Online Companion.

This electronic companion provides full mathematical details of the three examples of linear information aggregation problems listed in Section 3.3, as well as a discussion of additional findings from a different viewpoint of participant responses in Study 1.

Examples of Linear Information Aggregation Problems

Example 1: A Normal Random Variable. This frequently used model assumes that X is distributed normally with unknown mean θ and known precision λ, $(X|\theta) \sim \mathrm{N}(\theta, \lambda)$. All judges share a normally distributed common prior with mean μ_0 and precision $m_0\lambda$, $\pi_0(\theta) \sim \mathrm{N}(\mu_0, m_0\lambda)$, and observe the same shared signal s_1, where $(s_1|\theta) \sim \mathrm{N}(\theta, m_1\lambda)$. Private signals are distributed according to $(t_i|\theta) \sim \mathrm{N}(\theta, \ell\lambda)$. The posterior distribution for θ, given π_0 and a collection of signals $\{s_1, t_1, \ldots, t_n\}$, is $(\theta|\pi_0, s_1, \{t_i\}_{i=1}^n) \sim \mathrm{N}(\frac{ms+\ell\sum_{i=1}^n t_i}{m+n\ell}, (m+n\ell)\lambda)$ and the corresponding posterior predictive distribution for X is $(X|\pi_0, s_1, \{t_i\}_{i=1}^n) \sim \mathrm{N}(\frac{ms+\ell\sum_{i=1}^n t_i}{m+n\ell}, \frac{m+n\ell}{m+n\ell+1}\lambda)$.

Example 2: A Binomial Random Variable. As another example, suppose that $X \in \{0, 1, \ldots, N\}$ is a Binomial random variable, equalling the sum of N independent Bernoulli trials $X_j \in \{0, 1\}$. The decision analyst would like to estimate the probability $\mathbb{P}(X_j = 1) = \theta$ of a success on each trial $j = 1, \ldots, N$ (note that forecasting the probability of a single event is the special case of $N = 1$ here). Judges share a common prior over θ which follows a Beta distribution, $\pi_0(\theta) \sim \mathrm{Beta}(m_0\mu_0, m_0 - m_0\mu_0)$. The shared signal s_1 is the average of m_1 independent realizations X_k from the Bernoulli process, $s_1 = \frac{1}{m_1}\sum_{k=1}^{m_1} X_k$. Private signals t_i each equal the average of ℓ independent realizations X_{ik} from the Bernoulli process, $t_i = \frac{1}{\ell}\sum_{k=1}^{\ell} X_{ik}$. The posterior distribution for θ, given π_0 and a collection of signals $\{s_1, t_1, \ldots, t_n\}$, is $(\theta \mid \pi_0, s_1, \{t_i\}_{i=1}^n) \sim \mathrm{Beta}(ms + \ell\sum_{i=1}^n t_i, m(1-s) + \ell\sum_{i=1}^n (1-t_i))$, and the posterior predictive distribution for X is Beta-Binomial,[5] $(X|\pi_0, s_1, \{t_i\}_{i=1}^n) \sim Bb(ms + \ell\sum_{i=1}^n t_i, m(1-s) + \ell\sum_{i=1}^n (1-t_i), N)$, with expectation $\frac{ms+\ell\sum_{i=1}^n t_i}{m+n\ell}N$.

Example 3: A Gamma Random Variable. We can also consider a family of continuous random variables with skewness using the following gamma information structure. Suppose the decision analyst would like to estimate the mean of a positive random variable $X > 0$, which follows a Gamma distribution with known shape parameter $\alpha > 0$ and unknown rate parameter $\alpha/\theta > 0$, $X \sim Ga(\alpha, \alpha/\theta)$.[6] Judges share a common Inverted Gamma prior over θ with shape

[5] A random variable Z follows a Beta-Binomial distribution, $Z \sim Bb(\alpha, \beta, N)$, if it has the probability mass function $p(z; \alpha, \beta, N) = \binom{N}{z}\frac{\Gamma(\alpha+\beta)\Gamma(\alpha+z)\Gamma(\beta+N-z)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\alpha+\beta+N)}$ for $z \in \{0, 1, \ldots, N\}$ (Bernardo & Smith, 1994, p. 428).

[6] A positive random variable Z follows a Gamma distribution with shape $\alpha > 0$ and rate $\beta > 0$, $Z \sim Ga(\alpha, \beta)$, if it has the probability density function $f(z; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} z^{(\alpha-1)} \exp(-\beta z)$ over $z > 0$, with mean $\frac{\alpha}{\beta}$. A positive random variable Z follows an Inverted Gamma distribution with shape α and rate β, $Z \sim Ig(\alpha, \beta)$, if it has the probability density function $f(z; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} z^{-(\alpha+1)} \exp(-\beta/z)$ over $z > 0$, with mean $\frac{\beta}{\alpha-1}$ (Bernardo & Smith, 1994, p. 430-431).

$m_0\alpha+1$ and rate $m_0\alpha\mu_0$, $\pi_0(\theta)\sim Ig(m_0\alpha+1, m_0\alpha\mu_0)$. All judges observe the same shared signal s_1, which is distributed according to a Gamma distribution with shape $m_1\alpha$ and rate $m_1\alpha/\theta$, $(s_1 \mid \theta)\sim Ga(m_1\alpha, m_1\alpha/\theta)$. Private signals t_i follow a Gamma distribution with shape $\ell\alpha$ and rate $\ell\alpha/\theta$, $(t_i \mid \theta)\sim Ga(\ell\alpha, \ell\alpha/\theta)$. In this case, the posterior distribution for θ, given π_0 and a collection of signals $\{s_1, t_1, \dots, t_n\}$, is $(\theta \mid \pi_0, s_1, \{t_i\}_{i=1}^n)\sim Ig((m+n\ell)\alpha+1, (ms+\ell\sum_{i=1}^n t_i)\alpha)$, and the corresponding posterior expectation of X is $\frac{ms+\ell\sum_{i=1}^n t_i}{m+n\ell}$.

Additional Analyses of Participants' Responses in Study 1

Since the shared and private information provided to participants in Study 1 is controlled, we can examine how individuals use their information to respond and how this compares to the optimal strategy. In the main manuscript, we use this information to infer the relative weights applied to private and shared signals, since these implied weights convert responses from many different coins to a standardized weighting scale between 0 and 1. However, we can also consider the differences between actual and optimal responses provided by participants.[7] This allows us to directly study the patterns of behavior across a spectrum of different combinations of coin biases and signals that individuals could encounter in this setting.

Scatterplots of the average differences between the actual and optimal Bayesian responses across participants for each of the 72 coins in Study 1 for forecasts and guesses of others, respectively, are displayed in Figure EC.1. Both $\bar{f}$ and $\bar{g}$ tended to regress toward 50, a pattern which we believe may result from external priors that some participants brought with them to the experiment centered on the belief that coins are fair (i.e., have an equal chance of landing on heads or tails).

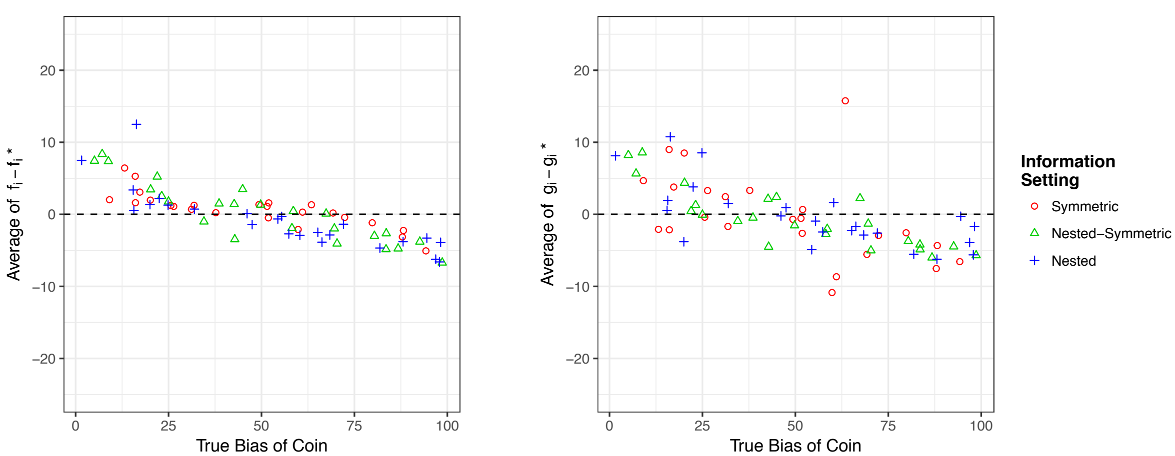

Figure EC.1 Average difference between actual and optimal for forecasts and guesses of others versus true bias.

[7] We are grateful to an anonymous reviewer for suggesting this analysis.

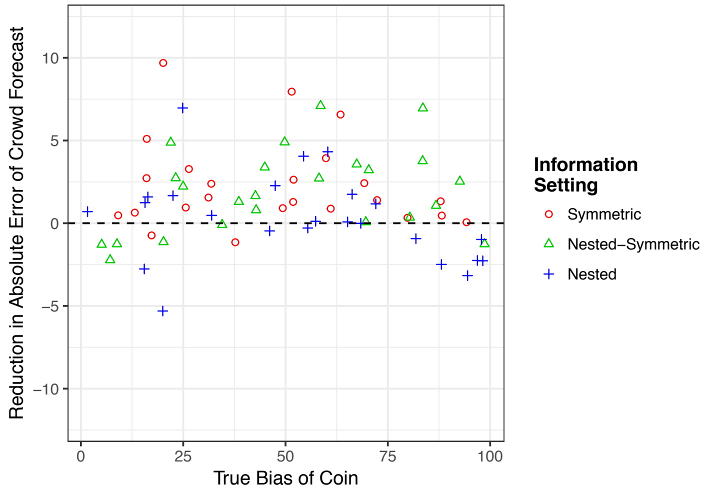

Figure EC.2 **Reduction $|\bar{f}-\theta| - |\hat{\theta}_M - \theta|$ in the absolute forecast error of the minimal pivoted forecast relative to the simple average forecast versus true bias.**

As a result, minimal pivoting tended to be most effective when the true bias of the coin was close to 0.5 and less effective when the bias of the coin was close to 0 or 1, as shown in Figure EC.2. This pattern contrasts with the predictions of extremizing procedures (Baron et al., 2014), which provide the greatest improvement when the bias of the coin is close to 0 or 1.

References

[1] Baron, J., Mellers, B. A., Tetlock, P. E., Stone, E., Ungar, L. H. (2014). Two Reasons to Make Aggregated Probability Forecasts More Extreme. *Decision Analysis* **11**(2) 133–145.

[2] Bernardo, J. M. Smith, A. F. M. (1994). *Bayesian Theory.* John Wiley & Sons, Ltd.