Behavioral Economics

Lecture 8: Overconfidence and Crowdsourcing

Daniel Martin

danielmartin@ucsb.edu

What Epping et al. Estimate

$$\operatorname{logit} r_i = \alpha_i \operatorname{logit} p_i + \beta_i$$

Data engineering move

- Each annotator reports a probability $p_i(x)$ that an image is a blast cell
- A small calibration set has known labels $y(x)$
- Fit a logistic regression of truth on the annotator's log odds
- Use the fitted map to produce recalibrated elicited beliefs $r_i(x)$

Parameter interpretation

- α_i: spread correction
- $|\alpha_i| < 1$ compresses overconfident reports; $|\alpha_i| > 1$ amplifies underconfident reports
- β_i: class/response bias correction
- $\alpha_i < 0$: flip anti-informative annotators

The One-Line Connection

The recalibration rule in Epping, Caplin, Duhaime, Holmes, Martin & Trueblood (2026) is the **binary-state version** of the theory paper's **power-weighted distortion** in Chambers, Masatlioglu & Raymond (2023)

Epping et al. recalibration rule

$$\operatorname{logit} r_i(x) = \alpha_i \operatorname{logit} p_i(x) + \beta_i$$

Power-weighted belief distortion

$$\phi(p)(\omega) = \frac{\psi(\omega)p(\omega)^\gamma}{\sum_{\omega'} \psi(\omega')p(\omega')^\gamma}$$

Epping et al. LLO recalibration equations, CMR Theorem 1

Binary Algebra: LLO Is Power Weighting

Let $\Omega = \{1, 0\}$ and write $p = p(1)$, looking at state 1:

$$\phi(p)(1) = \frac{\psi_1 p^\gamma}{\psi_1 p^\gamma + \psi_0 (1-p)^\gamma}$$

$$\frac{\phi(p)(1)}{1-\phi(p)(1)} = \frac{\psi_1}{\psi_0}\left(\frac{p}{1-p}\right)^\gamma$$

$$\operatorname{logit}\phi(p)(1) = \gamma \operatorname{logit} p + \log\left(\frac{\psi_1}{\psi_0}\right)$$

The LLO rule has the same form, with $\alpha = \gamma$ and $\beta = \log(\psi_1/\psi_0)$

What Coherence Requires

A distorted belief is a map

$$\phi : \Delta(\Omega) \to \Delta(\Omega)$$

Coherence:

$$\phi(p(\cdot \mid E)) = \phi(p)(\cdot \mid E)$$

whenever $p(E) > 0$

The timing of distortion and conditioning should not matter

CMR Definition 1 and Theorem 1

$$\begin{tikzcd}
\boxed{p} \arrow[r, "\text{distort}"] \arrow[d, "\text{condition}"'] & \boxed{\phi(p)} \arrow[d, "\text{condition}"] \\
\boxed{p(\cdot|E)} \arrow[r, "\text{distort}"'] & \boxed{\phi(p)(\cdot|E)}
\end{tikzcd}$$

CMR show: positivity + continuity + coherence $\iff$ power-weighted distortion

Connection to Grether

- CMR consider a generalized class of distortions where the DM can distort, independently, both their prior and the Blackwell experiment (belief updating)
- The coherency conditions imply that two distinct operations must be equivalent: (i) distorting the prior, distorting the signal, and then Bayesian updating; and (ii) Bayesian updating with true probabilities, and then distorting the posterior
- CMR find that coherence requires:
    1. Distortions to priors must take on the familiar power-weighted form
    2. Distortions to signals take on a non-normalized power weighted form
    3. The powers of the two distortions are the same
- The Grether model is coherent only when the prior and signal powers are equal!

References I

Chambers, C. P., Masatlioglu, Y. & Raymond, C. (2023), 'Coherent distorted beliefs', *arXiv preprint arXiv:2310.09879* .

Epping, G. P., Caplin, A., Duhaime, E., Holmes, W. R., Martin, D. & Trueblood, J. S. (2026), 'Improving crowdsourcing for ai through cognitive-inspired data engineering', *Preprint, submitted December* **30**.