



**S4 Fig. Fits of different aggregation rules to the observed data at various levels of added noise in Experiment 1.** (A) Log-likelihood of the observed consensus values being produced from the exact value of each rule tested for a range of noise values (curves). Highlighted are the log-likelihood values corresponding to the observed level of deviation around the rule (circles). (B) As A, but only for groups with a low range ( $\leq 40$ ) of estimates. (C) As A, but

only for groups with a high range ( $>40$ ) of estimates. The rules tested are: the lowest estimate ( $x_1$ ), mean of the lowest and median estimate ( $x_{12}$ ), the median ( $x_2$ ), geometric (geom) and arithmetic (arit) means, mean of the lowest and highest estimates in each group ( $x_{13}$ ), mean of median and highest estimate ( $x_{23}$ ), and highest estimate ( $x_3$ ). At each level of noise, the geometric mean consistently provides amongst the best fits to the observed data.