Pre-Analysis Plan

for *Associative memory, beliefs and investments: A replication and extension*

Companion document to the AEA RCT Registry pre-registration

---

1. Overview

This document contains the pre-analysis plan of *Associative memory, beliefs and investments: A replication and extension*: variable definitions, general estimation conventions, and the estimating equations and tests for each pre-registered hypothesis. Hypothesis numbers refer to the Hypotheses field of the AEA RCT Registry entry for this trial.

2. Variable definitions

Signals.
Signals ("news") are coded +1 (positive) and −1 (negative). For each subject–company pair, s_{k+1} denotes the second-period signal, $k \in \{0, 1, 2, 3, 4\}$ the number of first-period signals, N_p and N_n the number of positive and negative first-period signals, and

$$z \;:=\; \sum_{x=1}^{k} \mathbf{1}\{s_x = s_{k+1}\}$$

the number of first-period signals *congruent* with the second-period signal. Following Enke et al. (2024, §4.4), the signed sum of congruent first-period signals is $N^z := z\, s_{k+1}$ and the signed sum of incongruent first-period signals is $N^u := -(k - z)\, s_{k+1}$; note that $N^z + N^u = N_p - N_n$.

Beliefs.
$b_1, b_2 \in [0, 1]$ denote the stated first and second beliefs that the company is good (numeric input, 0–100%). Beliefs are transformed into *normalized log posterior odds*,

$$lpo_t \;:=\; \frac{\ln\big(b_t/(1 - b_t)\big)}{\ln\big(0.65/0.35\big)},$$

so that the perfect-memory Bayesian benchmark coefficient on each signal equals one (Grether 1980; Enke et al. 2024, Eqs. 2–3).

Boundary beliefs.
Following Enke et al. (2024, footnote 8), beliefs of 0% and 100% (for which log odds are undefined) are recoded to 1% and 99%, respectively, before the transformation. As a robustness check we re-estimate the main specifications dropping these observations instead.

Recall.
For each company, subjects report the number of positive and the number of negative signals they recall having seen, elicited after the second-period signal and before the second belief. Because this report covers the entire experiment including the second-period signal, the *effective*

recall of first-period signals of a given type equals the reported number of that type minus one if the second-period signal is of that type, and the reported number otherwise (Enke et al. 2024, footnote 11).

Investment.
$I \in [0, 1]$ denotes the investment share: the proportion of the per-company budget (1,000 ECU) allocated to the company's stock.

3. General conventions

- All confirmatory analyses are OLS regressions at the subject–company level.
- Standard errors are clustered at the subject level in all regressions.
- Following the estimating equations derived from the model in Enke et al. (2024, Eqs. 6–7), the belief regressions do not include a constant; we also estimate all specifications with a constant as a robustness check.
- $T^{i,f}_{\text{Cue}}$ is a treatment dummy equal to one if company f of subject i is shown in the Cue condition and zero if it is shown in the NoCue baseline; $T^{i,f}_{\text{Sim}}$ is defined the same way for the SimilarCue condition. Because the two conditions of an experiment are manipulated within subject, each dummy varies across the 14 companies of a subject.
- Each directional hypothesis is evaluated using the cluster-robust t-statistic of the relevant coefficient, requiring a point estimate of the hypothesized sign with two-sided $p < 0.05$.
- Following the original study, we do not adjust for multiple hypothesis testing; each hypothesis is evaluated separately.

4. Estimating equations and tests

Each hypothesis is a coefficient test in one of the OLS regressions below, all at the subject–company level with subject-clustered standard errors (see the general conventions above). The dependent variable is the normalized log posterior odds of the second belief (lpo_2) for the belief hypotheses, effective recall (ER) for the recall hypotheses, and the investment share (I) for the investment hypotheses. Each specification is described below with its economic interpretation.

Role of associative memory (original Result 1).
(*Pre-registered as Hypothesis 1 of the AEA RCT Registry entry for this trial.*) The interaction $s_{k+1} \times T_{\text{Cue}}$ isolates the causal effect of the cueing context: the *extra* weight the second-period signal receives when its repeated context triggers associative recall of congruent past signals. Cue and NoCue observations of the Beliefs Cue+NoCue experiment (Enke et al. 2024, Eq. 6):

$$lpo_2^{i,f} = \beta_1 s_{k+1}^{i,f} + \beta_2\, s_{k+1}^{i,f}\, T_{\text{Cue}}^{i,f} + \beta_3 T_{\text{Cue}}^{i,f} + \beta_4 \big(N_p^{i,f} - N_n^{i,f}\big) + \varepsilon^{i,f}. \tag{1}$$

Test: $\beta_2 > 0$ (associative memory amplifies the response to the second-period signal).

Variation in the number of cued signals.
(*Pre-registered as Hypothesis 2 of the AEA RCT Registry entry for this trial.*) Within Cue, this tests history-dependence (Enke et al. 2024, Result 2): the more Part 1 signals share the second-period signal's context (z), the more past signals are cued, and the stronger the overreaction should be. Cue observations of the Beliefs Cue+NoCue experiment only (Enke et al. 2024, Eq. 7):

$$lpo_2^{i,f} = \beta_5 s_{k+1}^{i,f} + \beta_6\, s_{k+1}^{i,f}\, z^{i,f} + \beta_7 z^{i,f} + \beta_8 \big(N_p^{i,f} - N_n^{i,f}\big) + \varepsilon^{i,f}. \tag{2}$$

Test: $\beta_6 > 0$ (overreaction increases in the number of cued signals).

Weight on cued versus non-cued signals (evidence for original Result 3). (*Pre-registered as Hypothesis 3 of the AEA RCT Registry entry for this trial.*) This decomposes the Part 1 signals into congruent (cued) and incongruent (non-cued) signed sums and asks through which channel the Cue effect operates: associative memory should reweight congruent signals only. It is the beliefs-side evidence for the original's Result 3 (Enke et al. 2024, Table 1, cols. (6)–(7)); the recall-side evidence is the asymmetric-recall specification below. Cue and NoCue observations of the Beliefs Cue+NoCue experiment:

$$lpo_2^{i,f} = \beta_1 s_{k+1}^{i,f} + \beta_2 T_{\text{Cue}}^{i,f} + \beta_3 N^{z,i,f} + \beta_4 N^{u,i,f} + \beta_5 \, N^{z,i,f} \, T_{\text{Cue}}^{i,f} + \beta_6 \, N^{u,i,f} \, T_{\text{Cue}}^{i,f} + \varepsilon^{i,f}. \quad (3)$$

Tests: $\beta_5 > 0$ (Cue raises the weight on congruent signals); β_6 not statistically different from zero (Cue does not change the weight on incongruent signals).

SimilarCue also triggers associative memory (beliefs). (*Pre-registered as Hypothesis 4 of the AEA RCT Registry entry for this trial.*) This repeats the belief tests on the new SimilarCue condition, asking whether *theme-level* similarity (same theme, different instance) is sufficient to produce the amplification. The three belief specifications above are estimated unchanged on the Beliefs SimilarCue+NoCue experiment, with $T_{\text{Sim}}^{i,f}$ in place of $T_{\text{Cue}}^{i,f}$ and NoCue as the within-subject baseline, and tested with the same signs; the headline test is $\beta_2 > 0$ in the associative-memory specification.

Exploratory comparison. To compare SimilarCue against Cue, stack the Beliefs Cue+NoCue and Beliefs SimilarCue+NoCue experiments (subjects are distinct across experiments). Let D_{Sim}^i be an experiment dummy, equal to one for subjects run in the SimilarCue+NoCue experiment and zero for subjects run in the Cue+NoCue experiment; it is an experiment fixed effect and absorbs any level difference between the two samples. The treatment dummies $T_{\text{Cue}}^{i,f}$ and $T_{\text{Sim}}^{i,f}$ are as defined above, so that a company is either in Cue, in SimilarCue, or in the NoCue baseline of its own experiment:

$$\begin{aligned} lpo_2^{i,f} = {} & \beta_1 s_{k+1}^{i,f} + \gamma_1 \, s_{k+1}^{i,f} \, T_{\text{Cue}}^{i,f} + \gamma_2 \, s_{k+1}^{i,f} \, T_{\text{Sim}}^{i,f} + \delta_1 T_{\text{Cue}}^{i,f} + \delta_2 T_{\text{Sim}}^{i,f} \\ & + \eta \, D_{\text{Sim}}^i + \kappa \, s_{k+1}^{i,f} \, D_{\text{Sim}}^i \\ & + \beta_4 (N_p^{i,f} - N_n^{i,f}) + \beta_5 (N_p^{i,f} - N_n^{i,f}) D_{\text{Sim}}^i + \varepsilon^{i,f}, \end{aligned} \quad (4)$$

where the NoCue companies of both experiments form the common reference category. The experiment dummy is interacted with every regressor, so the specification is algebraically equivalent to estimating each experiment separately: γ_1 and γ_2 are exactly the treatment interactions of their own experiment, each measured against that experiment's own NoCue arm, and $\gamma_1 - \gamma_2$ is a difference-in-differences. Pooling is only a device for obtaining that difference and its standard error in one step; the coefficients coincide with the separate estimates and the clustered standard errors do so up to the finite-sample correction. Without the interactions, a common baseline would be imposed on the two samples and any difference in it would load onto γ_1 and γ_2. We report the difference $\gamma_1 - \gamma_2$ (Cue versus SimilarCue) descriptively, without a pre-registered direction; because Cue and SimilarCue are administered to different samples, this comparison is across-subject.

Asymmetric recall of cued signals. (*Pre-registered as Hypothesis 5 of the AEA RCT Registry entry for this trial.*) This tests the memory mechanism directly on what subjects recall (Enke et al. 2024, Result 3): whether the cueing context raises recall of *congruent* Part 1 signals (those sharing the second-period signal's context) over and above recall of *incongruent* ones. Each subject–company pair contributes two observations, one for the congruent first-period signals and one for the incongruent ones; index them by $j \in \{c, u\}$. Let $ER_j^{i,f}$ denote effective recall in observation j, let $C_j^{i,f}$ be a congruent-signal dummy that equals one in the congruent observation and zero in the incongruent one, and let $k_j^{i,f}$ be the true number of signals of that type ($k_c^{i,f} = z$, $k_u^{i,f} = k - z$); observations with $k_j^{i,f} = 0$ are dropped. Pooling both types (the specification of Enke et al. 2024, Online Appendix Table C5, col. (3)):

$$ER_j^{i,f} = \gamma_1 T_{\text{Cue}}^{i,f} + \gamma_2\, C_j^{i,f} + \gamma_3\, T_{\text{Cue}}^{i,f} C_j^{i,f} + \mu_{k_j^{i,f}} + \varepsilon_j^{i,f}, \tag{5}$$

where μ_k are fixed effects for the true number of signals of that type. Test: $\gamma_3 > 0$ (the cue boosts recall of congruent signals differentially). As supplementary descriptive evidence we report the cue effect separately for congruent and incongruent signals (Table C5, cols. (1)–(2)).

Role of associative memory in investment (exploratory).

(*Pre-registered as Hypothesis 7 of the AEA RCT Registry entry for this trial.*) This is the associative-memory test with the investment share as the outcome: whether the Cue amplifies how strongly *investment* responds to the second-period signal, so that the memory bias shows up in actual allocations and not only in stated beliefs. Cue and NoCue observations of the Invest Cue+NoCue experiment:

$$I^{i,f} = \pi_1 s_{k+1}^{i,f} + \pi_2\, s_{k+1}^{i,f}\, T_{\text{Cue}}^{i,f} + \pi_3 T_{\text{Cue}}^{i,f} + \pi_4 (N_p^{i,f} - N_n^{i,f}) + \pi_0 + \varepsilon^{i,f}. \tag{6}$$

Test: $\pi_2 > 0$. (Unlike the belief regressions, this specification includes a constant, since the investment share is not derived from the log-odds transformation.)

Belief–investment consistency (exploratory).

(*Pre-registered as Hypothesis 8 of the AEA RCT Registry entry for this trial.*) This links the two responses elicited for the same company, asking whether subjects who state a higher probability that the company is good also allocate more of the budget to its stock:

$$I^{i,f} = \pi_0 + \pi_1 b_2^{i,f} + \varepsilon^{i,f}. \tag{7}$$

Test: $\pi_1 > 0$. The magnitude is interpreted descriptively: $\pi_1 = 1$ corresponds to the proportional-investment benchmark $I = b_2$; a risk-neutral investor follows a step function in b_2 (generating π_1 well above one), while sufficiently risk-averse investors exhibit attenuated responses. We pre-register an order-effect test: estimate $I^{i,f} = \pi_0 + \pi_1 b_2^{i,f} + \pi_2 O^i + \pi_3\, b_2^{i,f} O^i + \varepsilon^{i,f}$, with $O^i = 1$ if the belief was elicited first, and test $\pi_3 = 0$; the result is robust only if $\pi_1 > 0$ under both orders.

SimilarCue also triggers asymmetric recall.

(*Pre-registered as Hypothesis 6 of the AEA RCT Registry entry for this trial.*) This repeats the asymmetric-recall test on the SimilarCue condition, asking whether theme-level similarity is by itself enough to boost recall of congruent signals. *Confirmatory test.* Estimate the pooled effective-recall specification above on the Recall SimilarCue+NoCue experiment alone, with $T_{\text{Sim}}^{i,f}$ in place of $T_{\text{Cue}}^{i,f}$ and NoCue as the within-subject baseline:

$$ER_j^{i,f} = \gamma_1 T_{\text{Sim}}^{i,f} + \gamma_2\, C_j^{i,f} + \gamma_3\, T_{\text{Sim}}^{i,f} C_j^{i,f} + \mu_{k_j^{i,f}} + \varepsilon_j^{i,f}. \tag{8}$$

Test: $\gamma_3 > 0$ (theme-level similarity suffices to produce asymmetric recall of congruent signals).

Exploratory comparison. To compare SimilarCue against Cue, stack the Recall Cue+NoCue and Recall SimilarCue+NoCue experiments (subjects are distinct across experiments). Using the pooled specification, with the treatment dummies $T_{\text{Cue}}^{i,f}$ and $T_{\text{Sim}}^{i,f}$ and the experiment dummy D_{Sim}^i defined as above:

$$\begin{aligned} ER_j^{i,f} = {} & \gamma_2\, C_j^{i,f} + \psi_1\, T_{\text{Cue}}^{i,f} C_j^{i,f} + \psi_2\, T_{\text{Sim}}^{i,f} C_j^{i,f} + \delta_1 T_{\text{Cue}}^{i,f} + \delta_2 T_{\text{Sim}}^{i,f} \\ & + \eta\, D_{\text{Sim}}^i + \kappa\, C_j^{i,f}\, D_{\text{Sim}}^i + \mu_{k_j^{i,f}, D_{\text{Sim}}^i} + \varepsilon_j^{i,f}, \end{aligned} \tag{9}$$

where the NoCue companies of both experiments form the common reference category. Here $\mu_{k,D}$ are fixed effects for the true number of signals of that type, estimated separately by experiment. As in the belief comparison, the experiment dummy is interacted with every regressor, so the specification is equivalent to estimating each experiment separately and $\psi_1 - \psi_2$ is a difference-in-differences. We report this difference (Cue versus SimilarCue) descriptively, without a pre-registered direction; as in the belief case the comparison is across-subject.

References

- Enke, B., Schwerter, F., & Zimmermann, F. (2024). *Associative memory, beliefs and market interactions*. Journal of Financial Economics, 157, 103853. doi:10.1016/j.jfineco.2024.103853.
- Grether, D. M. (1980). *Bayes rule as a descriptive model: The representativeness heuristic*. Quarterly Journal of Economics, 95(3), 537–557.