Perceived cost of sampling new options predicts decision biases in economic contexts

Didrika S. van de Wouw, Ryan T. McKay, Nicholas Furl

Royal Holloway, University of London

Corresponding author:

Didrika van de Wouw

Department of Psychology

Royal Holloway, University of London

Egham, TW20 0EX, United Kingdom

[sahiravandewouw@hotmail.com](mailto:sahiravandewouw@hotmail.com)

Data and code availability: <https://github.com/nicholasfurl/Model_fitting_hybrid_study>

Abstract

Considerable research has shown that people make biased decisions in “optimal stopping problems”, such as flat hunting or finding a spouse, options can be encountered sequentially, with little opportunity to recall rejected options or to know upcoming options in advance. Here, we use computational modelling to identify computational mechanisms that best explain biased in the context of an especially realistic version of this problem, the full-information problem. We show that the extent of deviation of participants’ sampling rates from an optimality model (bias) depends on the sequence length and the distribution of payoffs, while we find null effects for a variety of other methodological alternatives. We fit a variety of Bayesian models of bias and a heuristic model and found participants’ biased sampling rates were best explained if participants perceived sampling as especially costly. We therefore propose a new theoretical viewpoint for how human solve full information problems, and we consider implications for more widely studying this diverse class of decision problems, including real-world sequential decision scenarios.

Introduction

General Methods

Paradigm summary

We describe here how we use computational models to gauge the optimality of participants’ decisions and to build theories about the sources of participants’ bias in these tasks. First, we briefly describe the relevant features of our paradigms that are relevant for understanding the operations of the models. More specific methods for individual studies will be described in separate sections later.

We implemented full information optimal stopping problems in which participants attempted to choose the most preferred mobile phone contract price that they could. Prices in all studies reported herein were for flagship models by the top brands (e.g., iPhone, Samsung, Huawei), on an up to 5GB plan with unlimited texts and minutes. The 90 prices were actual prices (in GBP) of 2-year contracts offered by various UK retailers, as harvested from internet advertisements in the year before data collection. The use of these real-world prices was intended to maximise the likelihood that the distribution of option values used in our studies would approximate the “true” generating distribution of smartphone price options in the participants’ local market and thereby also approximate any prior expectations participants derived from their experience with smartphone contract prices prior to the study.

In some study conditions (Pilot full, Study 1 full and ratings conditions, Study 2 and both sequence length conditions of Study 3), the paradigm began with a “Phase 1” ratings task, in which participants viewed the full distribution of prices that might (or might not) appear as options later and rated each for its “attractiveness” or subjective value. As described below, some models operate over objective raw / prices (OV) and other models operate on the subjective value of the prices (SV). Participants also have the opportunity to learn the “generating” distribution of option values and thereby establish prior expectations about option values to come, and these values can therefore also be used to fix the models’ prior on its generating distribution of option values.

Next, in an optimal stopping task, participants engage with several fixed length sequences of option values, populated by prices randomly -sampled without replacement from the Phase 1 generating distribution (12 option values in all studies except Study 3, which compares performance for 10 versus 14 options). In each sequence, participants sequentially encounter these prices and, for each, decide whether to reject that option value (rendering it no longer accessible) and sample a new one, or to take / choose that option value (attempting to maximise the objective or subjective value of the price when choosing), which stops the search through the sequence and renders the upcoming new price samples no longer accessible. If the last price in a sequence is reached, that price became their choice by default.

Ideal observer optimality model

On our optimal stopping tasks, the number of options sampled before taking / choosing an option by human participants was compared to that of an ideal observer model, for which performance is Bayes-optimal. This finite horizon, discrete time, Markov decision process (MDP) model has been used in previous studies and mathematical details are also given in these papers (Cardinale et al., 2021; Costa & Averbeck, 2015; Furl et al., 2019; van de Wouw et al., 2022). The Bayesian version of the full information problem optimality model builds on the classic Gilbert and Mosteller model (Gilbert & Mosteller, 1966) as a starting point. Like the Gilbert and Mosteller model, the Bayesian optimality model’s expectations about future option values are derived from the model’s belief about the distribution from which future options are assumed to be generated (i.e., the generating distribution). More precisely, the utility *u* for the state *s* at sample *t* is the maximal action value *Q*, out of the available actions *a* in *A*, which in turn depend on the reward values *r* and the probabilities of outcomes *j* of subsequent states (i.e., the generating distribution), weighted by their utilities.

The terms appearing inside the curly brackets are taken collectively as the action value *Q*. is the reward that will be obtained in state *s* at sample *t* if action *a* is taken. The model described here develops the classic Gilbert and Mosteller formulation by reducing r by costs incurred by sampling again. This is embodied in the “cost to sample” penalty term *C* (See formula for below). As there was no extrinsic cost-to-sample in any of our experimental designs herein, *C* was always fixed to zero for the ideal observer model. The integral is taken over the possible states subsequent to the current sample. Each of these states is weighted by the probability of transitioning into it from the current state, given by , as derived from the generating distribution.

Like the original Gilbert & Mosteller formulation, the model we consider here computes the utilities for sampling again based on backwards induction. See the Supporting Information for Baumann et al. (2020) for a clear intuitive description of this backwards induction procedure as applied to the Gilbert & Mosteller formulation. In the application of our ideal observer, the model first considers the utility for the final sample *N* in the sequence, which is simply the reward value associated with the *N*th state (because taking the option is the only available action for the final sample in a sequence).

Next, the model works backwards through the sequence, iteratively using the aforementioned formula for when computing each respective action value *Q* for taking the option and declining the option for each *t*.

The version of the model we use here adds to the classic Gilbert & Mosteller formulation by making the reward function *R* customisable to the distribution of payoffs over ranks and by adding a Cost to sample term . Whenever the reward value of taking the current option is considered, this function *R* assigns reward values to options based on their ranks. *h* represents the relative rank of the current option.

In contrast, the reward value of sampling again is simply the cost to sample *C*, which would be negatively valued if the experimenter imposes such a cost in the experimental design.

This flexibility allowed us to model multiple reward payoff schemes within our studies and to examine how the ideal observer model samples changes its sampling strategy under different schemes and to test whether human participants take similar sampling strategies. In Pilot full, the full condition of Study 1, Study 3 and both sequence length conditions of Study 4, participants were instructed to try to choose the best price possible. To match these instructions, we implemented a continuous payoff function (resembling that of the classic Gilbert & Mosteller formulation), in which each relative rank would be rewarded commensurate with the value of its associated option. In Pilot baseline and the baseline, squares, timing, and prior conditions of Study 1, we implemented the payoff scheme to match participants’ instructions that they would be paid £0.12 for the best rank, £0.08 for the second best rank, £0.04 for the third best rank and £0 for any other ranks. Lastly, in the payoff condition of Study 1, we matched the instructions given to participants by rewarding 5 stars for the best rank, 3 stars for the second best rank, one star for the third best rank and zero stars for any other ranks.

Another feature added to our implementation of the ideal observer, compared to the Glibert & Mosteller base model, is the ability to update the model’s generating distribution from its experience with new samples in a Bayesian fashion, instead of this generating distribution being specified in advance and then fixed throughout the paradigm. The Bayesian version of the optimality model we used treats option values as samples from a Gaussian distribution with a normal-inverse-*χ2* prior. Before experiencing any options, the prior distribution has four initial parameters: the prior mean *μ0*, the degrees of freedom of the prior mean *κ*, the prior variance *σ*20 , and the degrees of freedom of the prior variance *ν*. This initialised distribution plays the role of a prior generating distribution when the first option value is sampled. The *μ0* and *σ*20 parameters of the generating distribution are then updated by the model following presentation of each newly sampled option value as each sequence progresses.

Here, we set the prior values of *μ* and *σ*2 in two possible ways (IO OV and IO SV, as described below). In previous studies, the mean and variance of the generating distribution has been fixed in advance by the mean and variance of the empirical option value distribution (e.g., Baumann et al., 2020), sometimes under the assumption that participants will have experience with this distribution prior to the study (Cardinale et al., 2021; Costa & Averbeck, 2015). When computing the prior generating distribution, and when inputting price values to the model as option values, we reflected the values around their mean, and rescaled the values to span 1 (the highest / worst price) to 100 (the best price) to ensure better prices were always more positively-valued such that the models were always solving a maximisation problem and that parameters for all models (OV and SV) would be on the same scales. Our ideal observer objective values model (IO OV) followed this procedure, which was applied to all the study conditions reported herein, whether or not participants were familiarised with the distribution of potential price options in an initial phase. This procedure assumes that the raw prices can be treated as a proxy for participants’ subjective value of the prices, and that all participants have equivalent subjective price valuations, and so a model that optimises only the raw prices when making decisions would therefore be an appropriate basis for comparison. Some of our study conditions (Pilot full, Study 1 full condition, Study 1 ratings condition, Study 3 and both sequence length conditions of Study 4) additionally involved a first phase in which participants rated the subjective values of the price options. We considered here that participants’ subjective valuation of prices may not exactly equal the raw price values, especially in their scaling, which may be relevant to full information problems, which consider option value magnitude, rather than relative option rank. For these conditions, we therefore also computed a second version of the ideal observer, IO SV. In the conditions for which we had subjective values from the initial phase available, we used each participants’ individualised ratings (subjective valuations) of the prices as option values input to IO SV, and we used the mean and variance of individual participants’ ratings distributions when initialising the prior of the generating distribution of the ideal observer model. Because conditions with an initial rating phase had two versions of the ideal observer model, each providing separate optimality estimates (IO OV and IO SV), we were able to ascertain whether use of objective or subjective values affects the strategy taken by the optimality model and, consequently, whether it changes the assessment of participant bias. For both IO OV and IO SV, log transformation was always applied to the objective or subjective values of the prices so that the generating distribution could better approximate normality, an assumption of the ideal observer. Shapiro-Wilk tests of normality confirmed that objective valued prices were indeed not normally distributed *W* = 0.94, *p* < .001, nor were subjective valuations of prices in the Pilot full study *W* = 0.85, *p* < .001, Study 1 full condition *W* = 0.94, *p* < .001, Study 3 *W* = , *p* < .001, Study 4 short sequence length *W* = , *p* < .001 or Study 4 long sequence length *W* = , *p* < .001. The ratings condition of Study 1 came just short of significance (*W* = 0.97, *p* = 0.08).

Theoretical models

We implemented the ideal observer model described above to assess the degree to which humans are biased to oversample or undersample, assuming that they either optimise their choices according to the objective (IO OV) or subjective values (IO SV) of the prices. We always implemented the ideal observer model with its parameter values fixed to ground truths established by the experimental design. Ideal observer models, however, are not appropriate for use as theoretical models of potentially-biased human sampling and choice behaviour, without modification added to account for the sources of individual variability in bias. That is, the ideal observer only models the computations leading to accurate choices but not to systematic sources of error, like oversampling or undersampling. To better understand which computations might be responsible for participants’ biased choices, we formulated a number of such theoretical models and fitted them to participant’s take option versus sample again choices. As mentioned above with respect to the ideal observer model, some previous studies have implemented models which aim to optimise the objective values of choices (e.g., Baumann et al., 2020; Cardinale et al., 2021; Costa & Averbeck, 2015; Lee, 2006) while other model implementations optimise subjective values of those options, obtained via a separate rating task (Furl et al., 2019; van de Wouw et al., 2021). Because there is no obvious determination of which procedure is correct, we implemented both objective values (OV) and subjective values (SV) versions of all our theoretical models for study conditions with a preceding rating task that enabled both model implementations. Then, we could assess using model comparison whether OV or SV models best fit human participant choices, or whether OV and SV models are relatively interchangeable (as we in fact discovered, see Results).

For every sample, the probabilities of the two available choices (take current option versus sample again) were computed by transforming actions values to probabilities using Softmax and then negative log likelihoods were computed and summed over choices for each participant. In each model, we freed one key parameter which was theoretically interpretable (These free parameters and their models are described below) and shown during parameter recovery to be capable of modulating the sampling rate (Supplementary Procedures Text A and Supplementary Figure S2 and upper panel of S3), and the inverse temperature parameter beta from the Softmax function (the starting value for beta was always 1). These two free parameters were fitted by minimising the negative log likelihood using fminsearch.m in MATLAB (Mathworks, Natick MA). Parameter recovery analyses of all the models described below showed at least adequate correlations between pre-configured and recovered parameters and strong correlations between sampling rates associated with pre-configured parameters and sampling rates associated with recovered parameters (Supplementary Procedures Text A and Supplementary Figures S1, S2 and S3). We implemented two parallel model comparison methods based on negative log likelihood values converted to Bayesian information criterion (BIC) values. First, we submitted the BIC values to repeated measures pairwise statistical tests using Bayes factors to ascertain whether pairs of models differed or had equivalent BIC values on average over participants. Next, we computed which model had the lowest (best) BIC for each participant and then plotted histograms to ascertain which model(s) dominated the others in terms of participant “wins”.

The objective and subjective values versions of the *“cut off” heuristic (CO OV and CO SV)* is the first model type we considered (Todd & Miller, 1999). This heuristic derives from a mathematically-optimal solution to the “Secretary problem” (Ferguson, 1989), an optimal stopping problem whereby a relatively simple solution can be mathematically proved by making numerous assumptions not made by the full information problem (e.g., the secretary problem solution assumes participants use no prior knowledge of the generating distribution, considers only relative ranks of option values and agents are rewarded only when choosing the top-ranked option). Although this heuristic derives from the optimal solution to a different optimal stopping problem than the full information problem we consider here, Todd & Miller (1999) propose that this heuristic might be robust to violations of the secretary problem assumptions and, as a heuristic, would be relatively simple for humans to compute on the fly in realistic settings. More specifically Todd & Miller (1999) propose that such a CO model can explain undersampling bias, as the model can perform nearly optimally (on secretary problems), while incurring fewer samples, which “satisfices” under conditions where there is a cost to sample (Note that the CO model has no cost to sample parameter). This heuristic has previous been fitted to human behaviour on full information optimal stopping problems, although little evidence was found favouring it in that study (Baumann et al., 2020). In the theoretical CO models we implemented, the model chooses sample again for every option until it reaches a cut-off, where the sequence position of the cut-off is fitted as the key theoretical free parameter. Then, the model continues to sample until they reach the next option with the best relative rank. Here, we used the optimal cut-off value (37% of the sequence length, rounded to the nearest integer) as the starting value during model fitting. Cut-off values between the optimal one lead to undersampling and cut-off values above the optimal value lead to oversampling.

We also considered objective and subjective values of *the cost to sample model (CS OV and CS SV)*. Like all the other models described below in this section, CS OV and CS SV use the Bayesian ideal observer described above as a base model, but assume that participants’ otherwise rational Bayesian computations can be biased by a suboptimal free parameter value. In the case of the CO OV and CS SV models, the fitted parameter to account for such bias was the cost to sample value *C* (See computation of in Ideal Observer Optimality Model section above. In such a model, participants would undersample if they intrinsically perceived sampling as costly and so had a negative *C* (there was no extrinsic sample cost in any of our paradigms) and would oversample if they perceived sampling as rewarding as so had a positive *C*. We initialised model fitting with a starting *C* value of 0 (i.e., the optimal value).

We used a similar approach when building *the optimism model (O OV and O SV)*. In this model, we added a new free parameter to the *μ*, the mean of the posterior generating distribution. This additional constant alters the mean value after it is updated by the current sample value and before the use of this posterior generating distribution in computing utilities . Negative values of this parameter bias an agent to compute pessimistic estimates of future option values by shifting the posterior mean (i.e., expectation) to be lower. This can lead to undersampling by making the current option appear more appealing compared to the artificially deflated expectation of option values resulting from continued sampling. Conversely, positive values of this parameter encourage oversampling, as the agent would have too optimistic an expectation of future option values to be gained by continued sampling. We initialised model fitting with a starting value of 0 (i.e., the optimal value).

In the *biased values model (BV OV and BV SV)*, 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 choice 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). Here, we passed the option values through a logistic function prior to input as option values to the ideal observer, which effectively thresholds the option values such that option values less than the midpoint parameter of the logistic function are roughly minimal and option values above this midpoint are roughly maximal, leaving only option values above an input value threshold as eligible for choice. We fixed the logistic slope to equal .2 (on the basis of successful parameter recovery using this value) while freeing and fitting the midpoint parameter / threshold of the logistic function. We picked the centre of the input value range as the starting value for the free logistic midpoint parameter when fitting to [participants’ choices.

The biased rewards model (BR OV and BR SV) 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 values / payoffs to outcome relative ranks. As with BV, we passed the option values through a logistic function, with slope = 1 (based on experience with parameter recovery), with the logistic midpoint parameter as the free parameter. During fitting, we initialised this midpoint value as the centre of the input value range. Then, the transformed values were assigned as reward payoff values in place of the ones otherwise suitable for the model (See ideal observer optimality section for more information on how reward payoffs are otherwise implemented in these models). Increasing this midpoint parameter value / reward threshold leads to increased sampling while decreasing this value leads to decreased sampling.

Pilot Studies Methods

Participants

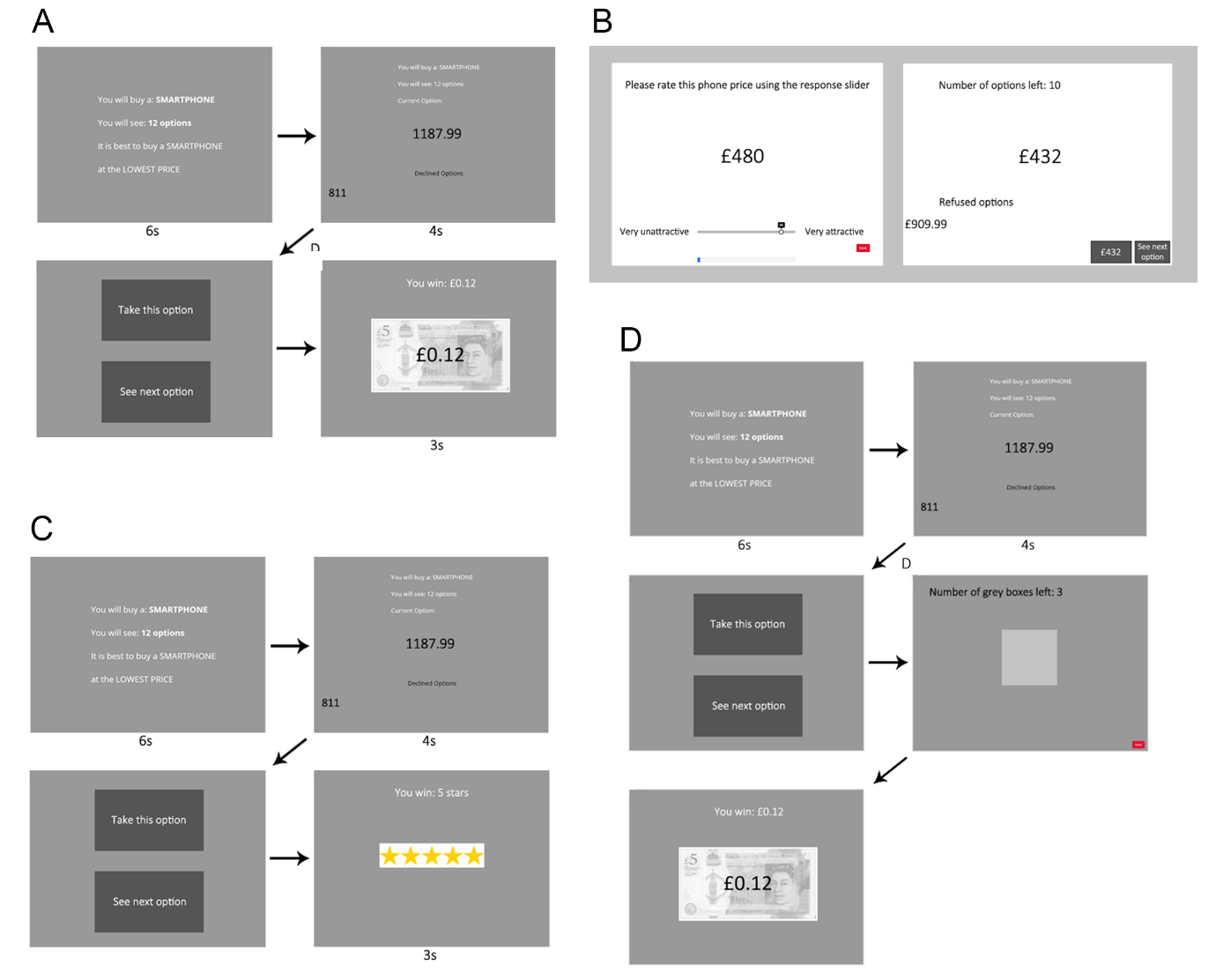
We recruited participants in both our pilot studies from the United Kingdom using the online Prolific platform (Prolific, 2014). We enrolled 50 participants into Pilot Baseline (there was no attention check available for exclusion of participants from analysis). We enrolled 51 participants into Pilot Full (Named “full”, as it possessed the full complement of additional methods taken from Furl et al., 2019, as described below in the Procedures). After excluding participants who failed the attention check used during the initial rating phase (See Supplementary Materials Text B), there remained for analysis 17 male and 29 female participants (Mage = 30.57, SDage = 11.36, range 18 to 75 years, four participants did not report their age).

Procedures

Gorilla Experiment Builder (Anwyl-Irvine et al., 2020) was used to create and host the studies. For the Pilot baseline study, we were interested in whether we could replicate participant undersampling bias, as observed previously (Cardinale, et al., 2021; Costa & Averbeck, 2015), in which participants sampled fewer options before decision relative to the Bayesian ideal observer model that was implemented in Costa and Averbeck (2015). We thus designed a paradigm that matched Costa and Averbeck (2015) in its methods particulars as closely as was practical, while concomitantly adapting it for an online setting. Consequently, there was no phase 1 ratings task in Pilot baseline. In the optimal stopping task (Figure 1A), participants attempted to choose the most highly-ranked smartphone prices as possible from five sequences of 12 price options each. The option value screen also presented the previously-rejected option values and the number of options remaining in the sequence. The five sequences were fixed in advance, so a given sequence’s option values and their order within the sequence was identical for every participant, although the sequences were presented in a random order.

Like Costa and Averbeck (2015), we rewarded participants financially for choosing one of the top three options in the sequence. Participants in Pilot baseline earned £0.12 per sequence if they chose the best price in the sequence, £0.08 if they chose the second best price, £0.04 if they chose the third best price, and £0 if they chose any other option. These bonus performance-based payments were earned on top of a flat fee, which for all our studies was set in line with Prolific’s recommended pay of £7.50 for one hour (participants typically finished the study in considerably less than this hour). Once a choice was made, participants viewed a feedback screen that informed them of their winnings for that sequence. The paradigm utilised fixed screen timings, meaning that participants automatically advanced through the screens, except when asked to make a decision (‘Take this option’ or ‘See next option’). Participants were warned about this feature in the instructions preceding the task.

Figure 1. Pilot and Study 1 paradigms. (a) Pilot baseline study and Study 1 baseline condition. (b) Pilot full and Study 1 full condition. (c) Study 1 payoff condition. (d) Study 1 squares condition.



For Pilot Full, we were interested in whether or not participant undersampling bias would continue to replicate using the same economic smartphone price task, but when implementing many of the methods particulars adapted from studies that revealed oversampling bias instead of undersampling bias (Cardinale et al., 2021; Furl et al., 2019). The logic is that, if one of these methods features is responsible for the oversampling bias seen in these earlier papers, then Pilot full should produce an oversampling bias, that contrasts with the undersampling bias we expected to see in Pilot baseline.

Pilot full (unlike Pilot baseline) added an initial ratings phase (Figure 1B), in which participants rated the “attractiveness” of the price, defined in the instructions as a willingness to purchase a phone at that price. Ratings were made by mouse click on a sliding scale from 1 to 100, with the slider initially hidden (to avoid slider biases)(Matejka et al., 2016), and showing the selected number shown it after the first click. Participants rated a total of 180 prices, presented one at a time in a random order, and comprising the 90 unique prices, but rated twice. The average over the two ratings for each price was then used as the subjective value input to the SV versions of the models. A blue progress bar was shown continuously at the bottom of the screen to visualise participants’ progression through the ratings phase.

The optimal stopping (second) phase of Pilot full (Figure 1B) included five sequences of 12 option values each. As in Pilot baseline, the option values in each sequence were fixed in advance but the sequences’ order was randomised. Unlike Pilot baseline, once participants chose one of the options, they then had to advance by button press through a series of grey squares that replaced the remaining options in that sequence. This ensured that participants could not finish the study early by choosing an early option. Also unlike Pilot baseline, the optimal stopping task was entirely self-paced - participants advanced by using their mouse to click on the buttons on the screen. After finishing a sequence, participants were directed to a feedback screen displaying their chosen price and the text: "This is the price of your contract! How rewarding is your choice?". Participants responded to this question using a slider scale ranging from not rewarding (1) to very rewarding (100). The purpose of this rating activity was to provide feedback about the quality of the participants’ choice, in lieu of the bonus payoff screen in Pilot Study 1, and to encourage them to reflect upon its reward value before moving on to the next sequence. These ratings do not provide hypothesis-relevant data and were not analysed. Participants were reimbursed a flat fee only - no bonus monetary payoff was awarded.

Pilot Studies Results and Discussion

As the two pilot studies are separate studies, with data collected at somewhat different times, we will descriptively, rather than statistically, compare them. Figures 2 and 3 show the number of human participant samples to decision for both of the pilot studies, which yielded similar numbers of samples, with a slight numerical increase for Pilot full.

Figure 2. Human participants’ numbers of samples to decision for all studies. Significant pairwise differences between conditions means within a study are shown as green horizontal lines (*p* < .05 after multiple comparison correction for the number of pairs in that study), which shows a significant difference only between the sequence lengths conditions in Study 3. Magenta horizontal lines connecting pairs of bars show conditions within each study where *BF*01 > 3 (i.e., moderate evidence favouring a null model with equal means). No pairs with *BF*10 > 3 were found.

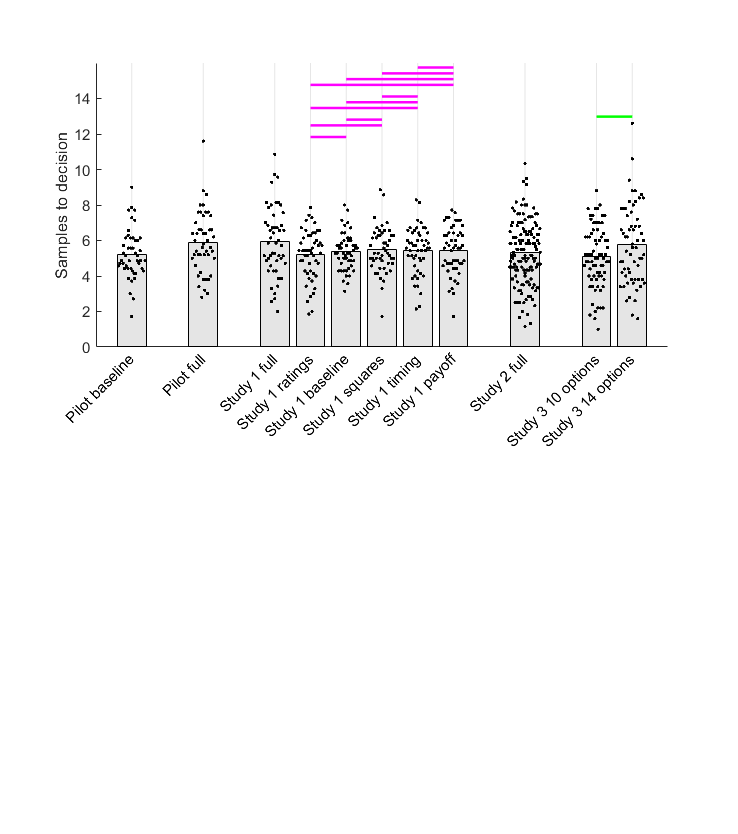


Figure 3. Model comparison for Pilot baseline (left column) and Pilot full (right column). Top and middle rows show individual participants as points and bars show their mean values. In the top row, human and IO samples are indicated by black horizontal lines when *BF01* > 3 (moderate evidence for equal means) and grey lines when *BF10* > 3 (moderate evidence for different means). Human participant data are the same as in Figure 1. The second row shows BIC values (lower values indicate better model fit) for participants (points) and their mean values (bars). Black horizontal lines indicate when *BF01* > 3. When *BF10* > 3, the horizontal line is coloured the same as the bar of the better model. The abundant orange and light green lines suggest that CO and CS are the best models. The third row plots the counts of participants for which each model was the best-fitting. Abbreviations: IO = ideal observer, CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values, SV = subjective values.

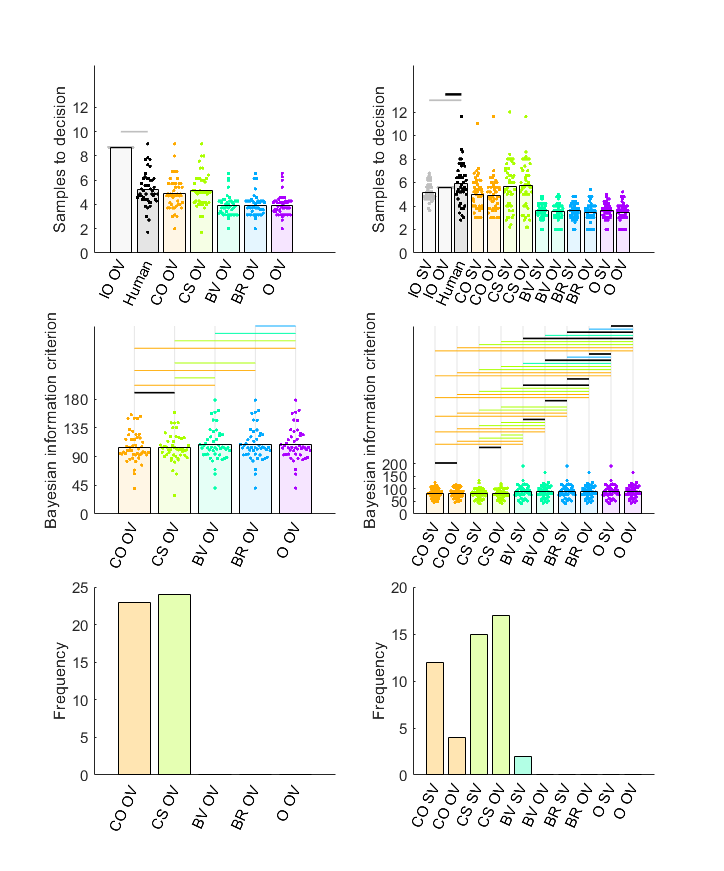


Figure 3 shows results from the comparison of human participants’ decisions with those of the models. As expected, we successfully replicated undersampling in the Pilot baseline condition (Figure 3, left column), where participants sampled fewer options than the ideal observer. The use of objective values by the ideal observer (i.e., IO OV) is the same as in previous studies showing undersampling (Baumann et al., 2020; Costa & Averbeck, 2015). No subjective values from an initial ratings phase were available in this Pilot baseline for comparison of IO SV with the IO SO model.

All the theoretical models, after fitting to Pilot baseline data, resembled the participants to some degree, as they all showed undersampling, compared to IO OV. Nevertheless, the CS model mostly closely approximated the human participants’ exact mean samples (Top row, Figure 3) and, moreover, CS was the best predictor of individual participants’ sampling rates (Figure S5). Model comparison using BIC scores, however, suggested that CS and CO models both might contribute to sampling bias in our participant sample. Bayesian pairwise tests (Figure 3, middle row), showed that CO and CS both had lower BIC values than any other model, although their mean BIC values were statistically equivalent to each other. And, approximately half of our participant sample each exhibited either CO or CS as their best-fitting model.

For Pilot full, however, there was no clear undersampling bias. Instead, participants’ sampling was statistically equivalent to optimal sampling, when compared to IO OV (the same optimality standard used in Pilot baseline) and they oversampled, when compared to IO SV. Although participants numerically sampled more in Pilot full than in Pilot baseline (Figure 2), this small increase did not appear to be sufficient to account for the disappearance of the undersampling bias in Pilot full, compared to Pilot baseline. Instead both IO SV and IO OV produced considerable fewer samples to decisions in Pilot full than IO OV did in Pilot baseline (grey bars, top left panel, Figure 3). We attribute this reduction in the optimal (IO) sampling rate in Pilot full compared to Pilot baseline to the different reward payoff function used. All relative ranks of choices were rewarded to some degree in Pilot full, depending on the magnitude of the option value, but only the top three ranks were rewarded in Pilot baseline. Moreover, optimal (IO) sampling was reduced in Pilot full compared to Pilot baseline regardless of whether or not subjective or objective values were used in the IO, and none of the other methods differences between Pilot baseline and Pilot full (e.g., the presence of a first phase, grey squares, timed screen advances, etc) can directly affect model implementation. It is therefore possible that reward payoff function might (at least partially) explain differences between studies that show undersampling (Cardinale et al., 2020; Costa & Averbeck, 2015) versus oversampling (Furl et al., 2019; van de Wouw et al., 2022) and we will test this possibility more rigorously in Study 1.

Although Pilot full appeared to show a different bias (i.e., somewhere between optimal and oversampling) than Pilot baseline (undersampling), the theoretical models that best fit the Pilot full behavioural data were similar to those in Pilot baseline. As in Pilot baseline, all the theoretical models tended to undersample to some degree, with CS (in the case of Pilot full, both CS OV and CS SV) most closely approximating participants’ mean number of samples (Figure 3, top right panel). At the individual participant level, the samples produced by CS and CO (with little difference between OV and SV) both predicted participants’ samples to decision better than any other theoretical models (Figure S6). Model comparison using BIC scores agreed with this pattern. CO (both OV and SV) and CS (both OV and SV) fit participants’ behaviour significantly better than any other model (See light green and orange horizontal bars indicating significant differences for these two models in the right panel in the second row of Figure 3). When considering the frequency of participants that best-fitted each model (right panel in the third row of Figure 3), CS appears to “win” the most participants, with little difference between CS SV and CS OV, while CO SV also appears to be the best-fitted model for a sizable number of participants.

In summary, the optimality IO model sampled more for Pilot baseline than for Pilot full, leading to undersampling in Pilot baseline and optimal (IO OV) or oversampling (IO SV) for Pilot full. As the difference in participants’ sampling between Pilot baseline and Pilot full was relatively small, and the difference with OV and SV versions of IO was relatively small, the different biases in the two studies presumably arose due to the differences in their reward payoffs. Nevertheless, both studies showed evidence that CO and, even more so, CS, well-fit participants choices.

Study 1

Fits of theoretical models to our Pilot baseline and Pilot full datasets produced some common results: namely, theoretical model fitting showed influences on participants’ sampling bias of both the cut-off heuristic (CO) and an intrinsic cost to sample in the context of a Bayesian solution to the full information problem. However, these two pilot studies showed some differences in results too. Although there was a small increase in participants’ sampling rates in Pilot full, compared to Pilot baseline, most of the differences pertained to the ideal observers’ sampling performance. Namely, both versions of the ideal observer (IO OV and IO SV) in Pilot full sampled less than IO OV in Pilot baseline. The result was that Pilot baseline replicated the classic undersampling effect (Costa & Averbeck (2015), while Pilot full showed a different result: Participants’ performance was somewhere between optimal (IO OV) and oversampling (IO SV). This finding is interesting, as most of the basic methods implemented in Pilot baseline were adapted from Costa & Averbeck (2015) and resulted in findings similar to that study. Meanwhile, we adapted many of the methods of Pilot full (i.e., the use of an initial ratings phase, grey squares to replace the remaining images after choice, self-paced screen timing, the absence of extrinsic / monetary payoff) from Furl et al., (2019) and then found results (a degree of oversampling) that resembled that study. This pattern raises a distinct possibility that at least one of these methods differences affects the nature of sampling bias.

Study 1 was therefore designed to put this possibility to the test by using six conditions to systematically vary the aforementioned methods differences and then analysing whether they affect sampling performance of participants and of the OV and SV versions of the IO (optimality) model. First, because participants’ samples to decision were so similar between Pilot baseline and Pilot full, we hypothesised that these methods differences would not substantially change participants’ number of samples to decision in Study 1. Second, because the computations used by the ideal observer models do not take into account the presence of a ratings phase, the use of grey squares to replace option screens after choice, self-paced timing and whether real money was used for incentivization, we do not expect the presence or absence of these task features to affect IO sampling behaviour. Third, because both IO SV and IO OV both sampled considerably less in Pilot full than Pilot baseline, we hypothesise that the use of objective (OV) or subjective values (SV) when computing the IO will have relatively little effect on participant bias. If all of the above holds, then the most likely possibility left over after eliminating these methods possibilities will be that sampling bias varies with whether the top three ranks are rewarded (as in Pilot baseline), as opposed to rewarding all ranks commensurate with the magnitude of the chosen option value (as in Pilot full). And finally, we hypothesise that the theoretical models that will best explain participants’ sampling biases will be either the CO heuristic or the CS model.

Study 1 Methods

Participants

As in the pilot studies, participants in Study 1 were enrolled from Prolific’s pre-screening facility to ensure that all participants were residents of the United Kingdom, to maximise familiarity with current UK smartphone market prices, denominated in in GPB). We enrolled independent participant samples into each of six conditions (See Procedures), aiming for fifty participants in each condition (chosen on the basis of our pilot studies, whose sample sizes proved sufficient to discriminate participant and IO sampling rates) . However, because of a technical difficulty with the participant recruitment platform, we overshot our data collection target by two participants, one in the timing condition and one in the ratings condition. Figure 1 describes participant dataset composition per condition, after exclusion of some participants from full and ratings conditions due to attention check failures (See Supplementary Text A).

Table 1. Demographics for each of the six Study 1 conditions

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Baseline  (*N* = 50) | Full  (*N* = 48) | Squares  (*N* = 50) | Payoff  (*N* = 50) | Timing  (*N* = 50) | Ratings  (*N* = 45) |
| Age |  |  |  |  |  |  |
| Mean  (SD) | 31.06 (10.63) | 32.45 (12.58) | 33.36 (10.40) | 30.41 (11.82) | 33.02 (11.66) | 33.36 (12.39) |
| Missing data points | 1 | 1 | 0 | 0 | 0 | 0 |
| Sex |  |  |  |  |  |  |
| Male | 15 | 13 | 12 | 18 | 10 | 12 |
| Female | 34 | 33 | 38 | 32 | 39 | 33 |
| Other | 1 | 2 | 1 | 1 | 0 | 0 |
| Prefer not to say | 0 | 0 | 0 | 0 | 1 | 0 |

Procedures

The study was developed using the study hosting software Gorilla Experiment Builder (Anwyl-Irvine et al., 2020). We implemented six conditions in Study 1, which systematically manipulated the presence or absence of four key task features. These features are summarised in the rows of Table 2 and the Baseline, full, squares and payoff conditions are visualised in Figure 1. We next cover each condition in turn. The *baseline condition* (Figure 1A) was nearly identical with the Pilot baseline study, except that it implemented seven sequences instead of five. Each sequence was fixed in advance, and then sequences were presented in random order. That means that, like Pilot baseline, Study 1 baseline adapted its methods from Cardinale et al. (2021) and Costa and Averbeck (2015). It is “baseline” in the sense that it possesses none of the methodological features being tested here and will serve as the basis for comparison against the other conditions, which each add one or more of these features. There was no initial rating phase, each sequence terminated and proceeded to the feedback screen immediately upon choice with no intervening grey squares, screens advanced with fixed timings (Figure 1A) and there was extrinsic monetary reward when the top three ranked options were chosen (as described for Pilot baseline). The *full condition* was identical to the Pilot full study (Figure 1B), except that it used seven sequences instead of five. Each sequence was fixed in advance, and then sequences were presented in random order. The condition is “full” in the sense that it resembles Furl et al. (2019) and van de Wouw (2022) in its design, which implements every one of the methodological features under test in this study. These features included the two-phase task structure (implemented using the same methods as Pilot full and Study 2 ratings condition), an incentivisation scheme in which participants are instructed to maximise the value of their choice, screen timings are self-paced timing and, when a choice is made, participants must page through grey squares, which replace the remaining options in the sequence. The *payoff condition* was the same as the baseline condition with the exception that participants did not receive the monetary incentivisation that they did in the baseline condition (Figure 2C). Participants were instructed to make choices to maximise the number of stars. Then, instead of receiving feedback regarding their earned bonus payments on the feedback screen (as in the baseline condition), participants were shown pictures of the number of stars that they earned for their choice: either five stars, three stars or one star, if they chose respectively the lowest, second lowest, or third lowest price in the sequence. The *squares condition* was the same as the baseline condition with the exception that, once participants had chosen an option that was not the last option, they had to press a key to advance through grey squares that replaced each forgone option until the end of the option sequence (Figure 1D). The *timing condition* was the same as the baseline condition with the exception that it incorporated a ‘next’ button in the top right corner of every option screen. This ensured that the entire paradigm was self-paced. The *ratings condition* was the same as the baseline condition with the exception that it added the same initial rating phase as in the Pilot Study and full condition (Figure 1B), while using the same optimal stopping task as the baseline condition (Figure 1A).

Table 2. Summary of conditions for Study 1

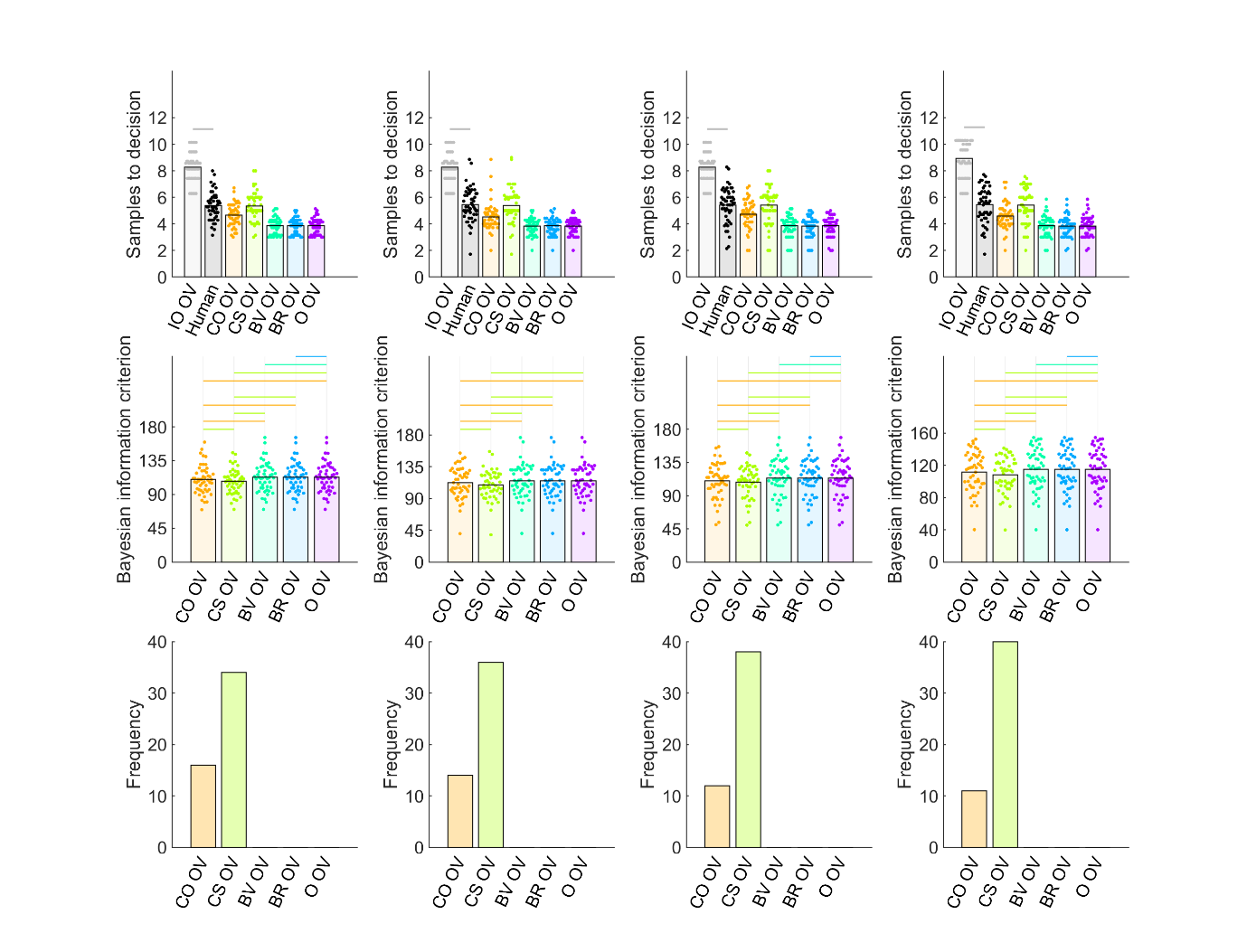
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Study 1 condition name | | | | | |
|  |  | Baseline | Full | Squares | Payoff | Timing | Ratings |
| Task feature | Grey squares |  | × | × |  |  |  |
| No monetary payoff |  | × |  | × |  |  |
| Self-paced timing |  | × |  |  | × |  |
| Rating phase |  | × |  |  |  | × |

Study 1 Results and Discussion

First, we tested whether any of the conditions would affect participants’ number of samples to decision. Similar to what we found with our pilot studies (Figure 2), there was a slightly higher number of participants’ samples in the full condition than any of other conditions. However, neither pairs of conditions including the full condition, nor any other pair showed a “significant” difference either by frequentist tests (using threshold *P* < .05, after multiple comparison corrected for the 15 condition pairs) or by Bayesian *t*-tests (using threshold *BF10* > 3, moderate evidence in favour of mean difference). However, nearly every pair of conditions had statistically equivalent means (magenta horizontal lines in Figure 1), according to Bayesian *t*-tests (all *BF01* > 3, moderate evidence in favour of null model), with the only exceptions being the five comparisons with the full condition, which were inconclusive. In short, our hypothesis was confirmed that participants’ sampling behaviour was not sensitive to presence or absence of any methodological feature.

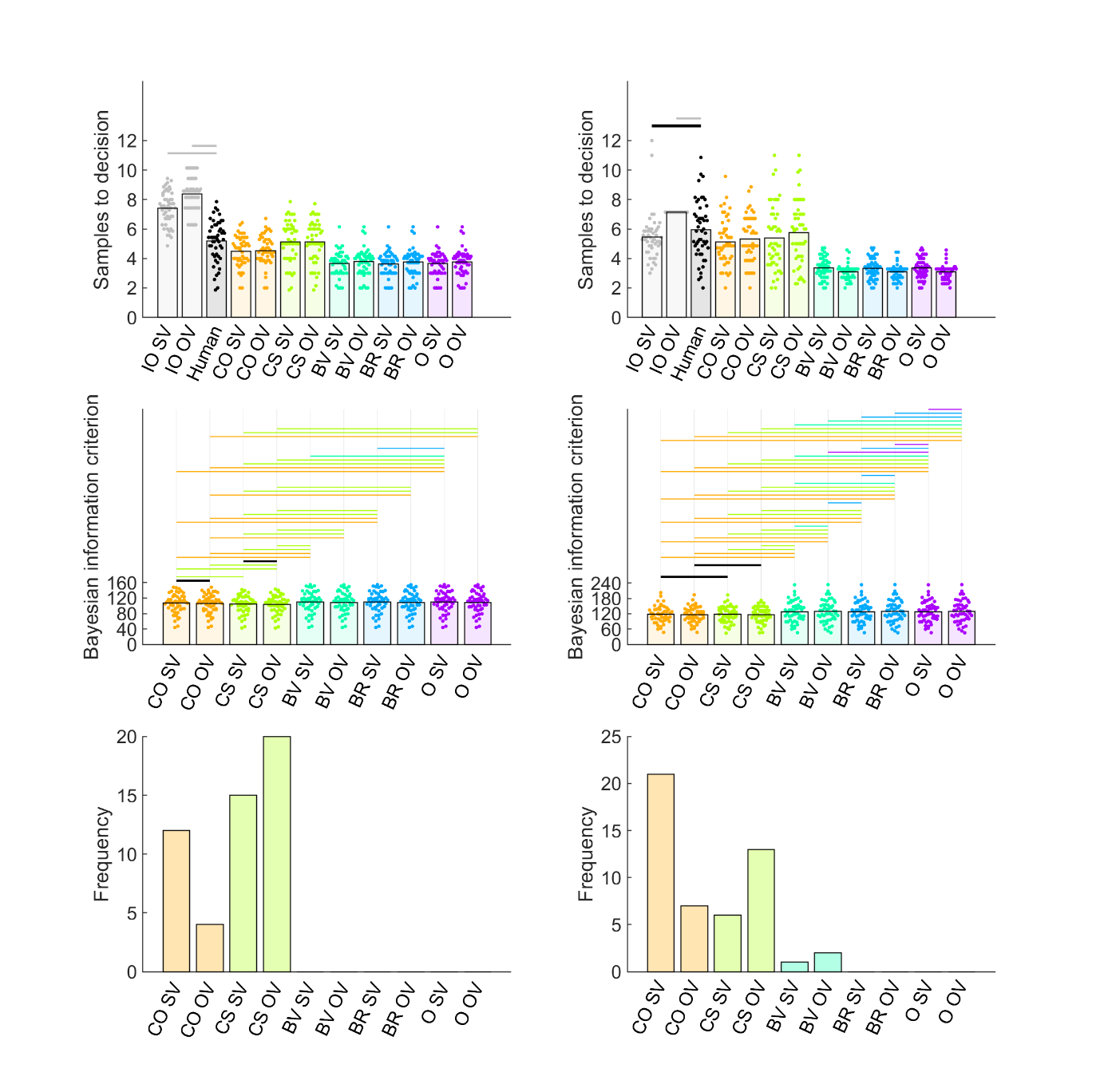
We next compared participants’ number of samples against those of the IO optimal models, to evaluate decision bias. The top row of Figure 4 shows results from the studies without any first phase, in which there is always an IO OV model with a payoff structure that rewards the top three ranks (in grey points), compared to participants (in black). All four conditions (baseline, squares, timing and payoff) show nearly-identical results, which collectively confirm our hypothesis: There is considerably higher sampling for IO OV than for human participants (undersampling), *BF10* > 3 (moderate evidence for different means). Thus, as expected, the presence or absence of grey squares, self-advanced timing and extrinsic monetary reward has no effect on IO performance and, therefore, the measurement of participant bias.

Figure 4. 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. In the top row, human and IO samples are indicated by black horizontal lines when *BF01* > 3 (moderate evidence for equal means) and grey lines when *BF10* > 3 (moderate evidence for different means). Human participant data are the same as in Figure 1. The second row shows BIC values (lower values indicate better model fit) for participants (points) and their mean values (bars). Black horizontal lines indicate when *BF01* > 3. When *BF10* > 3, the horizontal line is coloured the same as the bar of the better model. The abundant light green lines suggest that CS outperforms other models. The third row shows that the count of participants for which CS was the best-fitting was higher than for other models. Abbreviations: IO = ideal observer, CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values.



The top row of Figure 5 shows the same results for the two conditions with an initial rating phase, for which both IO SV and IO OV models are computable. Sampling bias in the ratings condition (left column, top row) for both IO SV and IO OV is similar to that observed for IO OV in all four conditions in Figure 4: the IO models sample more options than participants (undersampling). In contrast, the results of Study 1 full (left column) resemble those of Pilot full (Figure 3) in the sense that the numbers of samples chosen by IO SV and IO OV appear reduced compared to any of the other conditions (baseline, squares, timing, payoff, rating). The one remaining methodological difference between the IO models in Study 1 full and those of the remaining conditions, that could account for their decreased sampling behaviour, is that all choices in Study 1 full are rewarded to some degree, commensurate with the chosen option value, whereas choices in the other conditions are rewarded only when one of the top three ranked options are chosen. We conclude that this methods feature, rather than any of the methods features in Table 1, is what modulates the IO model’s number of samples to decision (for both SV and OV), and that human participants are not sensitive to this or any of the other methods features. The result is clear undersampling bias in every condition except the full condition. Note that the results of Pilot full and Study 1 full collectively can’t clearly indicate the direction of bias (if any), as the number of samples for IO models are much more similar to that of human participants in these conditions. We will resolve this issue in Study 2 by using a statistically more powerful study of the full condition.

Figure 5. Model comparison for Study 1 rating condition (left column) and full condition (right column). Top and middle rows show individual participants as points and bars show their mean values. In the top row, human and IO samples are indicated by black horizontal lines when *BF01* > 3 (moderate evidence for equal means) and grey lines when *BF10* > 3 (moderate evidence for different means). Human participant data are the same as in Figure 1. The second row shows BIC values (lower values indicate better model fit) for participants (points) and their mean values (bars). Black horizontal lines indicate when *BF01* > 3. When *BF10* > 3, the horizontal line is coloured the same as the bar of the better model. The abundant light green lines suggest that CS outperforms other models. The third row shows that the count of participants for which CS was the best-fitting was higher than for other models. Abbreviations: IO = ideal observer, CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values, SV – subjective values.



Last, we evaluated computational theoretical models that could explain biases in individual participants. All the conditions except the full condition produced similar results, and both CS OV and CS SV (in the ratings condition) produced similar results. Only CS models closely approximated participants’ mean number of samples to decision (Figures 4 and 5, first row) and CS models closely approximated participants’ mean rank of chosen price (Figures S7 and S12, top panel). Moreover, the CS OV and CS SV models were the only theoretical models to predict individual participants’ number of samples nearly-perfectly, while the other models do not (Figures S8-S11, S13, S14). CS OV and CS SV models fitted participants’ behaviour with better BIC’s (BF10 > 3) than any of the other models, including also the CO heuristic (Figure 4, third row) and, moreover, best-fit many more participants than any other models (Figure 4, third row). The results in Study 1 full were similar, though with a somewhat stronger contribution of the CO heuristic. Most notably, the CS and CO models did not differ in their BIC’s and both models best-fit a considerable number of participants, with CO SV and OV best-fitting somewhat more participants than CS SV and CS OV. Again, as with the comparison with IO performance discussed in the previous paragraph, Pilot full and Study 1 full do not produce as clear a result as the other conditions. We therefore will resolve this issue in Study 2 by using a statistically more powerful study in the full condition.

In summary, our hypotheses about the effects of methods features were largely confirmed. Participants sampled roughly the same amount across conditions, regardless of methods features. The IO models were also not sensitive to these methods details (indeed, they are not even programmed with information about whether there are grey squares, etc.), although the IO model appears to sample much more when only the top three ranking options are rewarded (in all conditions but the full condition), compared to when all choices are rewarded, depending on the value of the chosen option (in the full condition), leading to more prominent undersampling bias in all conditions, compared to the full condition. Participants’ sampling biases seem best explained, if they feel there is an intrinsic cost / reward value associated with further sampling (i.e., the CS model), with some potential contribution of the CO heuristic in the full condition. There was relatively little effect of SV versus OV for either the IO or theoretical modelling results.

Study 2

The Pilot full study and the Study 1 full condition showed that an optimal stopping task in which all choices are rewarded according to their value, leads to reduced IO sampling, compared to a variety of different conditions with different methods, but in all of which only the top three ranking choices were rewarded. Consequently, the mean number of samples of the IO SV and IO OV models were more similar to participants’ sampling rates in Pilot full and Study 1 full (Figures 3 and 5), making the exact direct difficult to determine, compared to the other conditions, where a substantial and robust undersampling bias was relatively obvious (Figures 3,4 and 5). The current study aimed to obtain a higher quality estimate of participant sampling bias in the full condition by overcoming a number of limitations of previous full condition designs. We increased the target sample size from approximately 50 to 151 and we generated a new set of sequence option values for every participant (In the two previous implementations, five or seven sequences of the same fixed option values were available to all the participants). This study further provided a robust effect size estimate that could be used for pre-registering an a priori sample size for Study 3.

Study 2 Methods

Participants

One hundred fifty one participants based in the UK enrolled, using the Prolific participant recruitment platform Prolific. This study forewent the attention check and all participants were included in analysis.

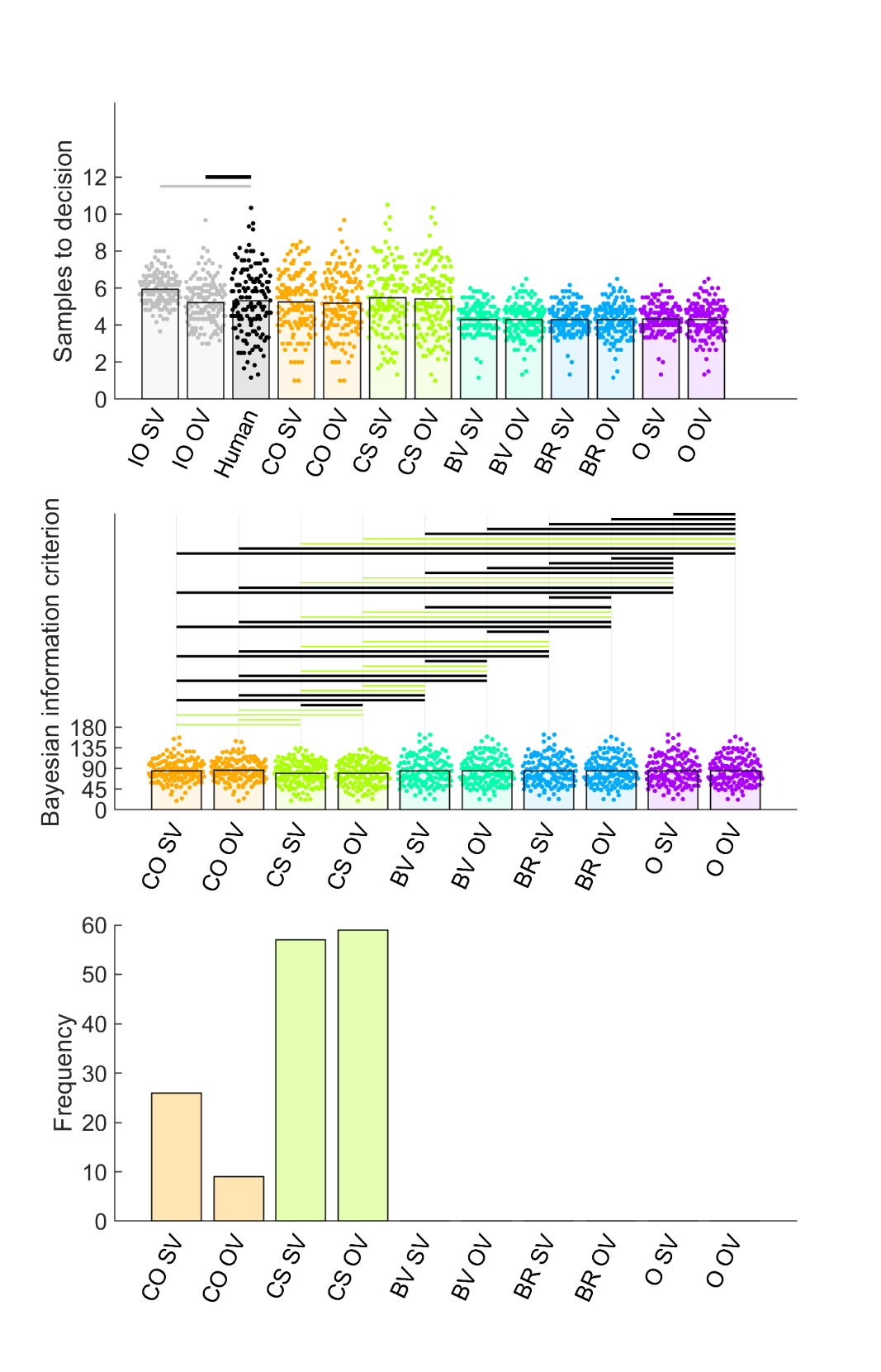
Procedures.

The study was developed in Javascript jsPsych 7.3.1 (de Leeuw et al., 2023), which repeated most of the methods of Pilot full and Study 1 full. In Phase 1, two lists of the 90 prices were concatenated and then its elements randomised and presented to participants sequentially above a 1 to 100 scale, in which they indicated the “attractiveness” of each price via mouse click. Then they performed an optimal stopping task with six sequences of 12 price option values, randomly sampled without replacement from the 90 phone contract prices. The study implemented self-paced screen timing. There were no grey squares. Instead, upon choice, the paradigm proceeded directly to the feedback screen (which is as Described above for Pilot full and Study 1 full). Participants were instructed to choose the best possible price.

Study 2 Results and Discussion

Participants appeared to sample about as many prices in Study 2 as in previous studies (Figures 1 and 6). Compared to the number of samples taken by IO OV, participants’ sampling was statistically equivalent (BF01 > 3, moderate evidence for null model). Meanwhile, compared to IO SV, participants slightly undersampled (BF10 > 3, moderate evidence for mean difference). The result does not exactly match that of either Pilot full or Study 1 full, although we note that both those studies delivered the same fixed option orders per sequence to every participant (Note how IO performance in Figures 3 and 5 show no variability) and so those results are tied to one specific option set. In contrast, here, our estimate of IO performance generated different option values for the sequences of for every participant, and so can be expected to better generalise between beyond one stimulus set. Study 2 also found no evidence for *over*sampling, either when comparing to IO SV or IO OV, which was found in several previous studies which used similar methods and the same ideal observer model (IO SV) as Study 2 (Furl et al., 2019; van de Wouw et al., 2022). We conclude that oversampling must not arise specifically from any of the methods details we consider here. Even the payoff scheme (i.e., we concluded above that IO samples less when choices are rewarded depending on the value of the chosen option, relative to rewarding only the top three ranks) or using the subjective values (SV) instead of the objective values (OV) in the IO model was not enough to induce oversampling. We hypothesise that it is the content domains of these studies, which used pictures instead of prices (e.g., attractive faces, foods, holiday destinations). Future research will need to explore this topic further.

Figure 6. Model comparison for Study 2 full. Top and middle rows show individual participants as points and bars show their mean values. In the top row, human and IO samples are indicated by black horizontal lines when *BF01* > 3 (moderate evidence for equal means) and grey lines when *BF10* > 3 (moderate evidence for different means). Human participant data are the same as in Figure 1. The second row shows BIC values (lower values indicate better model fit) for participants (points) and their mean values (bars). Black horizontal lines indicate when *BF01* > 3. When *BF10* > 3, the horizontal line is coloured the same as the bar of the better model. The abundant light green lines suggest that CS outperforms other models. The third row shows that the count of participants for which CS was the best-fitting was higher than for other models. Abbreviations: IO = ideal observer, CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values, SV – subjective values.



Even if our results do not explain oversampling on these picture-based tasks, our model fitting results do provide a theoretical explanation for the biases that we observe in the economic / price domain. Study 2 confirms the results we found throughout our studies reported herein, with the best evidence favouring CS as an appropriate model for most participants, with perhaps some contribution of the CO heuristic. CO and CS models reasonably reproduced participants’ mean number of samples to decision (Figure 6, first row), although CO models poorly reproduced the ranks of chosen prices that participants achieved (Figure S15). Both CS models (OV and SV) also more highly correlated with individual participants’ number of samples than CO or any other model (Figure S16). The two CS models showed a statistically better BIC than every other model, including both CO models (Figure 6, second row) and they were the best-fitting model for more individual participants than another model, including the CO models (although the CO models collectively “won” nearly 40 participants).

Study 3

Figure 1 suggests that participants are loth to change how much they sample, even when various methods features change and even though the IO model appears sensitive to the payoff scheme (i.e., whether every chosen rank is rewarded according to its option value as in Pilot full and Study 1 full, or whether only the top three ranks are rewarded as in all the other conditions). Indeed, our theoretical modelling fitting results so far suggest that participants’ sampling is controlled by a perceived intrinsic cost to sample, which would indeed limit them from increasing how much they sample, relative to an ideal observer which perceives no sampling cost.

The goal of Study 3 was to ensure that our implementation of the optimal stopping task was not somehow problematic and that it is in practice possible to experimentally modulate how much participants sample. Costa & Averbeck (2015) manipulated the sequence length (i.e., how many options were available in each sequence) and found participants were willing to increase the number of samples for longer sequences. Nevertheless, Costa & Averbeck found that undersampling was more pronounced at higher sequence lengths, because participants did not increase how much they sampled to the same degree as did the ideal observer, suggesting that participants were reluctant to change how much they sample also in that study. Here, we attempted to replicate this effect of sequence length on participants’ average number of samples, using sequence lengths of 10 and 14 options (We have reported results for 12 options so far). We hypothesise that participants will increase their number of samples to decision for longer sequences, although that change will appear small, relative to how much the ideal observer increases its number of samples to decision, leading to greater undersampling bias for longer sequences. We also predict that an intrinsic cost to sample will continue to best-explain participants’ sampling bias for both 10 option and 14 option sequences.

Study 3 Methods

The preregistration of Study 3 can be found at <https://osf.io/vcf7u>. We enrolled 140 participants from the UK using Prolific, where half engaged with sequences of length 10 and the other half engaged with sequences of length 14. As explained in the pre-registration, the sample size was intended to double that of Costa & Averbeck (who used a more powerful repeated-measures design and who were able to use more trials per participant in-lab, while we needed a shorter online study). The procedures were identical to Study 2, using the same jsPsych code, merely changing the sequence length.

Study 3 Results.

Figure 1 shows that our hypothesis was confirmed that participants would sample more for longer sequences, replicating the findings in Costa & Averbeck (2015). Moreover, Figure 7 showed that our hypothesis was confirmed that there would be more evidence for undersampling at longer sequences, also replicating Costa & Averbeck (2015). Indeed, at sequence length 10, participants actually oversampled compared to IO OV, with inconclusive evidence for IO SV. In contrast, at sequence length 14, participants undersampled compared to both IO SV and IO OV.

Our model-fitting also confirmed our hypothesis that participants’ sampling biases could be explained by an intrinsic / perceived cost to sample. The model-fitting results for the two sequence length conditions of Study 3 closely resembled those of Study 2 and of each other. Although CO and CS models both reasonably approximated how much participants sampled (Figure 7, first row), CO poorly reproduced participants’ rank (Figure S17, first row) and the two CS models outperformed all other models (including CO) when predicting individual participant sampling (Figures S18 and S19). The CS models has better BIC scores (Figure 7, second row) and better fit more individual participants (Figure 7, third row) than all other models. There was little differentiation between the results of CS SV and those of CS OV.

In summary, participants can and will change their sampling behaviour to a degree in some contexts. However, at least on tasks in the economic domain that we studied here, participants’ number of samples are “held in place” by their perception of an intrinsic cost of sampling further, that discourages them from increasing their sampling, and leading to ever-increasing undersampling bias, as sequences lengthen.

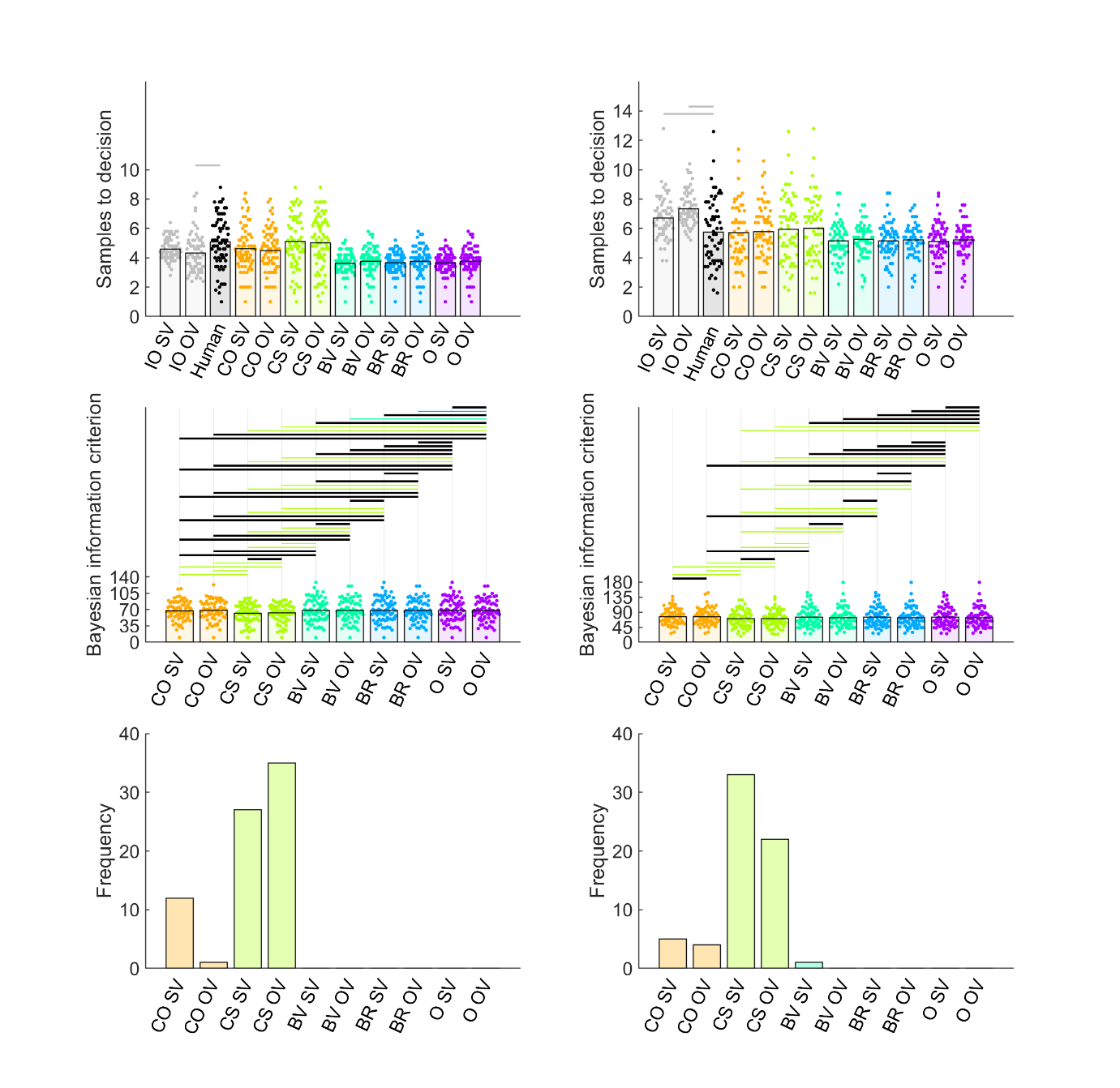


Figure 7. Model comparison for Study 3 10 options condition (left column) and 14 options condition (right column). Top and middle rows show individual participants as points and bars show their mean values. In the top row, human and IO samples are indicated by black horizontal lines when *BF01* > 3 (moderate evidence for equal means) and grey lines when *BF10* > 3 (moderate evidence for different means). Human participant data are the same as in Figure 1. The second row shows BIC values (lower values indicate better model fit) for participants (points) and their mean values (bars). Black horizontal lines indicate when *BF01* > 3. When *BF10* > 3, the horizontal line is coloured the same as the bar of the better model. The abundant light green lines suggest that CS outperforms other models. The third row shows that the count of participants for which CS was the best-fitting was higher than for other models. Abbreviations: IO = ideal observer, CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values, SV – subjective values.

Discussion

Subjective versus objective values make little difference.

Using multiple random sequences is important for stable results, as opposed to using sequences fixed in advance (as done routinely with the beads task).

Reference List

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. Behavior Research Methods, 52(1), 388–407. <https://doi.org/10.3758/s13428-019-01237-x>

Baumann, C., Singmann, H., Gershman, S. J., & von Helversen, B. (2020). A linear threshold model for optimal stopping behavior. *Proceedings of the National Academy of Sciences 117,* 12750–12755. <https://doi.org/10.1073/pnas.2002312117>

Cardinale, E. M., Pagliaccio, D., Swetlitz, C., Grassie, H., Abend, R., Costa, V., Averbeck, B. B., Brotman, M. A., Pine, D. S., Leibenluft, E., & Kircanski, K. (2021). Deliberative choice strategies in youths: Relevance to transdiagnostic anxiety symptoms. *Clinical Psychological Science*, 1–11. <https://doi.org/10.1177/2167702621991805>

Costa, V. D, & Averbeck, B. B. (2015). Frontal-parietal and limbic-striatal activity underlies information sampling in the best choice problem. *Cerebral Cortex* *25,* 972–982. <https://doi.org/10.1093/cercor/bht286>.

de Leeuw, J.R., Gilbert, R.A., & Luchterhandt, B. (2023). jsPsych: Enabling an open-source collaborative ecosystem of behavioral experiments. Journal of Open Source Software, 8(85), 5351, <https://joss.theoj.org/papers/10.21105/joss.05351>.

Ferguson, T. S. (1989). Who solved the secretary problem? *Statistical Science 4*, 282–289. <https://doi.org/10.1214/ss/1177012493>

Furl, N., Averbeck, B. B., & McKay, R. T. (2019). Looking for Mr(s) Right: Decision bias can prevent us from finding the most attractive face. *Cognitive psychology, 111,* 1–14. <https://doi.org/10.1016/j.cogpsych.2019.02.002>

Gilbert, J. P., & Mosteller, F. (1966). Recognizing the Maximum of a Sequence. *Journal of the American Statistical Association, 61*, 35–73. <https://doi.org/10.2307/2283044>

Lee, M. D. (2006). A hierarchical bayesian model of human decision-making on an optimal stopping problem. *Cognitive Science 30,* 1–26. <https://doi.org/10.1207/s15516709cog0000_69>

Matejka, J., Glueck, M., Grossman, T., & Fitzmaurice, G. (2016). The effect of visual appearance on the performance of continuous sliders and visual analogue scales. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 5421–5432

Prolific. (2014). Available at: <https://www.prolific.co>

Todd, P.M. & Miller, G.F. (1999). From pride and prejudice to persuasion: Satisficing in mate search. In G. Gigerenzer & P.M. Todd (Eds.), *Simple Heuristics that Make Us Smart*, pp 287–308. New York, NY: Oxford University Press.

Valone, T.J., Nordell, S.E., Giraldeau, L-A. & Templeton, J.J. (1996). The empirical question of thresholds and mechanisms of mate choice. *Evolutionary Ecology, 10,* 447–455.

van de Wouw, S., McKay, R., Averbeck, B. J., & Furl, N. (2022). Explaining human sampling rates across different decision domains. *Judgment and Decision Making* 17, 487-512.

Supplementary Procedures

Text A: 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 wished to ensure that our fits to human behaviour would provide more reliable fits than those simulated during parameter recovery, and so we simulated 20 participants (our empirical studies recruited at least approximately 50 participants) with only five sequence per participant with twelve options per sequence. To parallel the structure of our empirical paradigms, we created a generating distribution (separately for each simulated participant) of 426 option values, randomly-produced from a Gaussian distribution with mean 50 and standard deviation of 5 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. Configured and estimated parameters tended to correlate (Figure S1 and lower panel of Figure S3), especially for the model about which we will base our final conclusions, CS. 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).

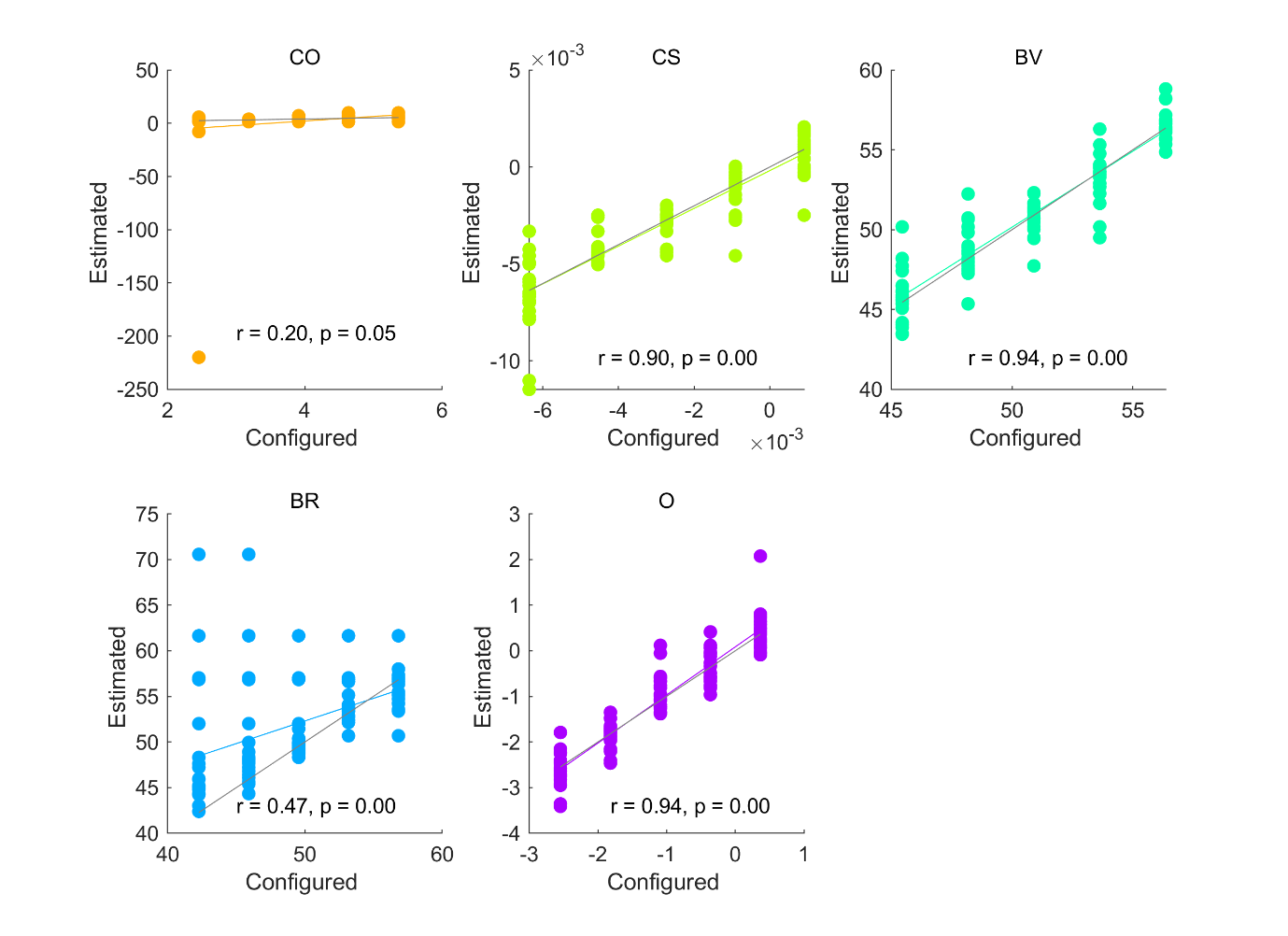


Figure S1. We compared configured parameter values (horizontal axes) for 20 simulated participants (each shown as an individual scatter point) against parameter values estimated (vertical axes) after fitting the five models to decisions simulated using the configured parameter values. Correlations can achieve > .9. 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. Note that the correlation for CO is disrupted by a single outlying point (r > .9 after removing this point). Generally, we found the CO model to be prone to occasional outlying fitted parameters for extremely low cut-off values like 1 or 2 sequence positions. Importantly, the model we later conclude to best explain human sampling choices, CS, shows highly accurate parameter recovery. Abbreviations: CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism.

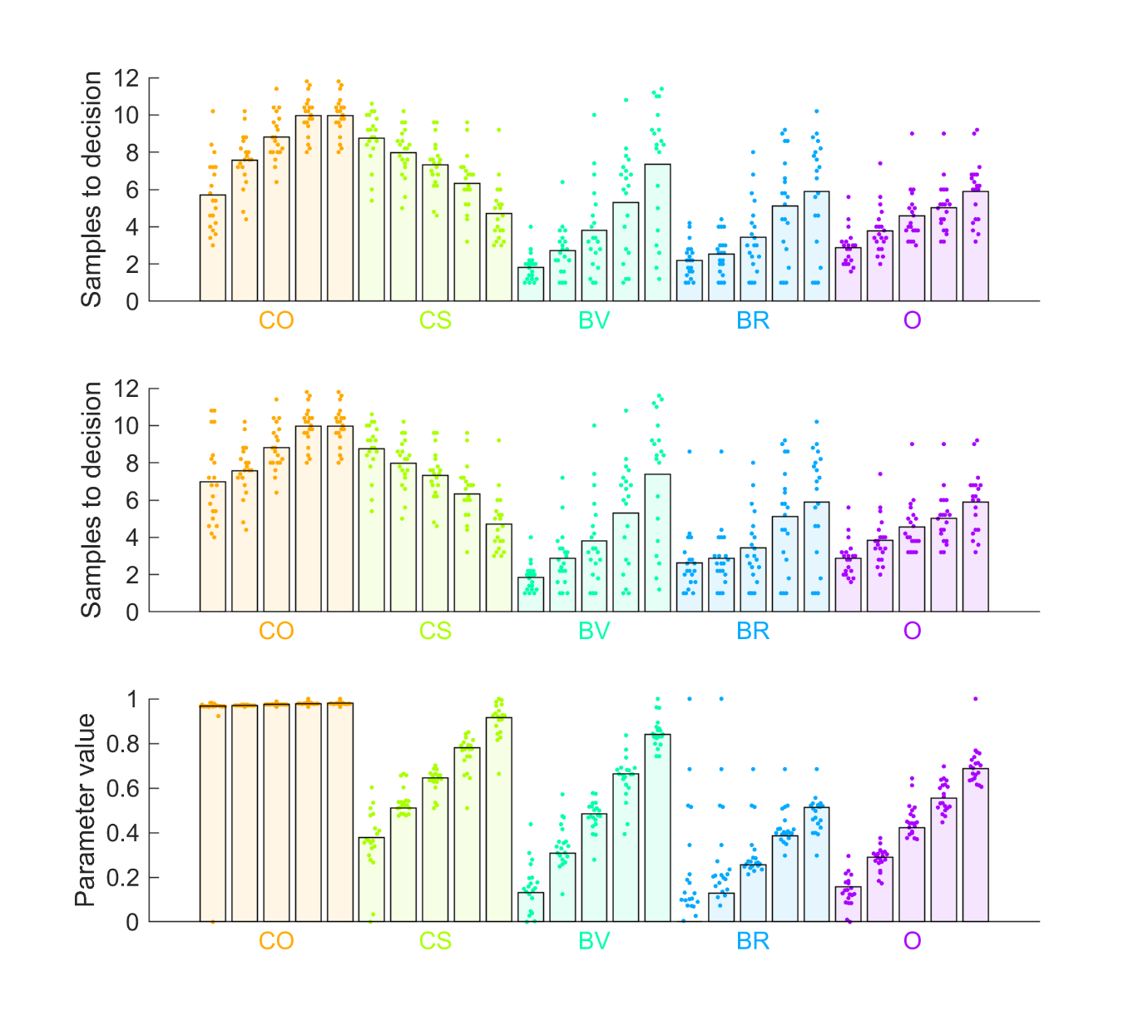


Figure S2. (Top panel) Mean sampling rates (bars) for simulated participants (points), for each configured parameter value for each model. Note that varying configured parameter values leads to systematic increase or decrease in simulated sampling rates. (Middle panel) Models were fitted to the data in the top panel, estimated parameters recovered, and then here we plot mean sampling rates (bars) produced from those recovered parameters in individual participants (points). The sampling rates of the fitted models closely approximate the sampling rates of the original configured models. (Lower panel) The points show how recovered / estimated parameters for individual participants cluster around their corresponding configured parameter values (bars). Note that, although we’ve normalised parameter values to be between 0 and 1 to facilitate plotting of all model parameters on the same scale, nevertheless, the CO model has one extreme outlying estimate parameter value that artificially widens its scale. Abbreviations: CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism.

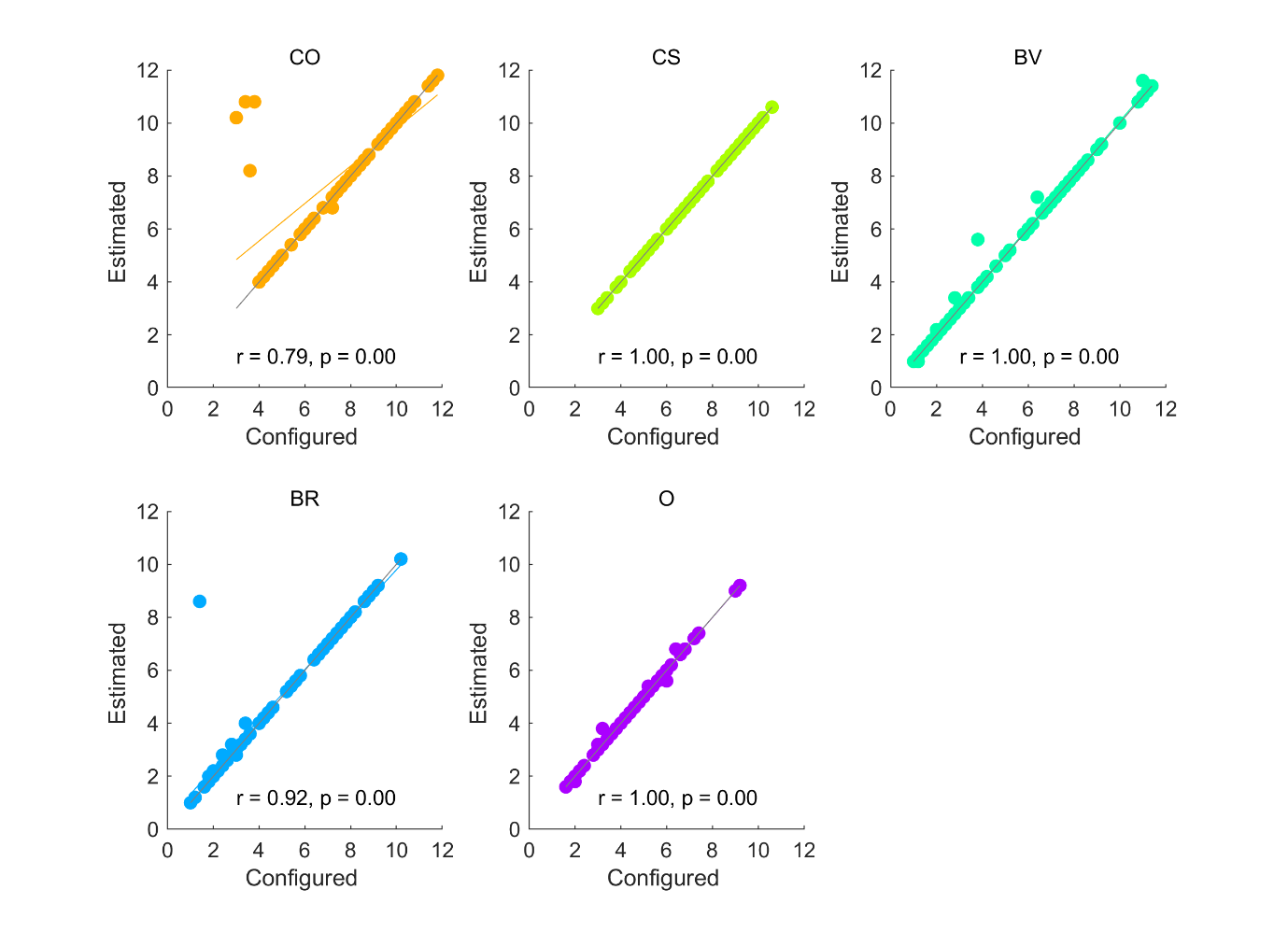


Figure S3. Sampling rates simulated using configured parameters (horizontal axis) are plotted against sampling rates computed from recovered (estimated) parameters. Recovered parameters are highly suitable for reproducing the sampling choices that they are intended to model. The grey diagonal indicates 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: CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism.

Text B: Attention check

Multiple 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. Participants who failed > 25% of the attention checks were excluded from analysis. Four participants failed the attention check in Pilot full, two failed in Study 1 full, and six failed in Study 1 ratings. These participants were subsequently excluded from analysis.

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 items. 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, CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values, SV = subjective values.

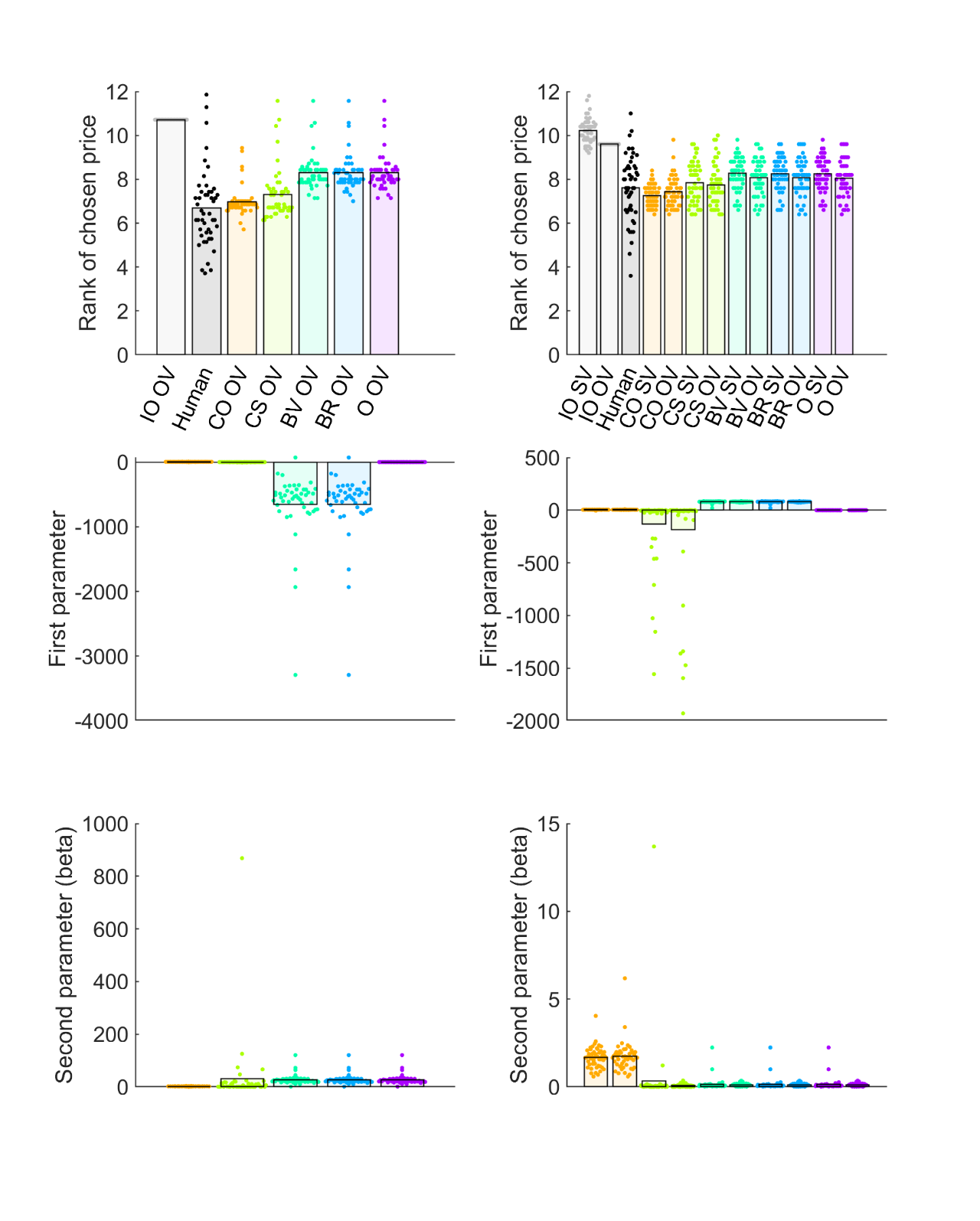


Figure S5. Linear relationships between human participant sampling in Pilot baseline versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured line represents the regression line, with corresponding R2 printed on plot. The CS model predicts the human data with the highest accuracy. Abbreviations: CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values.

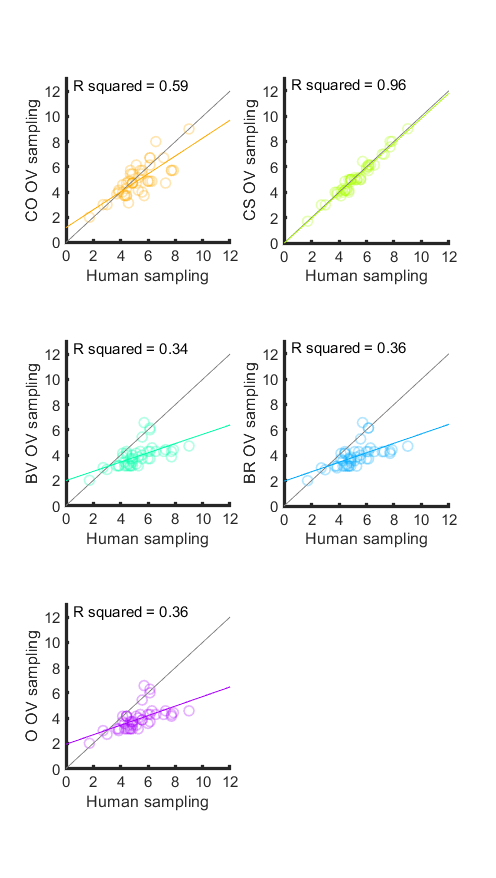


Figure S6. Linear relationships between human participant sampling in Pilot full versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured line represents the regression line, with corresponding R2 printed on plot. The CO and CS models predict the human data with the highest accuracy. Abbreviations: CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values, SV = subjective values.

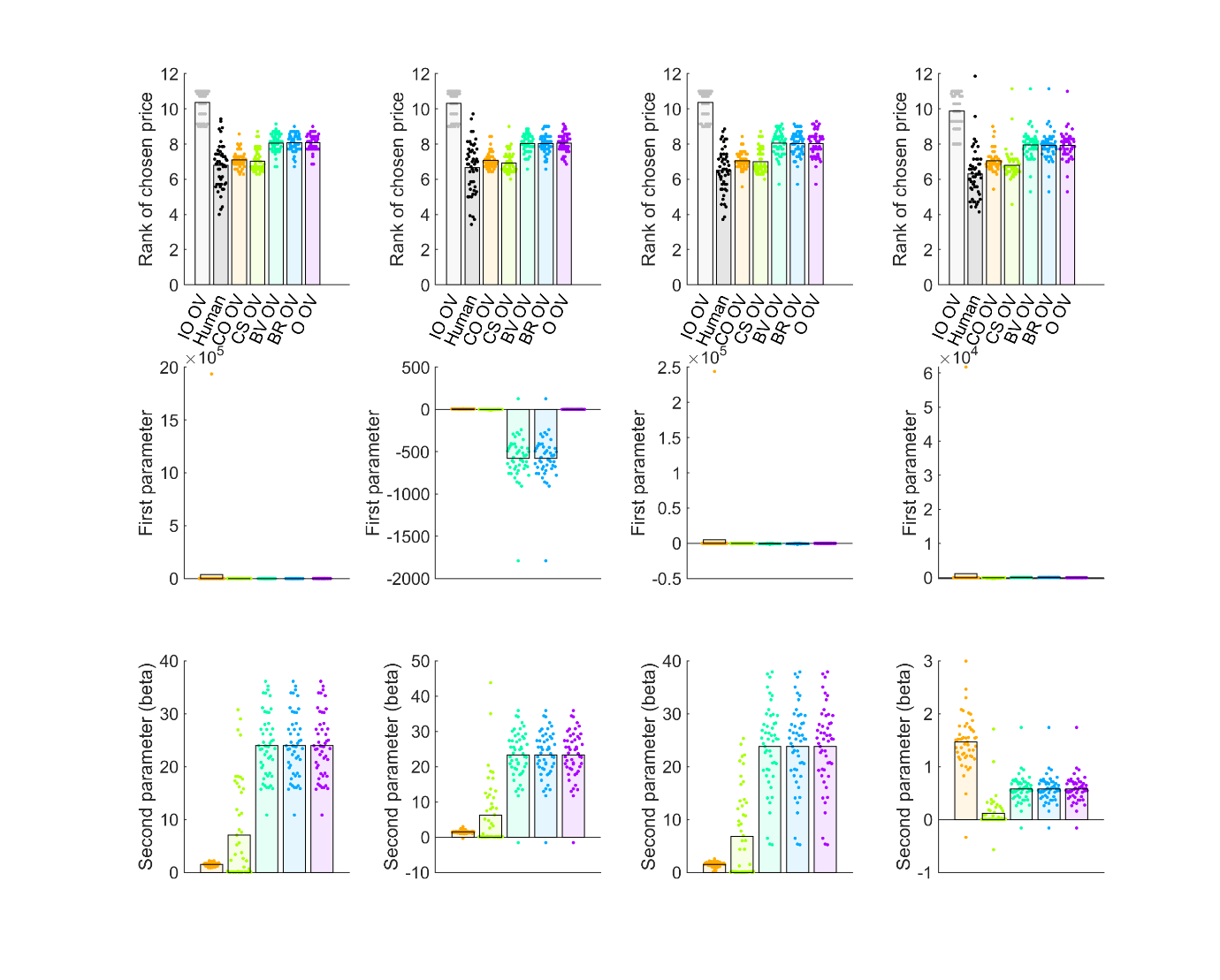
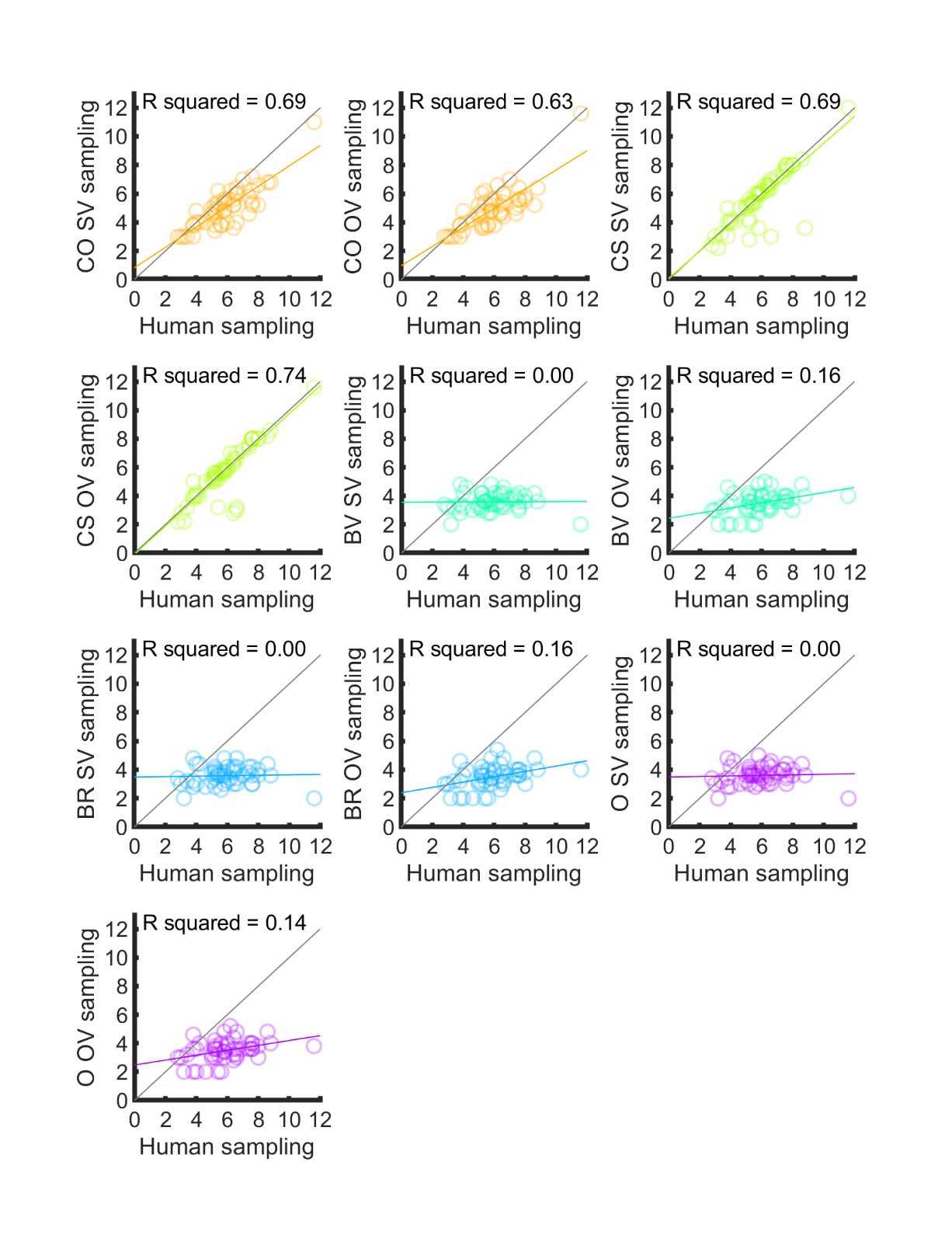


Figure S7. Model comparison for (columns from left to right): Study 1 baseline, squares, timing and payoff conditions. Top and middle rows show individual participants as points and bars show their mean values. The top row shows ranks of chosen prices. The second row plots the “first” or theoretical 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, CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values, SV = subjective values.

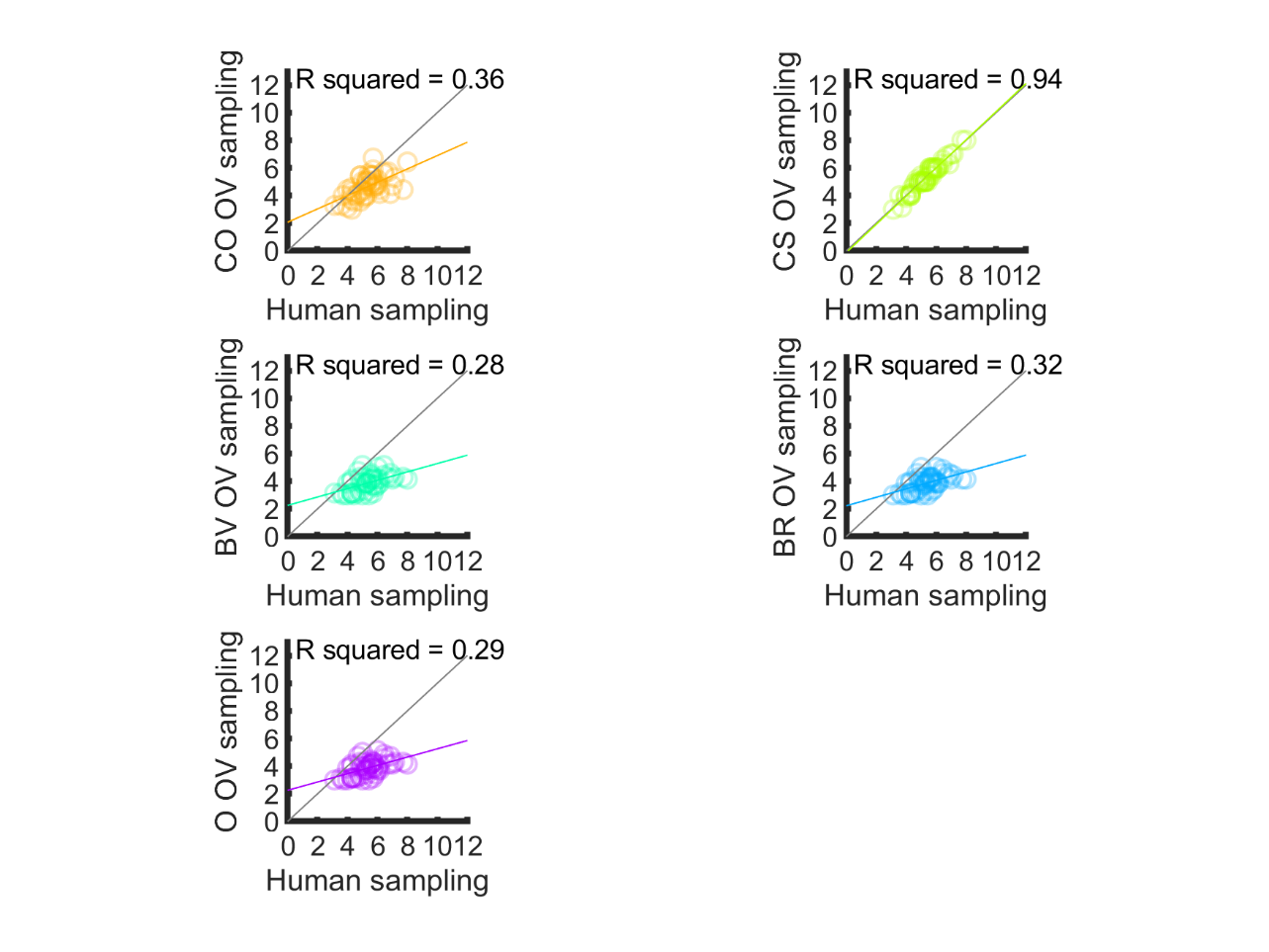


Figure S8. Linear relationships between human participants’ sampling in Study 1 baseline versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured line represents the regression line, with corresponding R2 printed on plot. The CS model predicts the human data with the highest accuracy. Abbreviations: CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values, SV = subjective values.

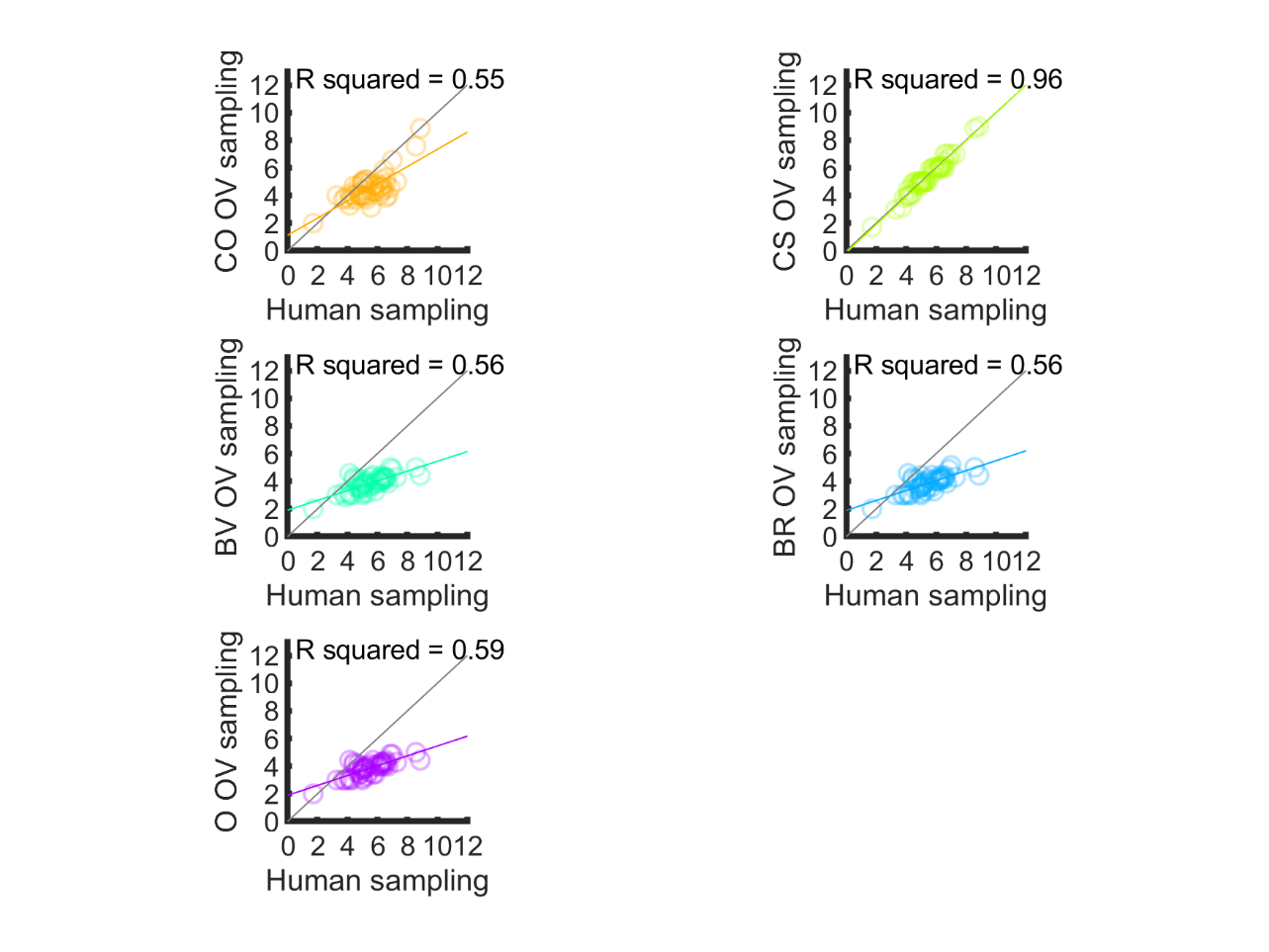


Figure S9. Linear relationships between human participants’ sampling in Study 1 squares condition versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured line represents the regression line, with corresponding R2 printed on plot. The CS model predicts the human data with the highest accuracy. Abbreviations: CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values.

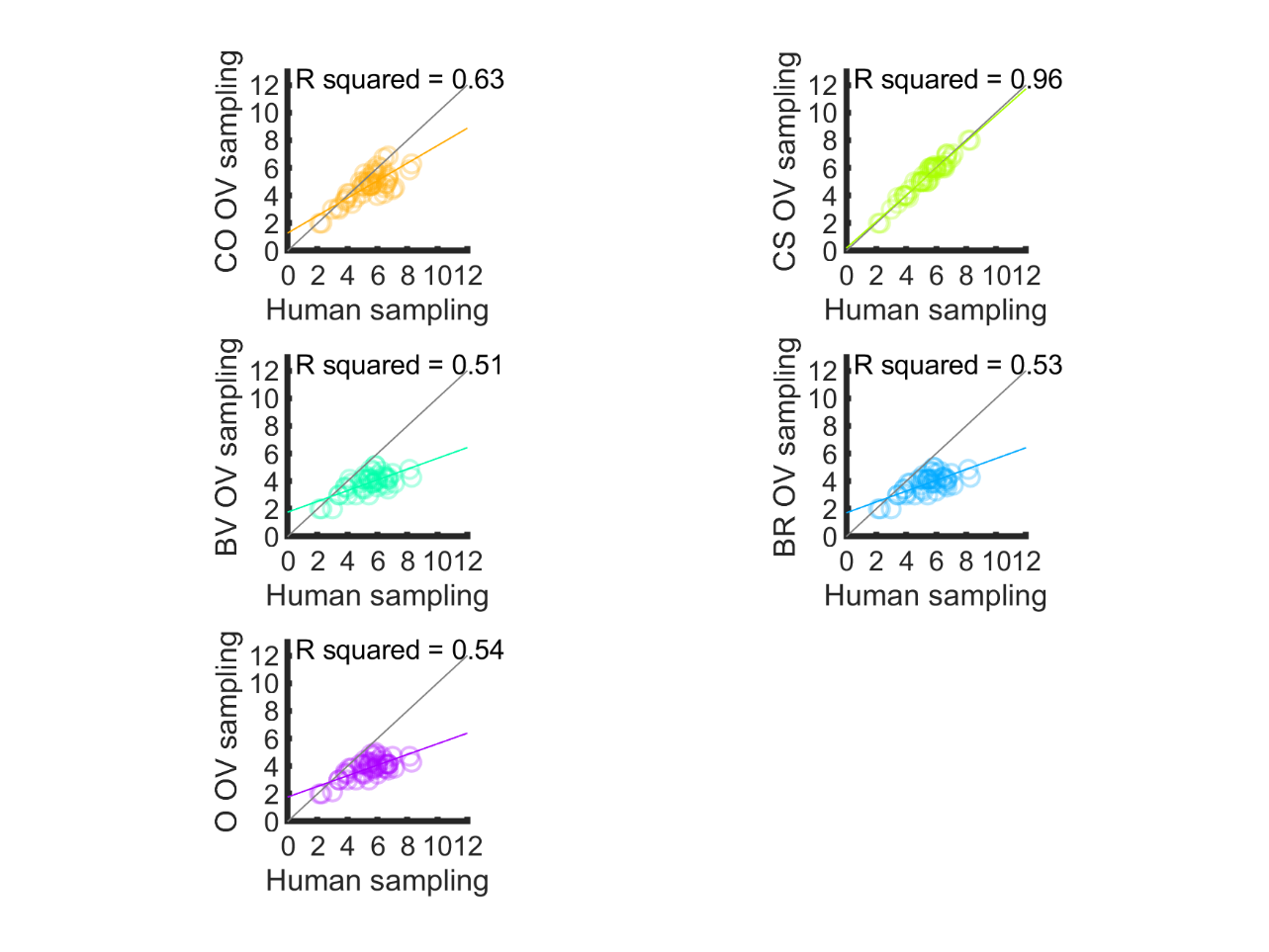


Figure S10. Linear relationships between human participants’ sampling in the Study 1 timing condition versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured line represents the regression line, with corresponding R2 printed on plot. The CS model predicts the human data with the highest accuracy. Abbreviations: CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values.

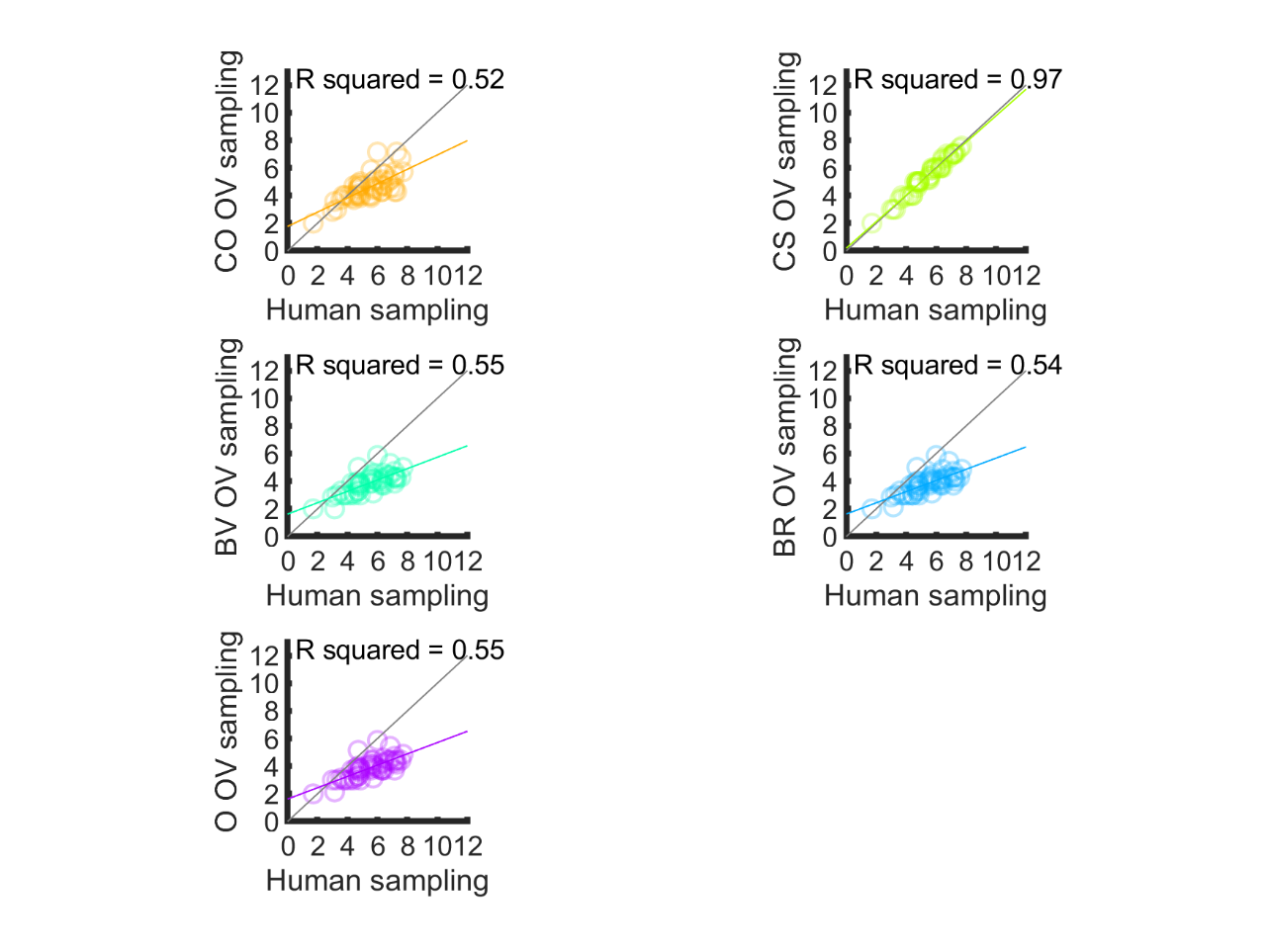


Figure S11. Linear relationships between human participants’ sampling in the Study 1 payoff condition versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured line represents the regression line, with corresponding R2 printed on plot. The CS model predicts the human data with the highest accuracy. Abbreviations: CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values.

Figure S12. Model comparison for Study 1 full (left column) and ratings (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, CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values, SV = subjective values.

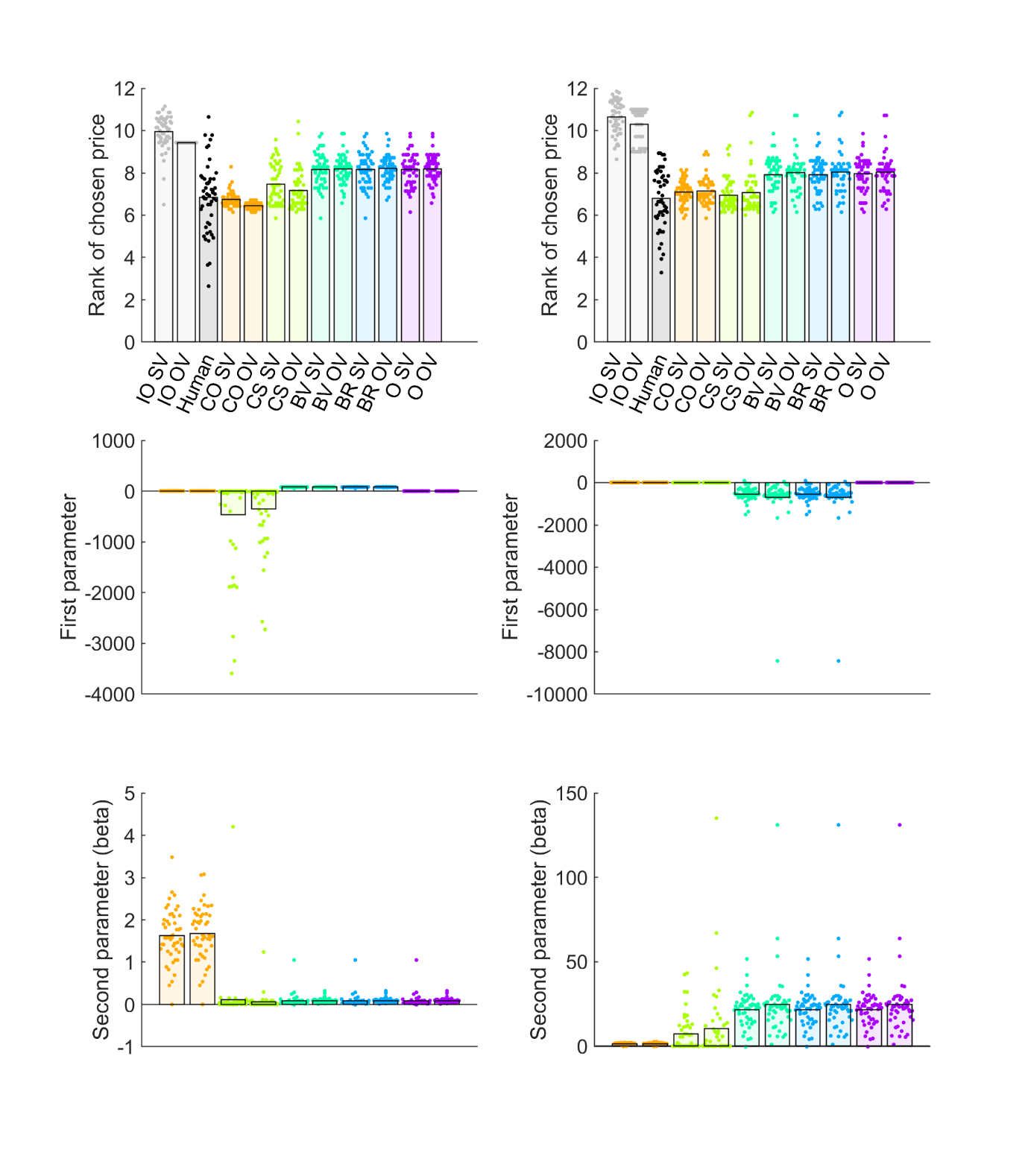




Figure S13. Linear relationships between human participants’ sampling in the Study 1 rating condition versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured line represents the regression line, with corresponding R2 printed on plot. The CS model predicts the human data with the highest accuracy. Abbreviations: CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values, SV = subjective values.

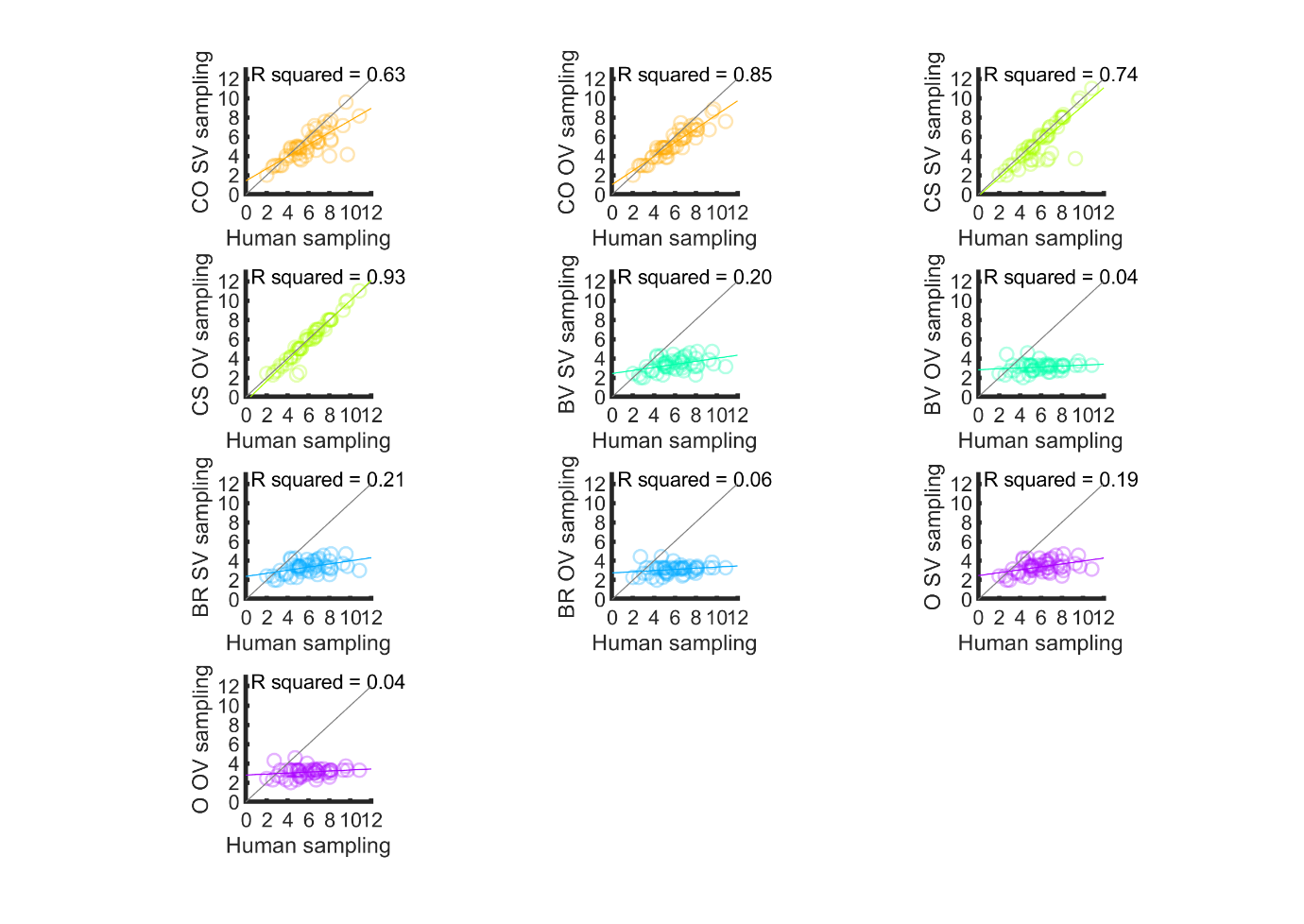


Figure S14. Linear relationships between human participants’ sampling in the Study 1 full condition versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured line represents the regression line, with corresponding R2 printed on plot. The CS model predicts the human data with the highest accuracy. Abbreviations: CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values, SV = subjective values.

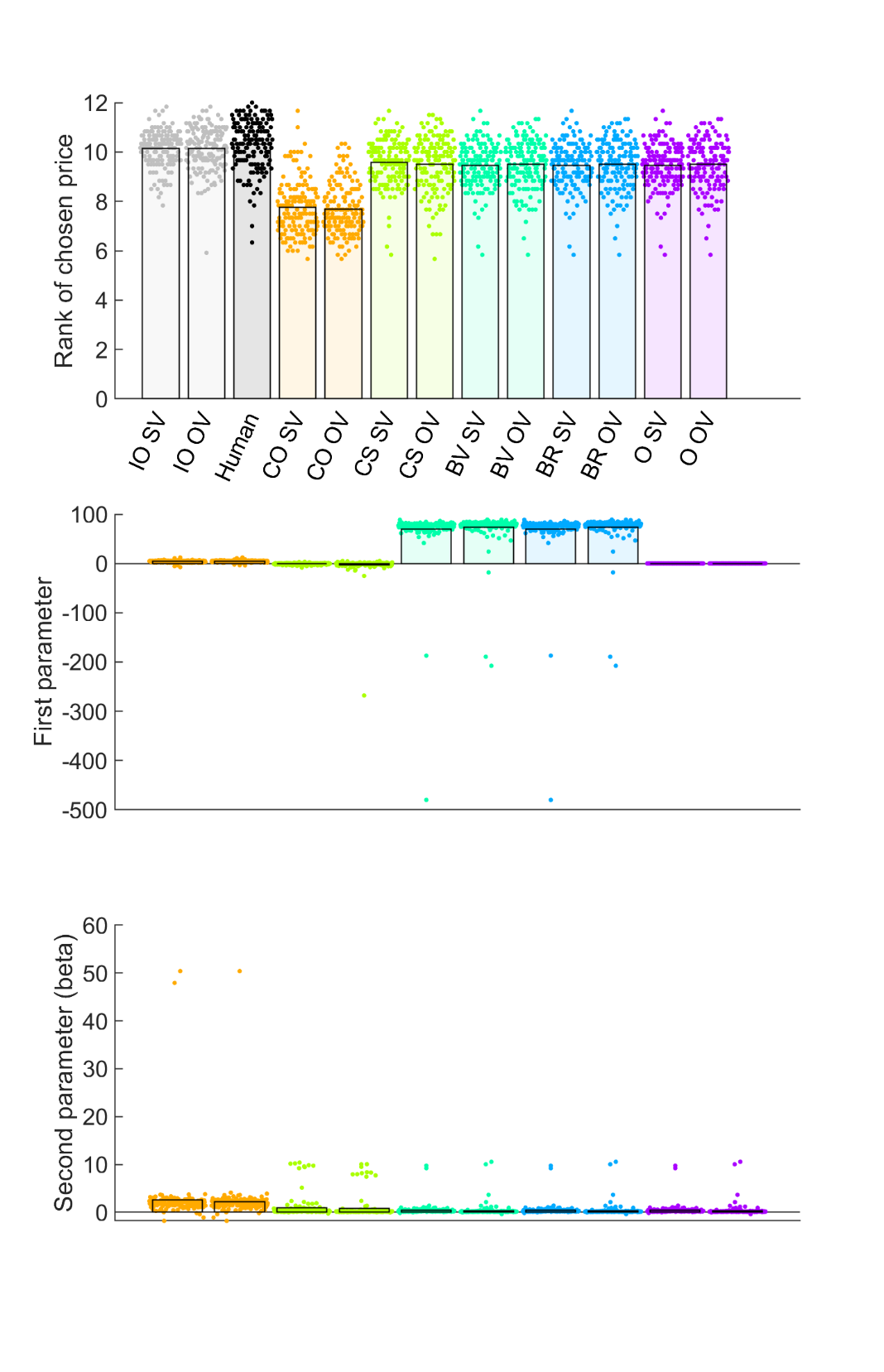


Figure S15. Model comparison for Study 2. Top and middle rows show individual participants as points and bars show their mean values. The top row shows ranks of chosen prices. The second row plots the “first” or theoretical parameter values, estimated for each fitted model. The third row shows the “second”, or inverse temperature parameter beta, estimated for each fitted model. Abbreviations: IO = ideal observer, CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values, SV = subjective values.

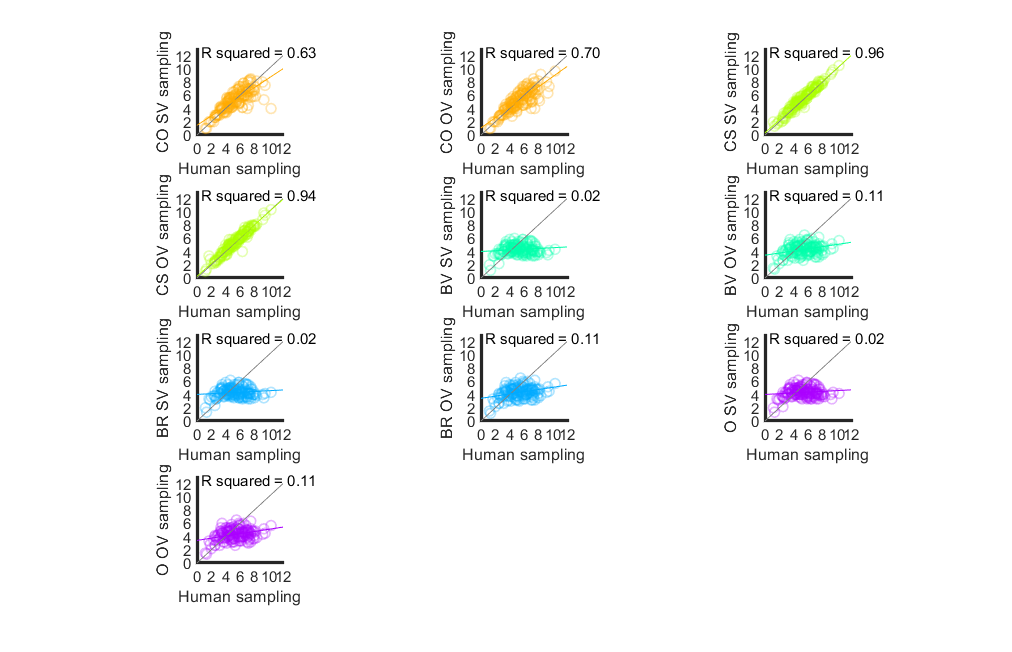


Figure S16. Linear relationships between human participants’ sampling in Study 2, versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured line represents the regression line, with corresponding R2 printed on plot. The CS model predicts the human data with the highest accuracy. Abbreviations: CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values, SV = subjective values.

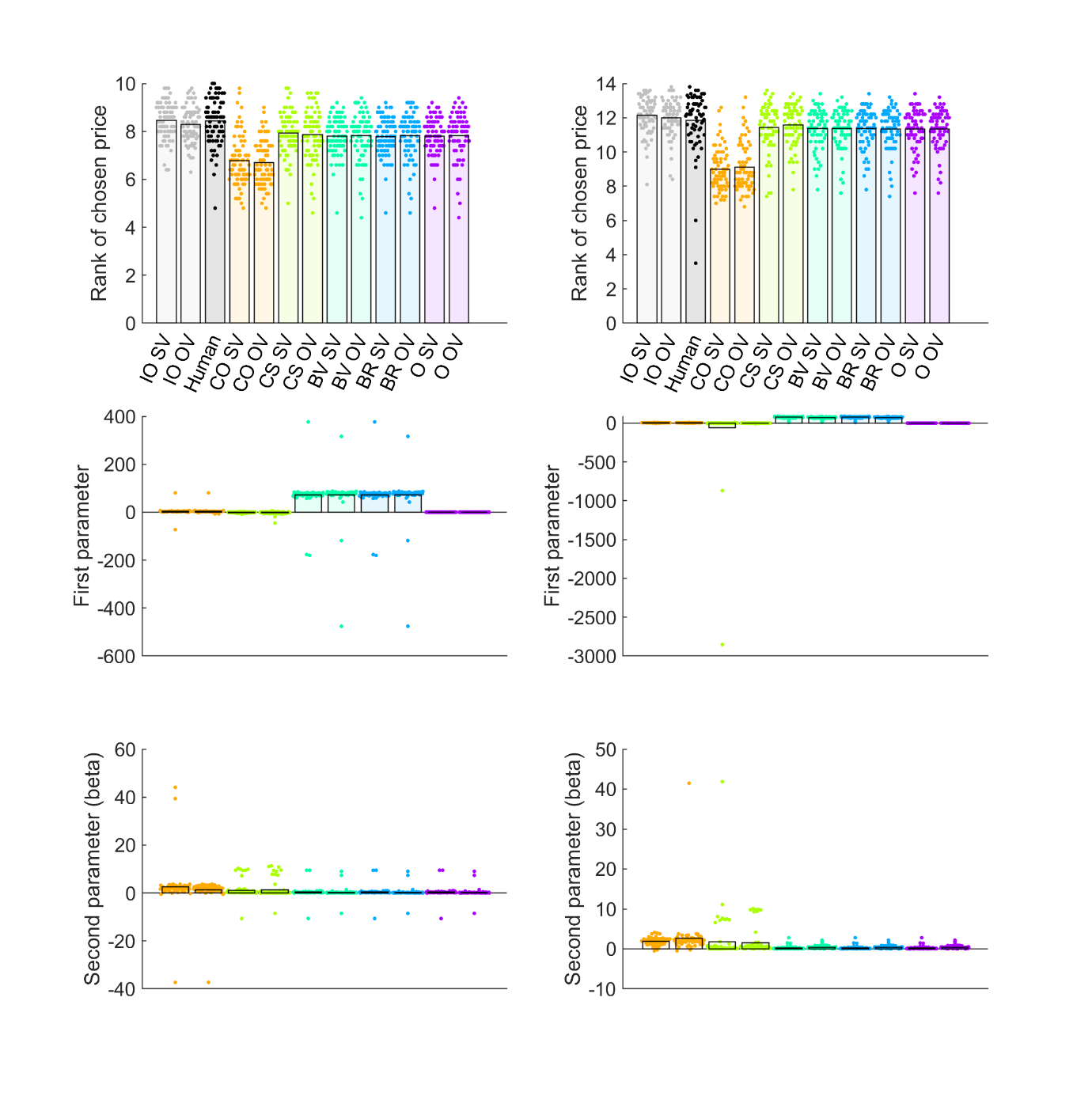


Figure S17. Model comparison for Study 3 10 options (left column) and 14 options (right column) conditions. Top and middle rows show individual participants as points and bars show their mean values. The top row shows ranks of chosen prices. The second row plots the “first” or theoretical parameter values, 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, CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values, SV = subjective values.

Figure S18. Linear relationships between human participants’ sampling in Study 3 10 options condition, versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured line represents the regression line, with corresponding R2 printed on plot. The CS model predicts the human data with the highest accuracy. Abbreviations: CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values, SV = subjective values.

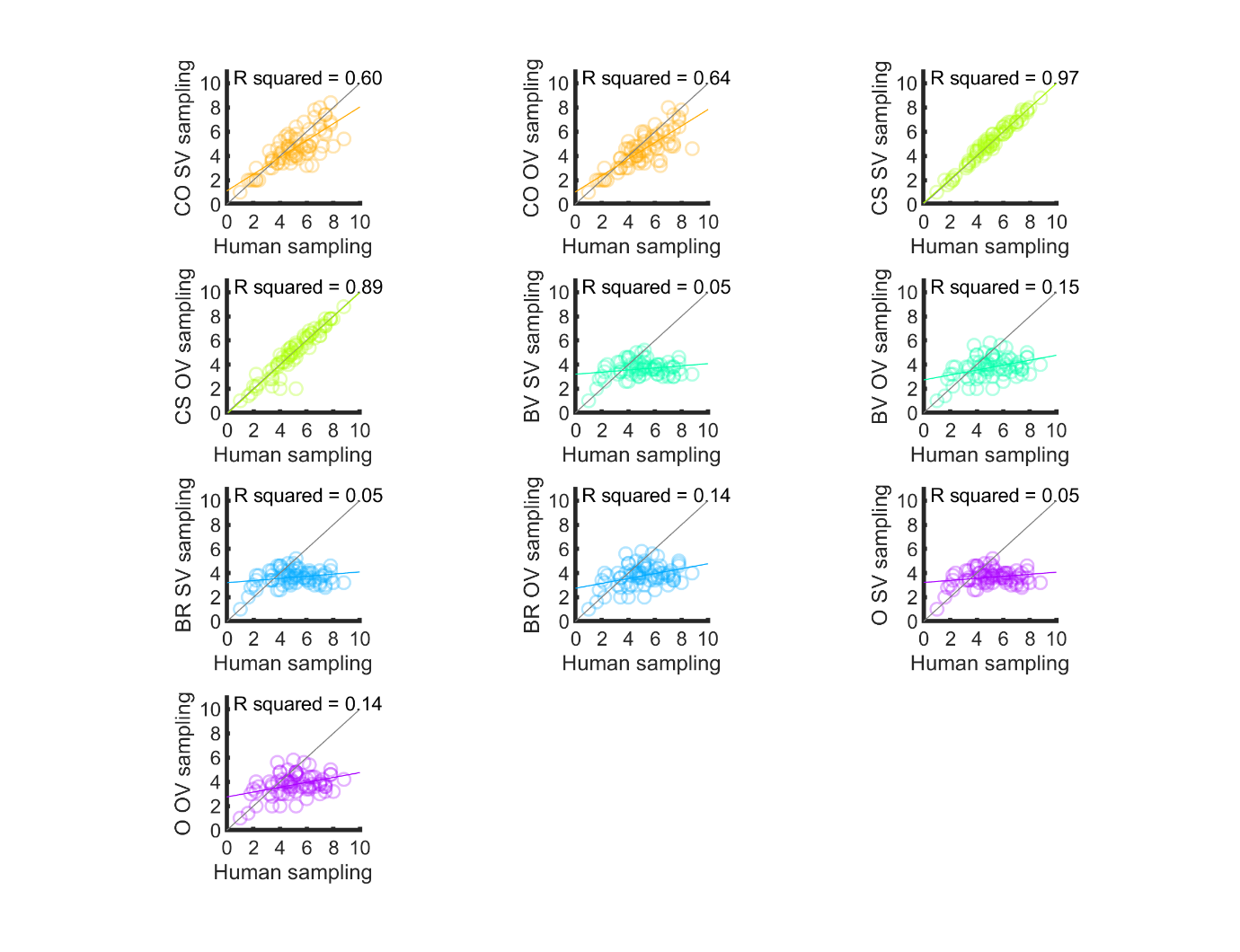


Figure S19. Linear relationships between human participants’ sampling in Study 3 14 options condition, versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured line represents the regression line, with corresponding R2 printed on plot. The CS model predicts the human data with the highest accuracy. Abbreviations: CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values, SV = subjective values.

