about which model is more likely to be correct. We assume that this induces the default representation

$$r^{\text{initial}} = (0.65, V_{\text{initial}}^{\text{context}}), \qquad \mathbf{R}_0 = \{r^{\text{initial}}\}.$$

We assume that $V^{\text{context}}_{\text{initial}}$ contains features relating to the general setup of the environment, in particular actions and beliefs about model uncertainty.

After going through the instructions, respondents provide the company value guesses, state their decision confidence and then finally their beliefs about which model is correct.

E.4 Predictions

E.4.1 Estimation of Company Values

Predictions. In the high-complexity condition, respondents will simplify the initial representation to $\hat{\pi} = 1$ and choose $a^* = a_A$. In the low-complexity condition, they will keep $\hat{\pi} = 0.65$ and choose $a^* = 0.65\, a_A + 0.35\, a_B$.

Proof. Since $R_0 = \{r^{\text{initial}}\}$, participants must decide whether or not to collapse that single representation. Hence they form

$$r^{\text{guess}} = \begin{cases} (1,\, V^{\text{context}}_{\text{guess}}) & \text{if } K(c, 0.65) - K(c, 1) > \Delta U(0.65), \\ (0.65,\, V^{\text{context}}_{\text{guess}}) & \text{otherwise.} \end{cases}$$

Substituting $K(c, \pi) = (1 - \pi)\kappa_c$ shows that simplification occurs exactly when $(1 - 0.65)\kappa_c > \Delta U(0.65)$. By assumption $(1-0.65)\kappa_{\text{low}} < \Delta U(0.65) < (1-0.65)\kappa_{\text{high}}$, hence only high-complexity agents collapse to $\pi = 1$, yielding the stated actions.

E.4.2 Confidence

Predictions. Assuming that the reported confidence for $\hat{\pi} = 1$ exceeds that for $\hat{\pi} = 0.65$, on average, respondents in the high-complexity condition will report higher confidence than those in the low-complexity condition.

Proof. Confidence is assessed within the formed mental representation. On the confidence screen, the set of available representations is $R_1 = \{r^{\text{initial}}, r^{\text{guess}}\}$. We assume the prompt 'How certain are you that your answer is within 10 percentage points of the best possible guess?' makes the guess context more salient, so

$$S(\xi_{\text{confidence}}, V^{\text{context}}_{\text{guess}}) > S(\xi_{\text{confidence}}, V^{\text{context}}_{\text{initial}}).$$

Therefore the most likely case is that r^{guess} is retrieved. Because $c = \emptyset$ on this screen, no further simplification occurs, so high-complexity subjects retain $\hat{\pi} = 1$ and low-complexity subjects retain $\hat{\pi} = 0.65$. By the assumption that higher stored $\hat{\pi}$ yields higher reported confidence, high-complexity subjects report greater confidence.

If r^{initial} is retrieved instead, once again no simplification occurs because of $c = \emptyset$ and participants in both conditions report a low confidence associated with $\hat{\pi} = 0.65$.

E.4.3 Belief

Predictions. Simplification in the action space might not carry over to the belief space. The most likely outcome is that respondents in both complexity conditions will state their belief as $\hat{\pi} = 0.65$. Respondents in the high-complexity condition may exhibit a greater tendency to state the simplified belief of $\hat{\pi} = 1$.

Proof. On the belief screen, $R_2 = \{r^{\text{initial}}, r^{\text{guess}}, r^{\text{confidence}}\}$. We assume the prompt 'Which of the two rules generated correct company value estimates?' explicitly echoes the original signal description, so

$$S(\xi_{\text{belief}}, V^{\text{context}}_{\text{initial}}) > \max\left\{S(\xi_{\text{belief}}, V^{\text{context}}_{\text{guess}}), S(\xi_{\text{belief}}, V^{\text{context}}_{\text{confidence}})\right\}.$$

Thus the most likely case is that r^{initial} is retrieved. As $c = \emptyset$ here too, no collapse occurs and the stated belief is $\hat{\pi} = 0.65$ in both complexity conditions.

If r^{guess} or $r^{\text{confidence}}$ are retrieved instead, once again no simplification occurs because of $c = \emptyset$. For respondents in the high-complexity condition this implies $\hat{\pi}(r^{\text{guess}}) = \hat{\pi}(r^{\text{confidence}}) = 1$, while under low complexity $\hat{\pi}(r^{\text{guess}}) = \hat{\pi}(r^{\text{confidence}}) = 0.65$.