Supplementary Materials

Detail on backwards induction.

The action value associated with continued sampling (and, therefore, effectively the decision threshold for each sequence position) was computed based on a backwards induction algorithm. Here, we give an intuitive description of this algorithm, using the simplest incentivisation scheme, where the reward value of each option equals its option value (as in the Gilbert & Mosteller formulation as well as the “full” conditions in our current studies). Backwards induction involves begins by considering the final option. When the agent encounters the final option, the decision threshold T1 must be zero, as the paradigm forces choice of this option and there are no future option values to consider. Before that, when the agent encounters the second to last option, the computation of the threshold T2 for this option remains relatively simple: the agent must compare the value of the second to last option against the *expected value* of the final option (i.e., the mean of generating distribution of option values). For the third to last option, its decision threshold T3 depends on the expected value of continuing to the second to last option, with the opportunity to also continue to the final option. Computation of T3 then involves integrating the density function of the generating distribution of option values to compute the expected value of options above the decision threshold T2 and then adding the expected reward for the final position. This process can continue backwards to obtain the expected value of getting any option in the future that is better than the current one, wherever in the sequence the current one is.Parameter recovery

To ascertain the ability of our models to derive the correct parameters from individual participant data, we performed parameter recovery analyses, in which we simulated model choices (take option or sample again) in response to randomly generated option values. We simulated models with 12 possible levels of each parameter, within which there were 25 simulated participants with five sequences per participant and 12 options per sequence. To parallel the structure of our empirical paradigms, we created a generating distribution (separately for each simulated participant) of 90 option values, randomly produced from a Gaussian distribution with mean 40 and standard deviation of 20 and within the range of 1 to 100 (recall that we normalised all our prices to this same range when fitting models to human participants). Then, we populated the sequences of input option values for the optimal stopping task from this participant-specific generating distribution. We configured our models with ranges of the key theoretical parameters (Figure S1, x axis) that produced sampling rates between roughly two and ten samples to decision (Figures S2 and S3). The aforementioned randomly generated option values were then presented to every configured model to extract simulated sampling rates associated with each configured parameter value. Varying the configured parameters in this way led to systematic variation in the sampling rate, as expected (Fig S3, top panel). We then fitted the models to these simulated take option / sample again decisions in the same way as we fitted human participants to obtain parameter estimates of the configured parameters. For the three models we cover in the main text, CO, CS and Biased Prior models, configured and estimated parameters tended to correlate (Figure S1 and lower panel of Figure S3). Also, the sampling rates simulated using configured parameters highly correlated with sampling rates simulated using the estimated parameters (Figure S2 and middle panel of Figure S3).

We formulated two more theoretical models, but these did not perform so well during parameter recovery and so were excluded from model fitting and comparison. In the *biased values model (BV)*, we considered the possibility that, although participants may use the optimal solution to solve the task, they might instead be biased to misperceive the magnitudes of the option values that are input into this optimal solution. This might especially be the case if participants perceive only the very most valued options as worthy of consideration at all, as might be the case in “high threshold” models of optimal stopping in mate choice (Furl et al., 2019; Valone et al., 1996). We incorporated a threshold into the BV model: Option values above this option value threshold are transformed to 100 (the maximal option value) and option values below this threshold are transformed to 1 (the minimal option value). The transformed option values, once rendered attractive only when above threshold, are then submitted to the Ideal Observer. We picked the centre of the input value range (i.e., 50) as the starting value for the threshold parameter when fitting to participants’ choices and bounded the parameter fitting to be within the option value range 1 to 100.

The *biased rewards model* *(BR)* is based on similar logic as BV. However, instead of assuming participants place a threshold on the option values being input to the model, we instead assumed such a threshold on the reward function *R* (See formula for above). Recall that this function assigns reward payoffs the relative ranks of choices. As with BV, we adopted a threshold, which was the free parameter, above which choices received a maximal reward of 100 and below which choices received a reward of 1. As in the BV model, the starting value of this threshold parameter was initialised at 50 and fitting was bounded between 1 and 100.

Text B: Attention check

Attention checks were added to phase one (i.e., the ratings phase) of Pilot full and the Study 1 full and Study 1 ratings conditions, to compensate for the unsupervised nature of online data collection. Every attention check showed a cross, a ‘next’ button, and the text "press ‘next’ when the cross disappears". The cross disappeared at a random time interval between one and five seconds. The ‘next’ button was active the whole time. If participants were paying attention, they would not press the ‘next’ button as soon as it appeared, but would instead read the text and respond only after the cross had disappeared. Thus, if participants’ response time exceeded the cross display time, they passed the attention check. Nevertheless, we found high correlations between phase 1 and phase 2 ratings (i.e., > .8) across our studies (see average correlations between phase1 ratings reported in Methods to pilot studies and Study 1) and so we elected to not remove participants based on attention check data in Pilot full, Study 1 full and Study 1 ratings and we discontinued the use of attention check trials in Studies 2 and 3.

Figure S1. For parameter recovery analysis, we examined parameter values (vertical axes) estimated from fitting models to decisions simulated using configured parameter values (horizontal axes) for simulated participants, each shown as an individual scatter point. The grey diagonal indicates when configured and estimated parameters would be exactly equal. The coloured line indicates the regression line relating configured and estimated parameter values. BR and BV models showed too poor parameter recovery to be entered into formal model comparison. Abbreviations: CS = cost to sample, CO = cut-off, BP = biased prior, BV = biased values, BR = biased reward.

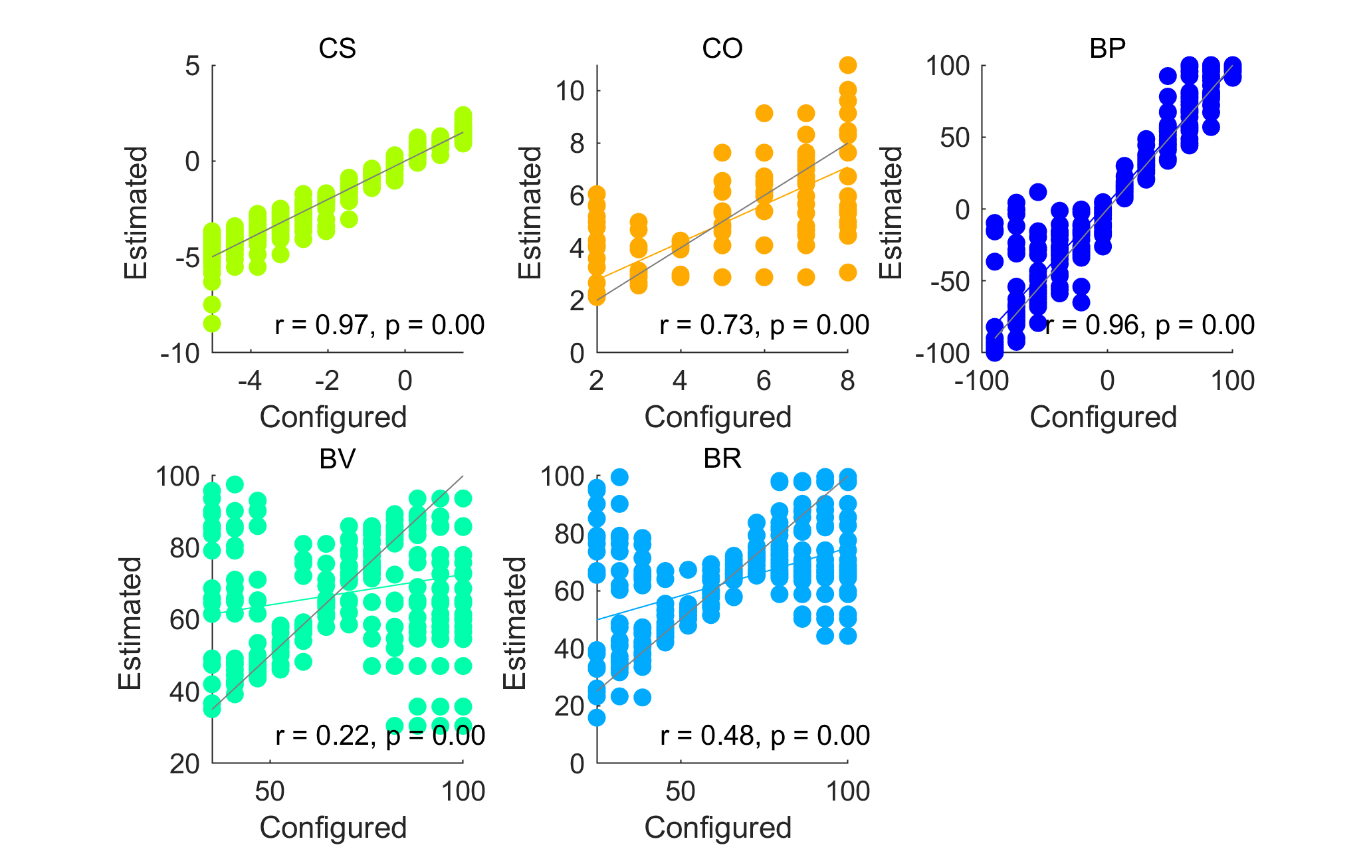


Figure S2. Top panel: We plot sampling rates for individual simulated participants (points) and their mean values (bars) for each configured parameter level in the parameter recovery analysis. Systematically varying configured parameter values successfully increases or decreases simulated sampling rates for Cost to Sample and Biased Prior models and the Cut Off heuristic. Middle panel: Models were fitted to the data in the top panel and parameters estimated. We plot the sampling rates simulated using each estimated parameter (points) and their mean sampling rates (bars). Lower panel: The estimated parameters (points) are plotted relative to their target configured parameter values (bars). Each model’s parameter values are normalised to a 0 to 1 range to facilitate plotting on one scale. Abbreviations: CS = cost to sample, CO = cut-off, BP = biased prior, BV = biased values, BR = biased reward.

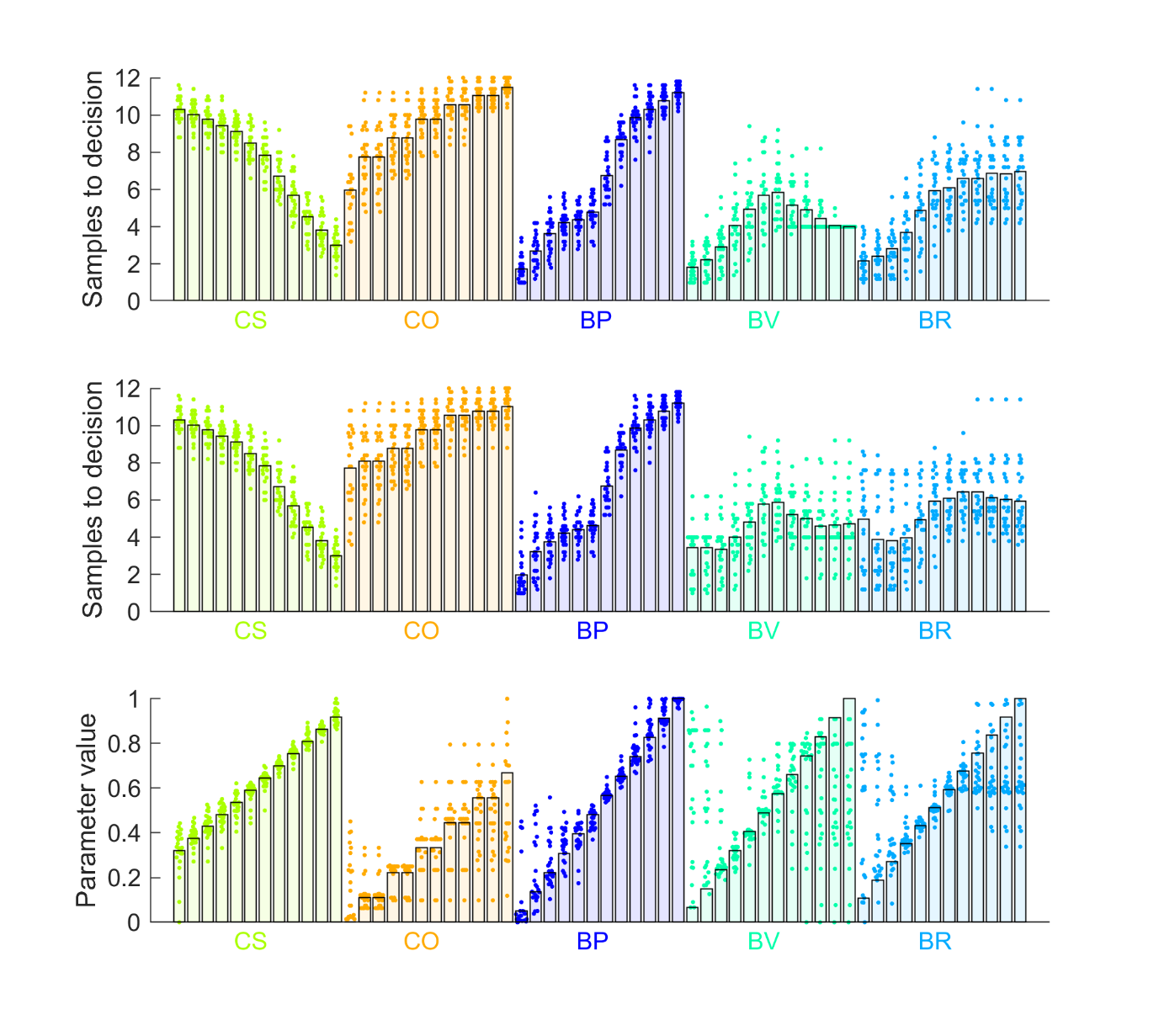


Figure S3. Sampling rates simulated using configured parameters (horizontal axis) are plotted against sampling rates computed from estimated parameters. The grey diagonal indicates when sampling rates based on configured and estimated parameters would be exactly equal. The coloured line indicates the regression line relating sampling rates based on configured and estimated parameter values. Abbreviations: CS = cost to sample, CO = cut-off, BP = biased prior, BV = biased values, BR = biased reward.

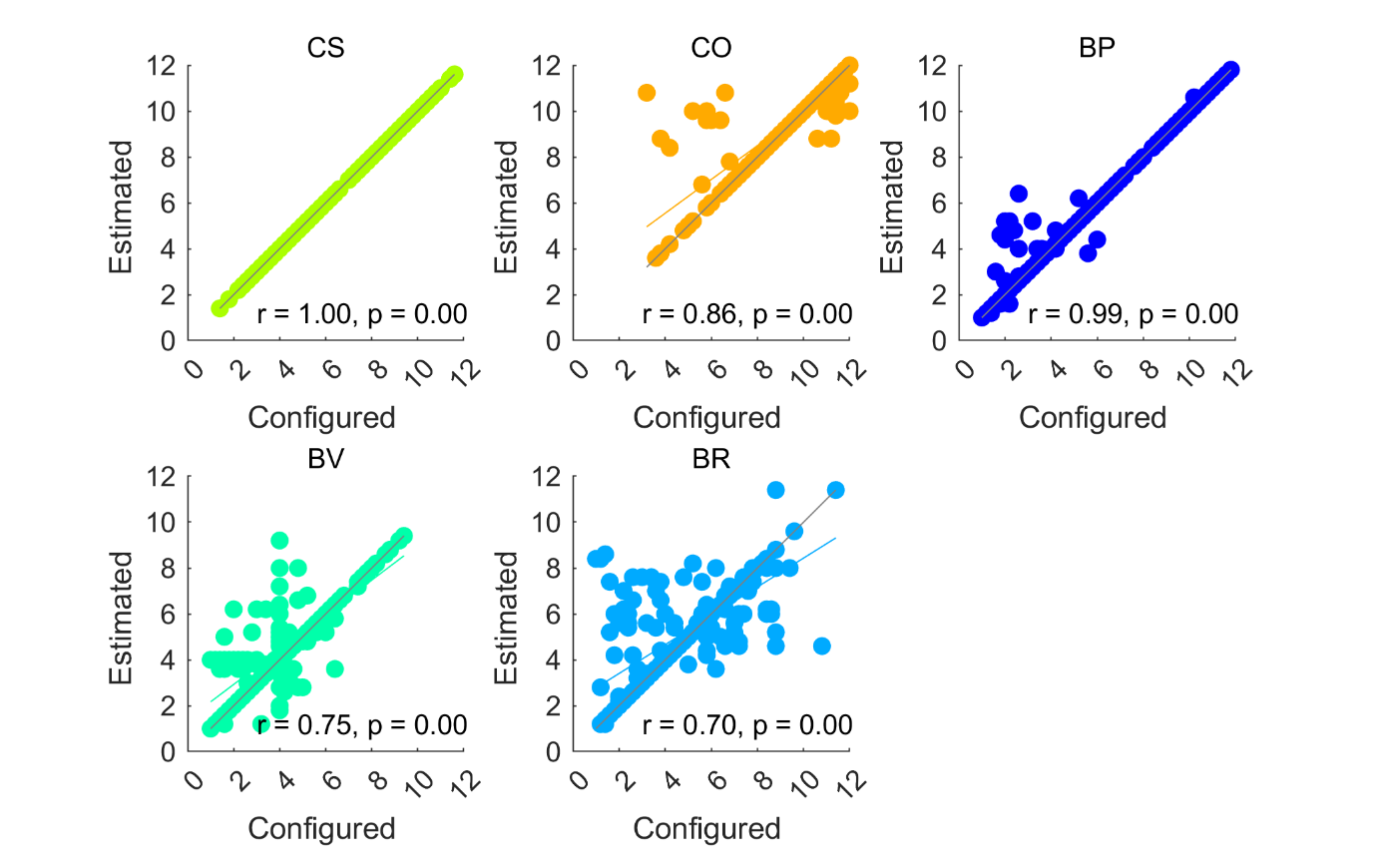


Figure S4. Model comparison for Pilot baseline (left column) and Pilot full (right column). Top and middle rows show individual participants as points and bars show their mean values. The top row shows the ranks of chosen prices. The second row plots the “first” or key theoretical parameter values, estimated for each fitted model. The third row shows the “second,” or inverse temperature parameter beta, estimated for each fitted model. Abbreviations: IO = ideal observer, CS = cost to sample, CO = cut-off, BP = biased prior, OV = objective values, SV = subjective values.

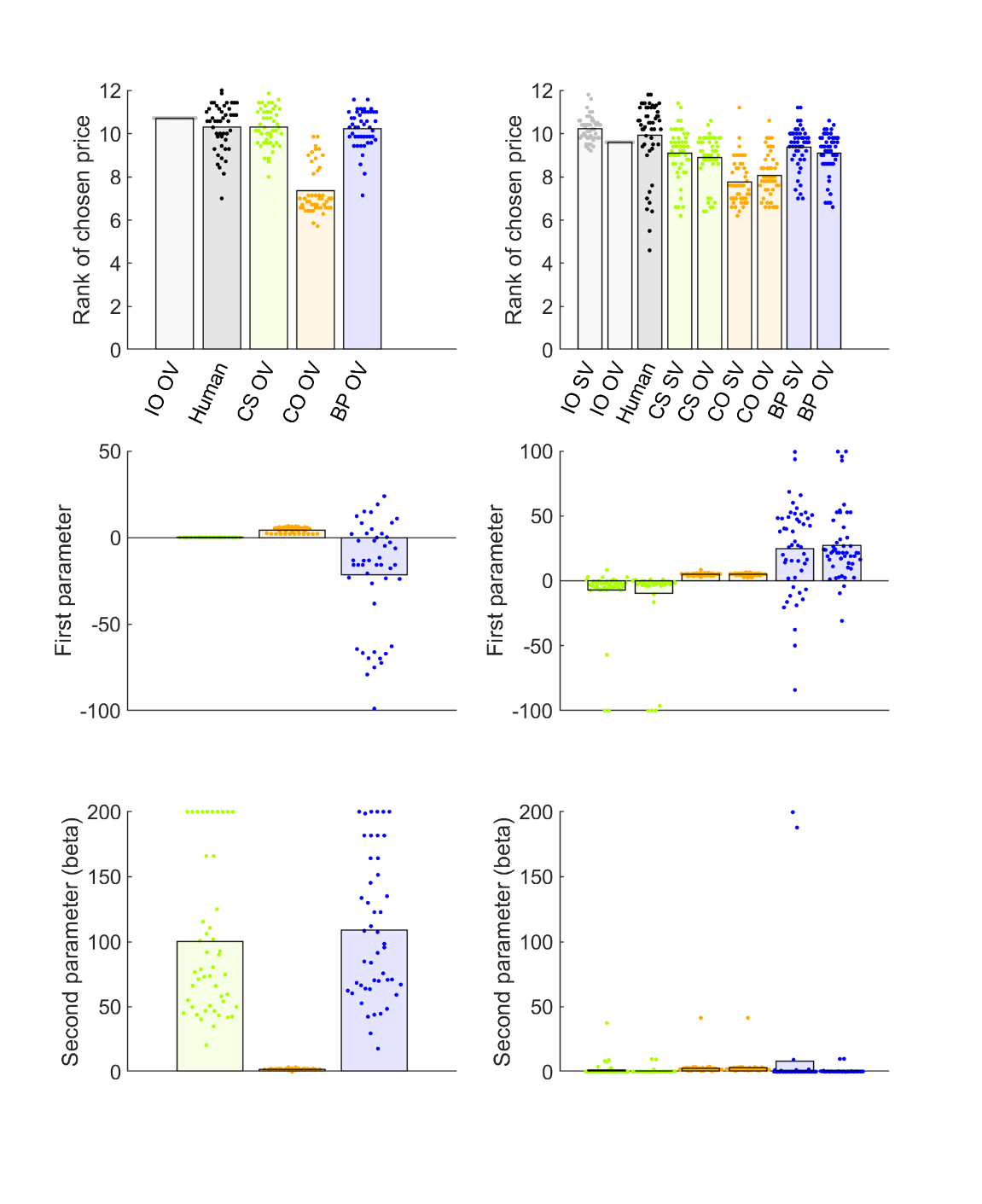


Figure S5. Cohen’s *d* effect sizes for pairwise comparisons of participants’ sampling rates in Study 1.

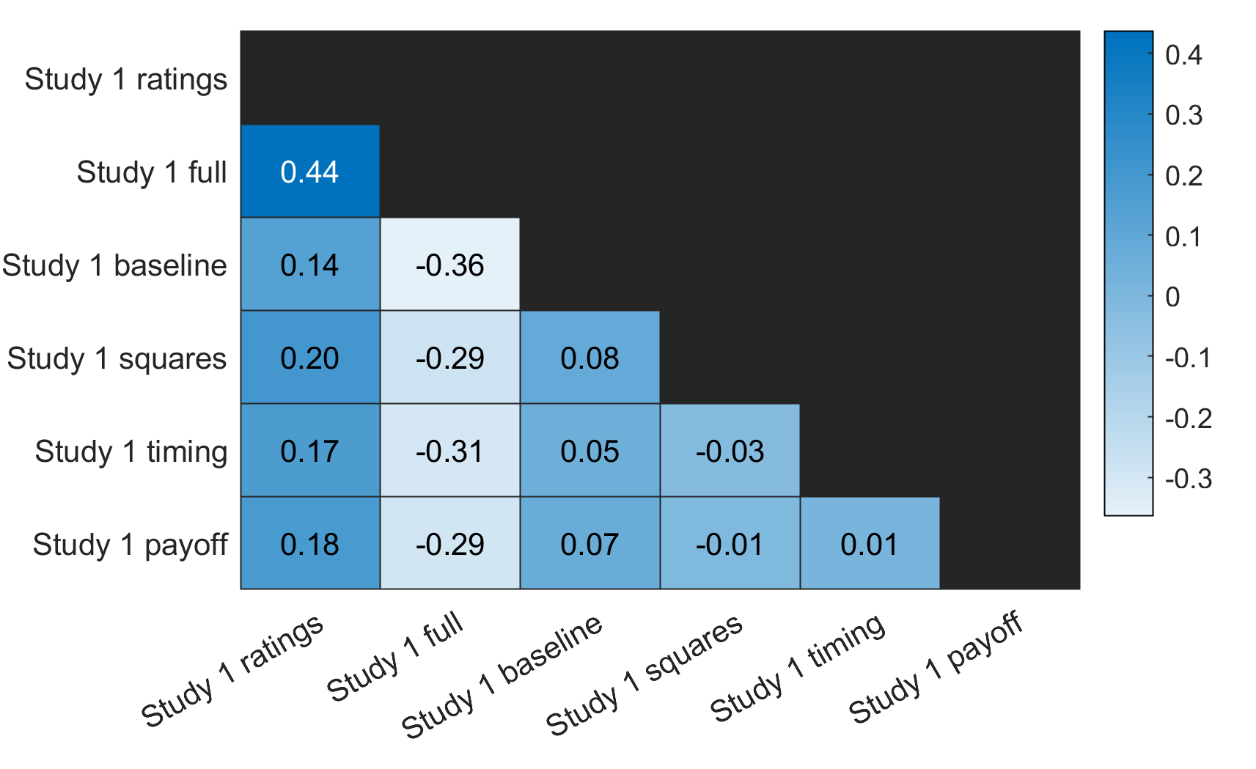


Figure S6. Model comparison for (columns from left to right): Study 1 baseline, squares, timing and payoff conditions. Top and middle rows show individual participants as points and bars show their mean values. The top row shows ranks of chosen prices. The second row plots the “first” or theoretical parameter values, estimated for each fitted model. The third row shows the “second,” or inverse temperature parameter beta, estimated for each fitted model. Abbreviations: IO = ideal observer, CS = cost to sample, CO = cut-off, BP = biased prior, OV = objective values.

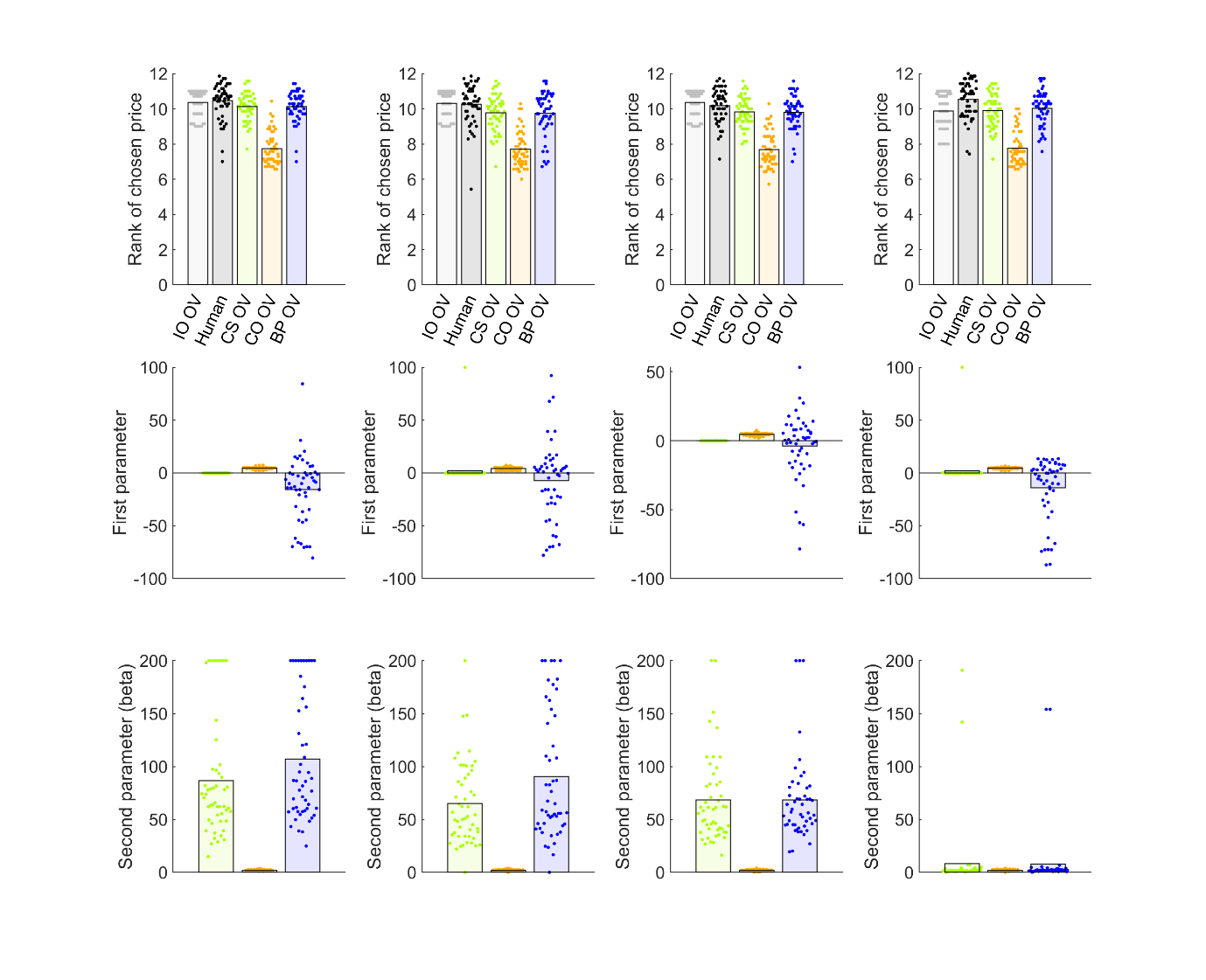


Figure S7. Model comparison for Study 1 ratings (left column) and full (right column) conditions. Top and middle rows show individual participants as points and bars show their mean values. The top row shows ranks of chosen prices. The second row plots the “first” or theoretical parameter values, estimated for each fitted model. The third row shows the “second,” or inverse temperature parameter beta, estimated for each fitted model. Abbreviations: IO = ideal observer, CS = cost to sample, CO = cut-off, BP = biased prior, OV = objective values, SV = subjective values.

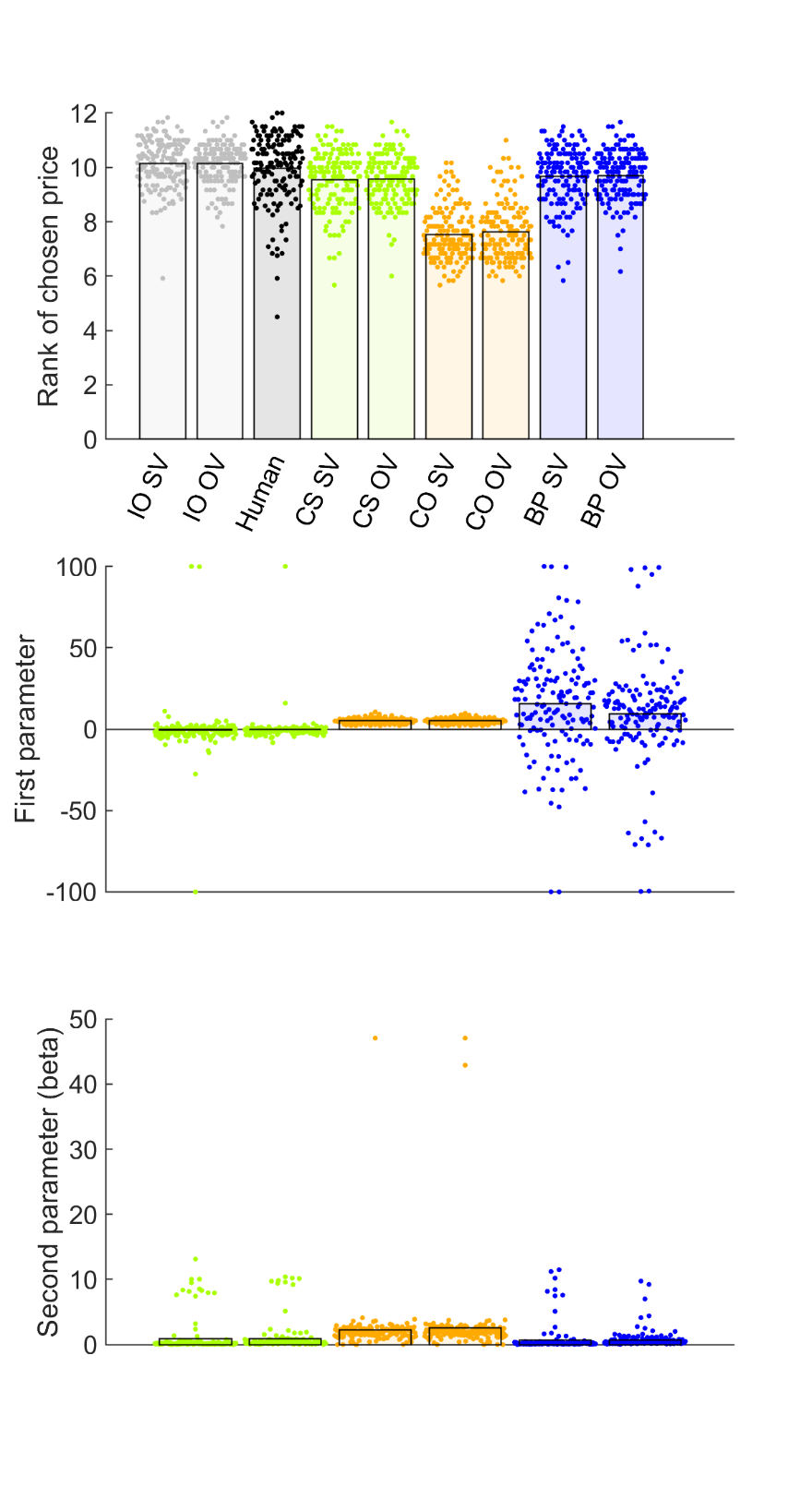
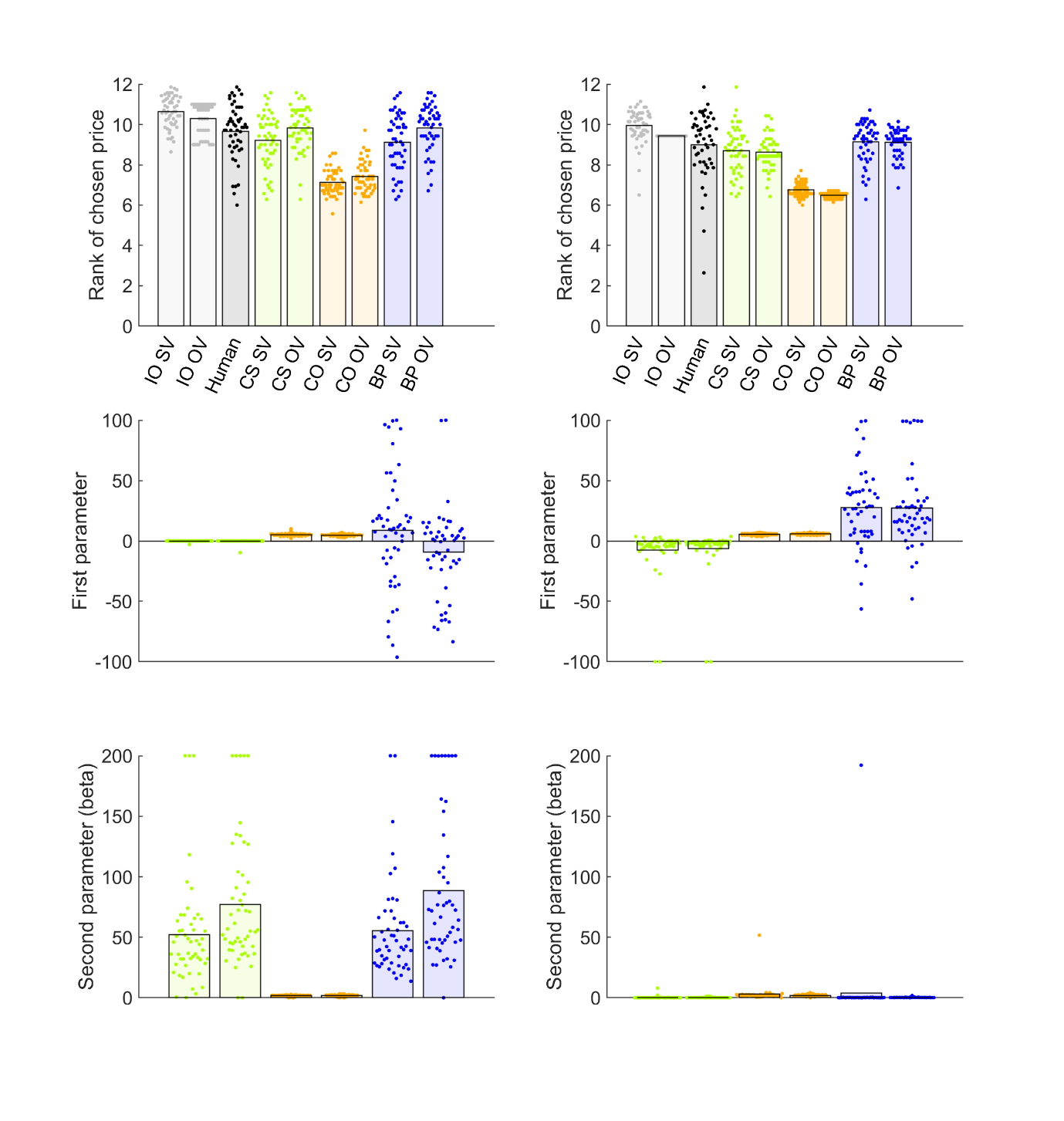


Figure S8. Model comparison for Study 2. Top and middle rows show individual participants as points and bars show their mean values. The top row shows ranks of chosen prices. The second row plots the “first” or theoretical parameter values, estimated for each fitted model. The third row shows the “second,” or inverse temperature parameter beta, estimated for each fitted model. Abbreviations: IO = ideal observer, CS = cost to sample, CO = cut-off, BP = biased prior, OV = objective values, SV = subjective values.

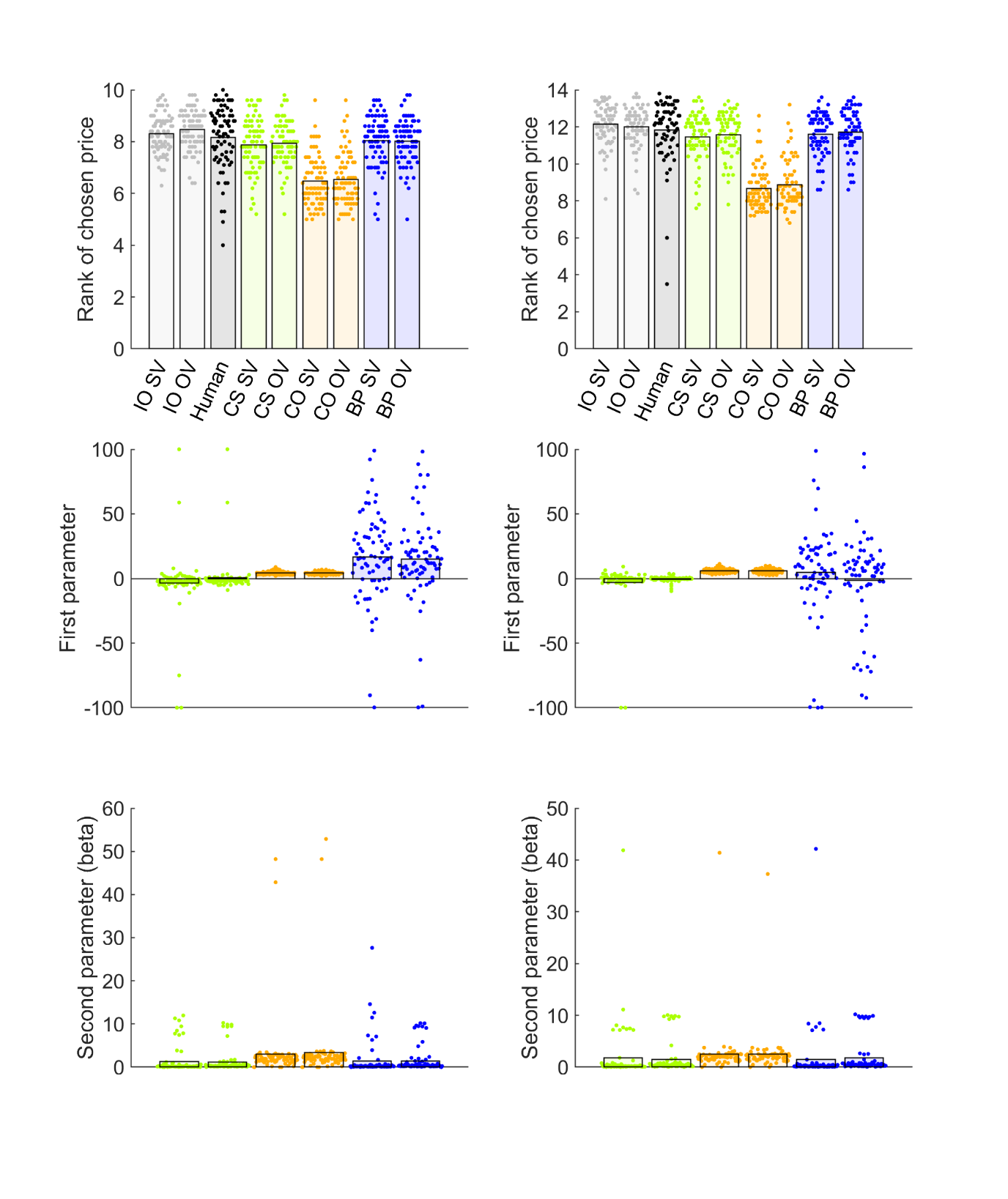


Figure S9. Model comparison for Study 3 10 options (left column) and 14 options (right column) conditions. Top and middle rows show individual participants as points and bars show their mean values. The top row shows ranks of chosen prices. The second row plots the “first” or theoretical parameter values, estimated for each fitted model. The third row shows the “second,” or inverse temperature parameter beta, estimated for each fitted model. Abbreviations: IO = ideal observer, CS = cost to sample, CO = cut-off, BP = biased prior, OV = objective values, SV = subjective values.