Biased expectations about future choice options predict sequential economic decisions

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”, where options are encountered sequentially, and there is no opportunity to recall rejected options or to know upcoming options in advance (e.g., when flat hunting or choosing a spouse). Here, we use computational modelling to identify the mechanisms that best explain decision bias in the context of an especially realistic version of this problem: the full-information problem. We show that participants’ bias - the extent to which their sampling rates deviate from an optimality model - depends on the sequence length and the nature of the payoff scheme. Meanwhile, we rule out a variety of other hypothetical sources of bias. Our comparison of several computational models of bias demonstrated that bias most often arises from inaccurate expectations of the quality of future options (i.e., a mis-specified prior distribution). From these results, we propose a new theoretical viewpoint for the human solution to full information problems. Understanding the causes of decision errors could enhance how we conduct real world sequential searches for options, for example how online shopping or dating applications present options to users.

Introduction

Often in everyday life, decisions must be made regarding options presented in sequence, like when attempting to find the best deal on a certain product or service. For such scenarios we can ask ourselves, when should someone stop evaluating new information and commit to a decision? This common real-life dilemma can be defined as an optimal stopping problem. For example, if one encounters a limited-time offer whilst shopping, should one accept it when it is available or miss it and wait for a better one? If a doctor needs a healthy organ for transplant, do they use what is available now or risk waiting for a healthier one? If an animal welfare charity is visiting homes to find a suitable environment to rehome an animal, do they accept the currently-visited home or continue to visit homes in hope of a better one? This general problem is often referred to as the "fiancé(e) problem", by analogy to decisions about whether to reject a current suitor in favour of meeting new ones in the future. We shall see below that solving many of these problems optimally is computationally challenging and that participants (when compared to the optimal solution) can under some circumstances show systematic decision biases. Our aim here is to delineate the experimental contexts in which participants exhibit these biases and to fit theoretical models to participants’ choices to identify the computational mechanisms that give rise to these biases.

There are many types of optimal stopping problems and their potential computational solutions have been discussed in the fields of mathematics (Ferguson, 1989), behavioural ecology (Castellano et al., 2012; Castellano & Cermelli, 2011), economic decision making (Baumann et al., 2020; Seale & Rapoport, 1997), cognitive science (Lee, 2006) and neuroscience (Costa & Averbeck, 2015). The computational solutions considered for optimal stopping problems are closely related to probabilistic reasoning and explore/exploit foraging decisions (Averbeck, 2015) and other sequential tasks that involve prospective reward prediction (Kolling et al., 2018; Scholl et al., 2022). The availability of optimal computational solutions to these decision problems enables researchers to use them as “ideal observer models”, which can identify when people make suboptimal decisions, including decisions that reveal systematic biases.

We focus in the present study on a bias that arises for *“full information problems”*. This version of optimal stopping problem arguably most closely resembles real-world decision problems. Imagine an agent is searching for a new flat in a competitive market. The agent can sample a limited number of options in sequence (e.g., twelve flats can be viewed, one at a time) and decides, for each option, whether to stop sampling and choose that option, under the condition that rejected options cannot be returned to later (e.g., refused flats are then offered to others and so become unavailable). Full information problems involve more realistic complexities. Flat hunters in full information problems directly know the value of each option (e.g., how nice the currently-viewed flat is or how much it costs). Full information problems can incorporate flexible payoff schemes (e.g., an agent might feel rewarded only if they achieve the best possible flat or their subjective reward might depend on the relative quality of whatever flat is chosen). Full information problems may involve a “cost to sample”. Each time a new flat is visited (i.e., a new option is sampled), our flat hunter may incur calculable costs such as time, money or effort, which may be subtracted from the final achieved reward value and so can limit how many options are sampled. Finally, full information problems allow agents to harness their prior belief about the probability distribution that is generating their decision options (i.e., the generating distribution). When flat hunting, consumers can use these *prior expectations* about the housing market to prospectively compute the probability that an even nicer flat might be sampled if the current one is refused.

Here, we will use experimental methods and computational modelling to test a raft of hypotheses related to an “undersampling bias”. When the sampling behaviour of ideal observers is compared to that of human participants, humans sample fewer options than is optimal (e.g., Baumann et al., 2020; Cardinale et al., 2021; Costa & Averbeck, 2015; Goldstein et al., 2020; Guan & Stokes, 2020). To date, this undersampling bias has mainly been demonstrated for optimal stopping problems cast in economic scenarios in which options are represented as numbers (e.g., prices). Here, we have adapted the economic task first reported by Costa and Averbeck (2015). In our version, participants attempt to choose high-ranking smart phone prices. However, undersampling bias is by no means universal. For example, some new studies have reported full information problems associated with oversampling rather than undersampling (Furl et al., 2019; van de Wouw et al., 2022). These studies employed several different experimental and computational methods that might have ameliorated the undersampling bias shown in other studies. Herein, we systematically manipulated each of these methods, demonstrating that undersampling bias is increased for longer sequence lengths and for payoff schemes that reward only top-ranking choices, while ruling out the other methods as potential sources of bias.

What computational mechanisms account for participants’ errors on this task? We created theoretical computational models, each with a free “bias” parameter that skews otherwise optimal performance. We show (replicated across multiple studies and conditions) that participants’ sampling decisions on our economic task are best explained by a theoretical model that makes inaccurate expectations of the quality of upcoming options, based on a mis-specified belief about the prior option distribution.

General Methods

Paradigm summary

First, we briefly describe the features of our paradigms that are relevant for understanding the operations of our computational 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 price of a mobile phone contract that they most preferred. Prices used for 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.

In some conditions, the paradigm began with a “phase 1” ratings task, in which participants viewed the full distribution of prices that could appear as options later and rated each for its “attractiveness” or subjective value. As described below, some models operate over objective values / raw prices (OV) and other models operate on the subjective value of the prices (SV), derived from the ratings measured during phase 1. In phase1, participants also could learn the “generating” distribution of option values and thereby establish expectations about the probabilities with which certain option values might appear in any given sequence, later in the optimal stopping task. The distribution of these ratings could also be used to set the models’ prior on its generating distribution of option values (See *Ideal observer optimality model* section).

Next, in the optimal stopping task, participants engaged with several fixed-length sequences of option values, populated by prices sampled randomly, without replacement, from the Phase 1 generating distribution. 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, 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 becomes their choice by default.

Ideal observer optimality model

To analyse the optimal stopping task, we compared the number of options sampled by our participants before choosing an option to that of an ideal observer model. The ideal observer model is a benchmark of optimality, for which performance is Bayes-optimal. This finite-horizon, discrete-time, Markov decision process (MDP) model has been used in previous studies (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). Expectations about 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 reduces r by costs incurred by sampling again using a “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.

The utilities for sampling again are computed based on backwards induction. 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*. Whenever the reward value of taking the current option is considered, the reward 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*.

This customisable *R* function allowed us to examine how the ideal observer model changes its sampling strategy under the different reward payoff schemes used in our studies. In Pilot full, the full condition of Study 1, Study 2 and both sequence length conditions of Study 3 (These studies and their associated experimental conditions will all be described in depth in later sections), participants were instructed to try to choose the best price possible. In study conditions using these instructions, we implemented a continuous payoff function (resembling that of the classic Gilbert & Mosteller formulation), in which the relative rank of each choice 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 programmed the reward payoff function to match participants’ reward of 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 Gilbert & 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. Our Bayesian version of the optimality model 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: Ideal observer objective values (IO OV) and ideal observer subjective values (IO SV). 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). We implemented the IO OV procedure for all the study conditions reported herein. This OV procedure assumes that the raw prices can be treated as a proxy for participants’ subjective value of the prices, so an IO model that optimises only the raw prices when making decisions would therefore be an appropriate basis for comparison with participants. However, we also had direct access to subjective values of options in some conditions, due to the presence of the initial rating phase. The IO SV model assumes that participants’ subjective valuation of prices may not necessarily exactly equal the raw price values, especially in their scaling, which may be relevant to full information problems. For conditions which had a ratings phase, 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 (Pilot full, Study 1 full condition, Study 1 ratings condition, Study 2 and both sequence length conditions of Study 3), 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 test the hypothesis that the use of objective or subjective values when modelling affects the strategy taken by the optimality model and, consequently, whether it changes the assessment of participant bias. We ensured for both OV and SV models that better options were always more positively-valued such that the models were always solving a maximisation problem. We further ensured that estimated parameters for OV and SV models would be on the same scales. We achieved this by reflecting the prices around their mean. Then we rescaled the values to span 1 (the highest / worst price) to 100 (the best price). These reflected and rescaled objective values were then used in OV models when computing the prior generating distribution (subjective value ratings were already made by participants on this 1 to 100 scale), and when inputting price values to the model as option values,

Theoretical models

The purpose of the ideal observer model described above was to assess bias, not to theoretically explain participants’ bias. By definition, the parameter values of an ideal observer model are fixed to ground truths established by the experimental design. Because of this feature, however, ideal observer models are not appropriate for use as theoretical models of potentially-biased human sampling and choice behaviour, without modification added to account for 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. To better understand which computations might be responsible for participants’ errors, we formulated a number of theoretical models and fitted them to participants’ 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., 2022). 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, whenever a study condition involved 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 (which 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 action values from each model to probabilities using Softmax and then summing negative log likelihoods over choices for each participant. In each model, we freed one theoretically interpretable key parameter (these free parameters and their models are described below) and the inverse temperature parameter beta from the Softmax function (the starting value for beta was always 1 and the fitting of beta was bounded between 0 and 200). Variability in each of the key theoretical parameters was confirmed during parameter recovery to be capable of modulating the sampling rate (Supplementary Procedures Text A and Supplementary Figure S2 and upper panel of Figure S3). The two free parameters per model were fitted using fminsearch.m in MATLAB (Mathworks, Natick MA). Parameter recovery analyses for three of the models we consider and describe below showed at least adequate correlations between configured and recovered parameters (Figure S1): The cut-off heuristic, the cost to sample model and the biased prior model. These models also showed strong correlations between sampling rates associated with configured parameters and sampling rates associated with recovered parameters (Supplementary Procedures Text A and Supplementary Figures S2 and S3). Two other theoretically motivated models – the biased values and biased rewards models (See Supplementary Materials) – performed more poorly during parameter recovery and so were excluded from the formal model comparison. We implemented two parallel model comparison methods based on negative log likelihood values converted to Bayesian information criterion (BIC) values. For the first model comparison method, 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. The better models show statistically lower BIC mean values. For the second model comparison method, 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 model that best-fit the most participants presumably was the sampling strategy most often used by participants in our sample.

The objective and subjective values versions of the *cut off heuristic (CO OV and CO SV)* derive from the mathematically-optimal solution to the “Secretary problem” (Ferguson, 1989), a distinct optimal stopping problem with a relatively simple mathematical solution, due to an abundance of required assumptions that need not hold for full-information problems. Namely, the secretary problem solution assumes the agent uses no prior knowledge of the generating distribution, considers only relative ranks of option values and feels 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 nevertheless 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, however, that the CO model has no formal cost to sample parameter). This heuristic has previously 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). The CO heuristic chooses to sample again for every option until it reaches a cut-off sequence position, which is fitted as the key theoretical free parameter. Then, the model continues to sample until it reaches the next option with the highest 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 and the parameter search was bounded between 2 and the sequence length minus 1 (as the learning period defined by the cut-off must contain at least one sample and be followed by at least one sample available for choice). Cut-off values below the optimal value lead to undersampling and cut-off values above the optimal value lead to oversampling.

We also considered objective and subjective values versions of *the cost to sample model (CS OV and CS SV)*. CS OV and CS SV use the Bayesian ideal observer described above as a base, while also assuming that participants’ otherwise rational Bayesian computations can be biased by a 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 the Ideal Observer Optimality Model section above. In such a model, participants would undersample if they intrinsically perceive sampling as costly and so adopt a negatively-valued *C,* and would oversample if they perceive sampling as rewarding as so adopt a positive *C*. We initialised model fitting with a starting *C* value of 0 (i.e., the optimal value) and, during fitting, bounded *C* to be between -100 and 100.

We used a similar approach when building *the biased prior model (BO OV and BP SV)*. In this model, we added a new free parameter to , the mean of the prior generating distribution. Negative values of this parameter can bias an agent to compute pessimistic estimates of future option values by shifting the prior 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) and the biased prior parameter was bounded during fitting to be between -100 and 100.

Pilot Studies Methods

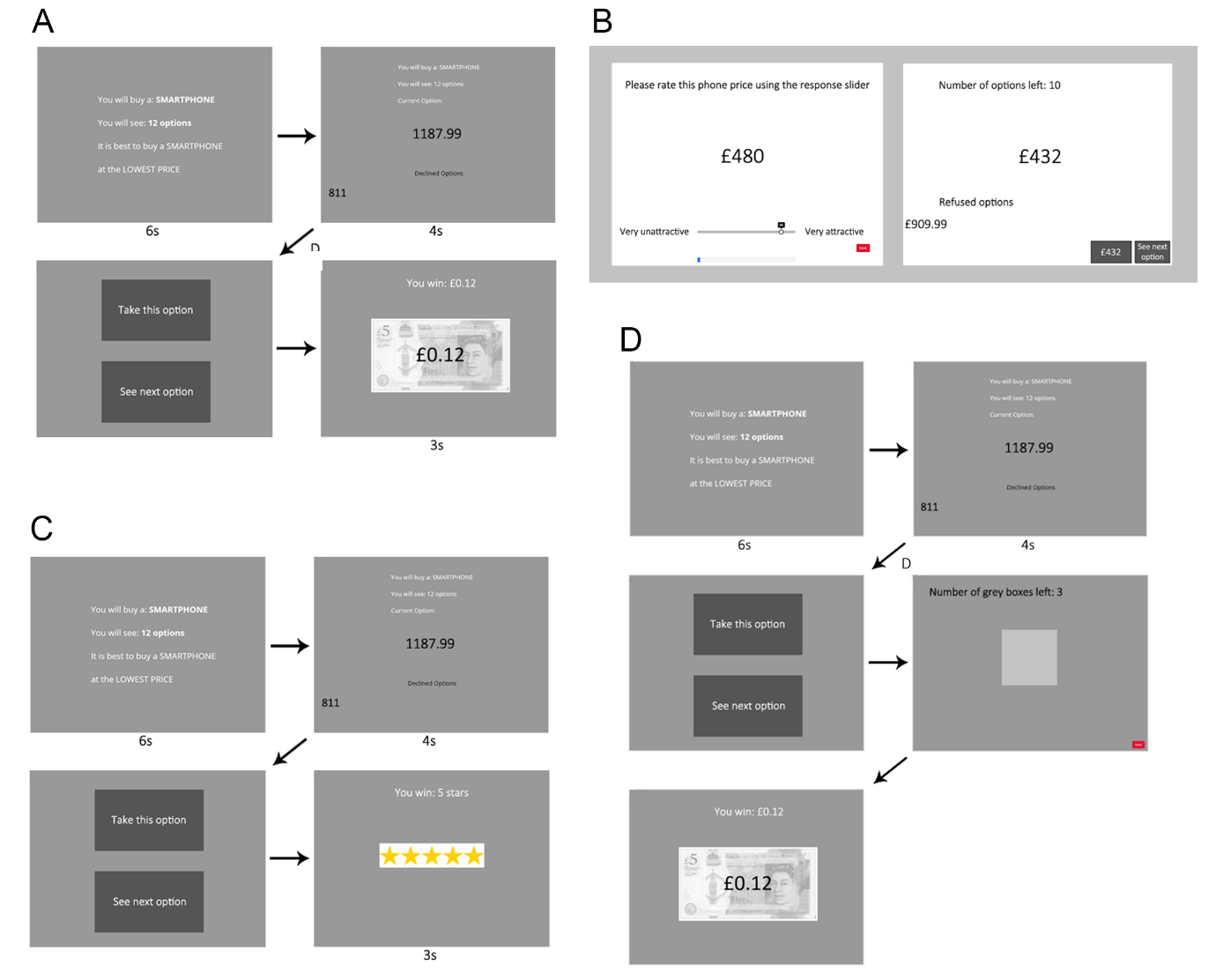
Participants

We recruited participants in both our pilot studies from the United Kingdom using the online data collection platform Prolific (Prolific, 2014). We enrolled 50 participants into Pilot Baseline and 51 participants into Pilot Full.

Procedures

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

Figure 1. Paradigm designs used in pilot studies and Study 1. (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.



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 performance-based bonus 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 per hour (participants typically finished the study in considerably less time than an 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.

For Pilot Full, we were interested in whether participant undersampling bias would continue to replicate using the same economic smartphone price task, but when implementing the “full” complement of methods particulars adapted from studies that revealed oversampling bias instead of undersampling bias (Furl et al., 2019). The logic is that, if any of these methods features is responsible for the oversampling bias seen in these earlier papers, then Pilot full should produce an oversampling bias, which would contrast with the undersampling bias we expected to see in Pilot baseline.

Pilot full 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 only appearing after the first click - to avoid slider biases (Matejka et al., 2016) - with the selected rating value shown above the slider. Participants rated 180 prices, presented one at a time in a random order, and comprising the 90 unique prices, each 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. In Pilot full, the mean (over participants) Pearson’s correlation coefficient between the two ratings was .83. 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 was intended to discourage participants from finishing the study early by choosing earlier options. 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 only to provide feedback to the participants about the quality of their choices, in lieu of the bonus payoff screen in Pilot baseline, and to encourage participants to reflect upon the choice’s 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. Figure 2 shows the mean number of samples to decision made by human participants for both 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 condition means within a study are shown as green horizontal lines (*p* < .05, multiple comparison corrected for the number of pairs in that study), which appear only for the Study 3 sequence length conditions. 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).

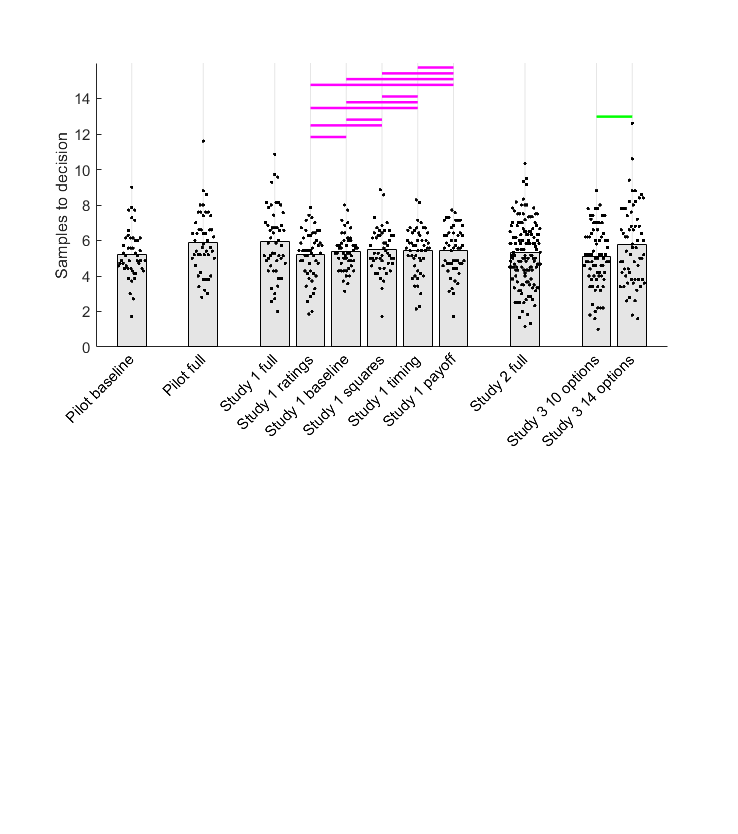


Figure 3. Model comparison for Pilot baseline (left column) and Pilot full (right column). Points in the first and second rows show data corresponding to individual participants, while bars show their mean values. Human participant data are reproduced from Figure 2. In the first row, horizontal lines above human and IO samples data indicate in thin black when *BF01* > 3 (moderate evidence for equal means) or in thick grey when *BF10* > 3 (moderate evidence for different means). The second row shows BIC values (lower values indicate better model fit) for participants (points) and their mean values (bars). Horizontal lines are shown in the colour corresponding to the better-fitting model when *BF10* > 3 or in black when *BF01* > 3. The third row demonstrates that BP models best fitted the most participants in both pilots. Abbreviations: IO = ideal observer, CS = cost to sample, CO = cut-off, BP = biased prior, OV = objective values, SV = subjective values.

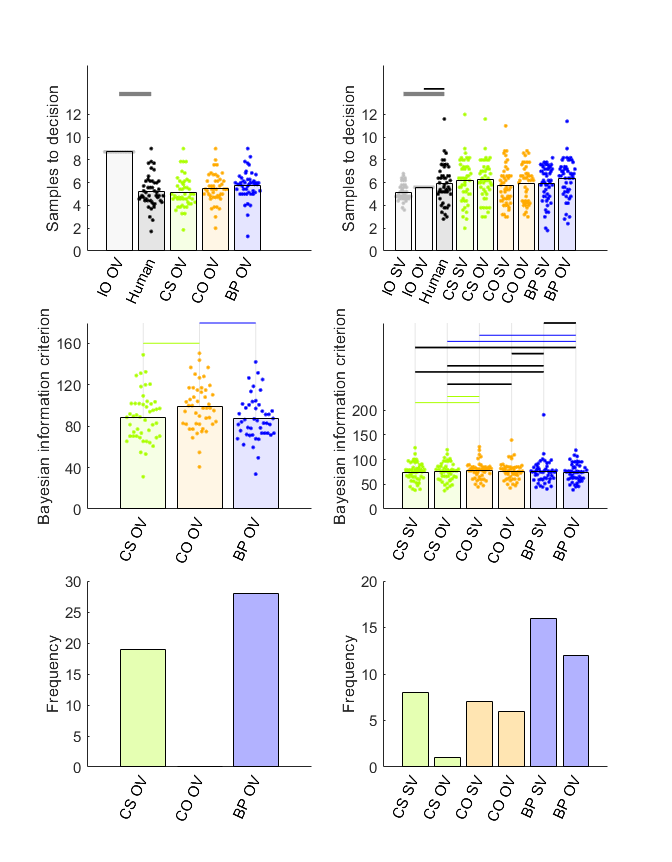


Figure 3 shows results from the comparison of human participants’ sampling with sampling of the ideal observer and theoretical models. As expected, we successfully replicated undersampling in the Pilot baseline condition (Figure 3, upper left), where participants sampled fewer options than the ideal observer (Cohen’s *d* = -2.52). All the theoretical models, after fitting to Pilot baseline data, resembled the participants to some degree, as they all showed some degree of undersampling, compared to IO OV. Bayesian pairwise tests (Figure 3, left column, second row), showed that CS and BP were not statistically distinguishable but both outperformed CO. BP was the best fitting model for most participants, though CS was the best fitting model for a sizable number.

Our hypothesis that some task feature in Pilot full would eliminate the undersampling bias observed in Pilot 1 was fulfilled. This contrast between pilot studies does not seem to arise because participants sampled differently, but rather because the ideal observers reduced sampling in Pilot, compared to Pilot baseline. In Pilot full, participants’ sampling (Figure 3, top right) was statistically equivalent to IO OV sampling (Cohen’s *d* = .17) and even significantly greater that IO SV sampling (Cohen’s *d* = .45). To anticipate, we shall systematically test these methods features in Study 1, where we will see that the IO models reduce their sampling rates when all relative ranks of choices are rewarded depending on the magnitude of the option value (as in Pilot full), and increase their sampling rates when only the top three ranks are rewarded (as in Pilot baseline), while participants are relatively insensitive to the payoff scheme.

What computational mechanisms account for participants’ discrepancies from optimality in these pilot studies? In both studies, statistical tests comparing pairs of participant BIC values give some evidence that BP and CS are both better than CO (Figure 3, middle row), though BP seems to better fit the most participants in both studies with a substantial contribution of CS (Figure 3, lower row). To anticipate, we will see a similar pattern replicated across all of our later studies: The most evidence favours BP as the most common model of participant performance, though there may also be a contribution of CS.

Study 1

The paradigm design that we adapted to use in Pilot baseline was taken from Costa & Averbeck (2015) and we thereby replicated that study’s findings of undersampling. Concomitantly, we adapted many of the design features for Pilot full from studies that instead showed oversampling (Furl et al., 2019, van de Wouw et al., 2022) and we found that Pilot full eliminated the undersampling bias by reducing the sampling rate of the IO optimality model. This pattern raises a distinct possibility – which we test in Study 1 - that a systematic manipulation of each of the task features added to Pilot full will show at least one of them affects IO sampling, though they may not also affect participant sampling. Study 1 will further give us six more datasets with which we can perform model fits and attempt to replicate our findings from the pilot studies that BP is the most commonly fitted explanation of participant performance.

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 GPB. We enrolled independent participant samples into each of six conditions (See Procedures), targeting 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.

Procedures

The study was developed using the experiment 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 1 and Figure 1 visualises the paradigm designs for Study 1 baseline (Figure 1A), full (Figure 1B), payoff (Figure 1C) and squares (Figure 1D) conditions. Next, we will 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. 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 new methodological features from Furl et al. (2019) under test here, and it will serve as the basis for comparison against the other conditions, which each add one or more of the methodological features. Similar to Pilot baseline, we fixed in advance the option values and their order within each of the sequences, and then these fixed-option sequences were presented in random order. However, in this case, to avoid as homogenous a set of sequences as was used in Pilot baseline, we created 10 such fixed sets of sequences and each participant was randomly assigned to one of these sets. This procedure was implemented in Study 1 baseline and in all the conditions based on it, described below (i.e., ratings, payoff, squares, timing). The *full condition* was identical to the Pilot full study (Figure 1B), except that it used seven sequences instead of five. The mean (over participants) Pearson’s correlation coefficient between the two ratings for each price collected in the first phase was .87. 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). In this condition, the correlation between the two ratings for each price (on average over participants) was .81. The *payoff condition* (Figure 1C) 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. 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 best, second best, or third best price in the sequence. The *squares condition* (Figure 1D) 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. 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.

Table 1. Summary of conditions for Study 1

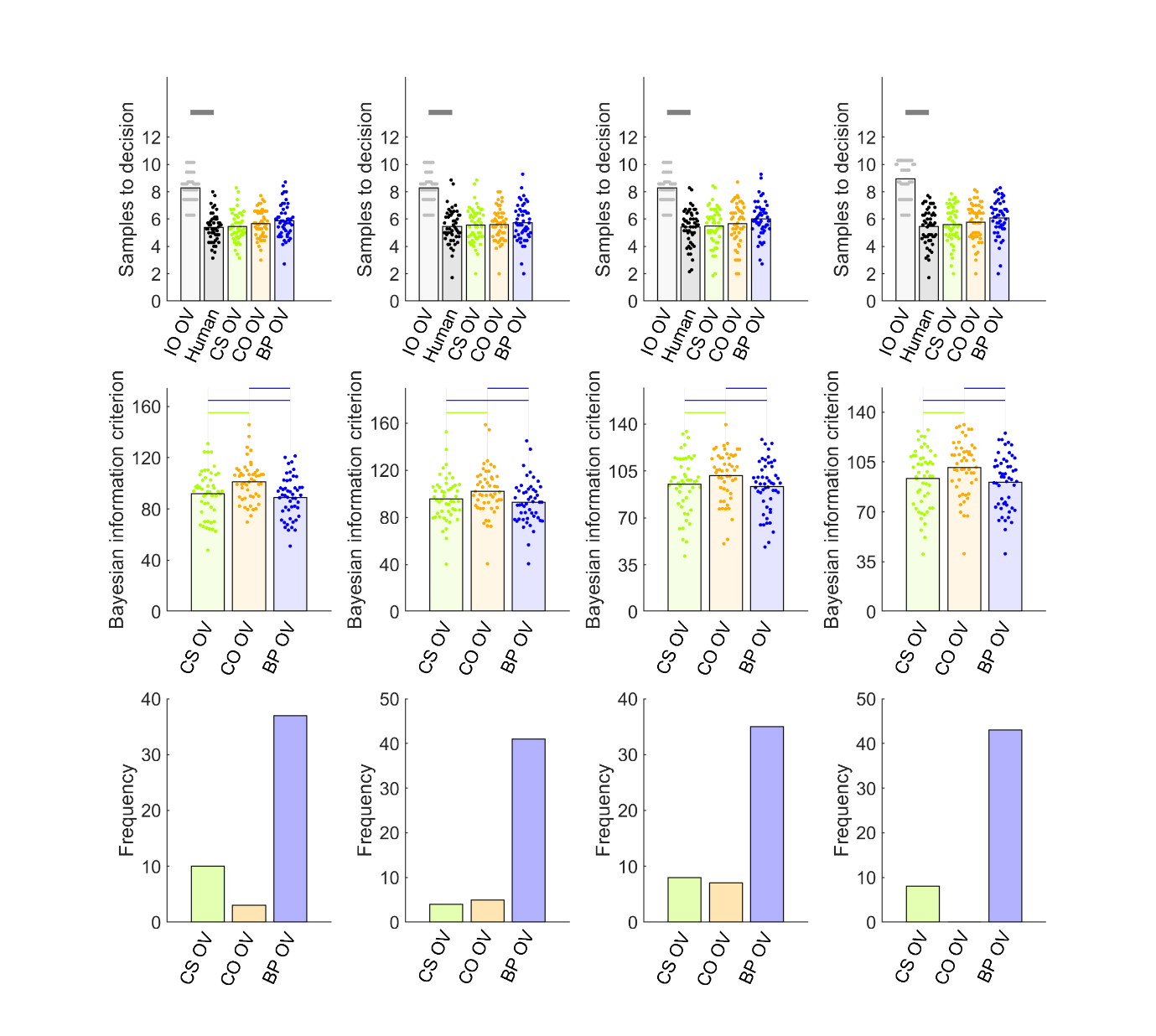
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Study 1 condition name | | | | | |
|  |  | Baseline | Full | Squares | Payoff | Timing | Ratings |
| Task feature | Grey squares |  | × | × |  |  |  |
| No monetary payoff |  | × |  | × |  |  |
| Self-paced timing | 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

Study 1 Results and Discussion

Our hypothesis was confirmed that none 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” statistically-substantiated mean 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). According to these Bayesian *t*-tests, nearly every pair of conditions showed statistically equivalent means, (all *BF01* > 3, moderate evidence in favour of null model and shown as magenta horizontal lines in Figure 2), with the only exceptions being the comparisons with the full condition, which were statistically inconclusive. Cohen’s *d* values for these comparisons are visualised in Figure S5 in the Supplementary Materials.

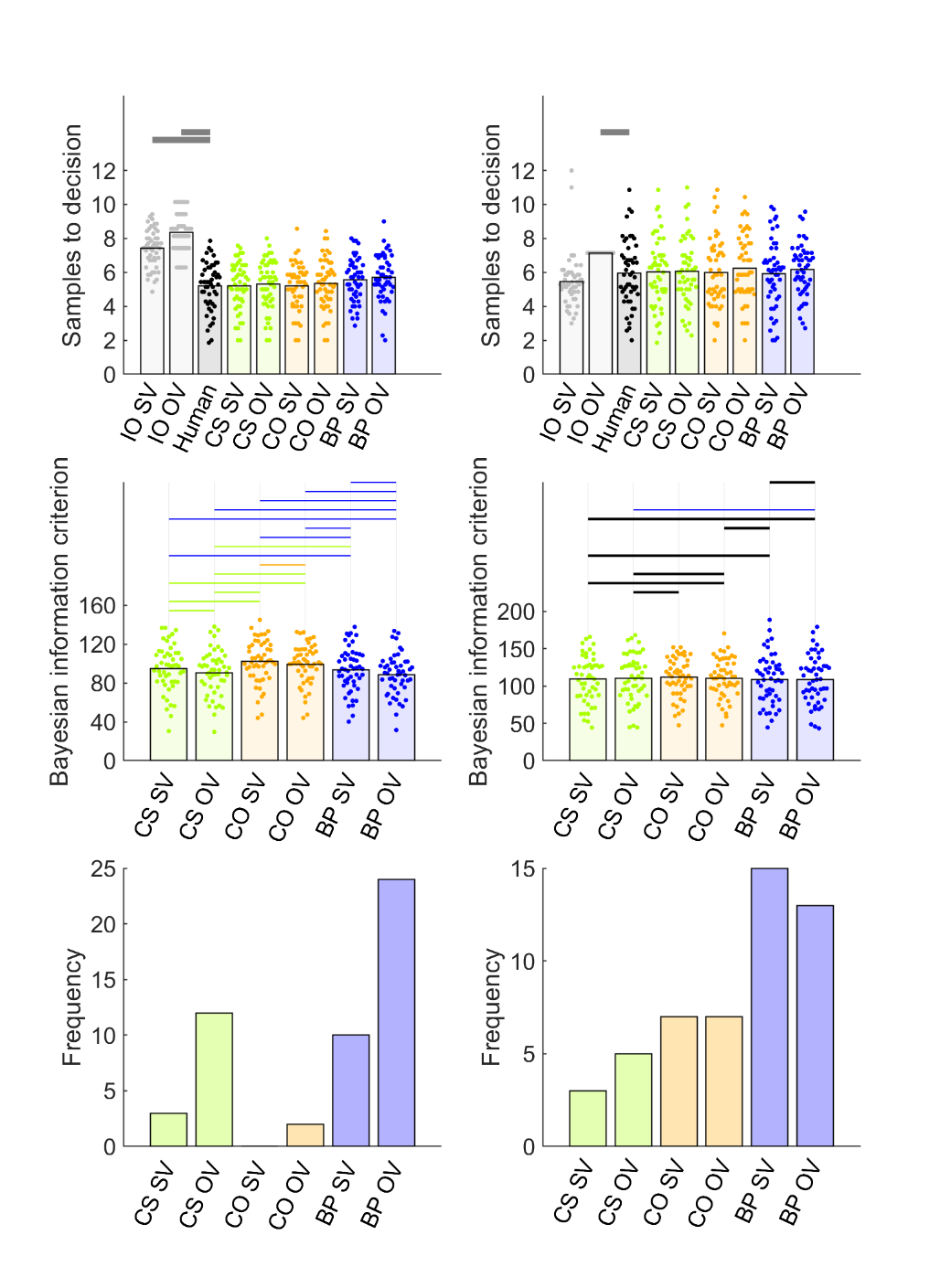
Nevertheless, our manipulation of payoff scheme, but not any other task feature, sufficed to modulate the sampling rates of the IO models, if not also those of the participants. The first row of Figure 4 shows Bayesian pairwise tests (threshold *BF10* > 3, moderate evidence for different means) from the studies without any first phase, comparing participants’ sampling (black points) against that of the IO OV model with a payoff structure that rewards only the top three ranks (grey points). We found nearly-identical undersampling bias in the baseline (Cohen’s *d* = -2.01), squares (Cohen’s *d* = -171), timing (Cohen’s *d* = -1.74) and payoff (Cohen’s *d* = -1.96) conditions. All four of these conditions used a comparable payoff scheme where participants were instructed to try to choose one of the top-three ranked options in each sequence.

Figure 4. Model comparison for (columns from left to right): Study 1 baseline, squares, timing and payoff conditions. Points in the first and second rows show data corresponding to individual participants, while bars show their mean values. Human participant data are reproduced from Figure 2. In the first row, above human and IO samples data, grey horizontal lines indicate *BF10* > 3 (moderate evidence for different means). The second row shows BIC values (lower values indicate better model fit) for participants (points) and mean values (bars). Horizontal lines are in the colour corresponding to the better-fitting model when *BF10* > 3 or shown in black when *BF01* > 3. The blue lines suggest that, in all four conditions, BP outperforms CS and CO. The third row demonstrates that, in all four conditions, BP best fitted the most participants. Abbreviations: IO = ideal observer, CO = cut-off, CS = cost to sample, BP = biased prior, OV = objective values.



Study 1 full is the only condition where participants (and IO models) were instructed to choose the highest-ranked option possible, instead of only one of the top-three ranked options. And, IO sampling decreased in this condition to some degree compared to the other conditions. The first row of Figure 5 shows results for the two conditions with an initial rating phase (ratings and full) and therefore with optimality measures from both IO SV and IO OV. Participant undersampling compared to both IO SV (Cohen’s *d* = -1.20) and IO OV (Cohen’s *d* = -1.72) is present for Study 1 ratings (left column), a condition in which the top three ranks are rewarded. However, undersampling in Study 1 full (right column) was completely eliminated for IO SV (Cohen’s *d* = 0.19) and the effect size of undersampling was reduced by more than half for IO OV (Cohen’s *d* = -0.61), though it retained significance. Note that the results for Pilot full showed an elimination of undersampling altogether for both IO OV and IO SV (and even oversampling for IO SV), which is a somewhat more striking result than what we obtained for Study 1 full. In our next study (Study 2), we will resolve this issue by implementing a full condition with improved design elements and a statistically better-powered sample size. Provisionally, we conclude that undersampling bias was greater in all the conditions that rewarded only the top three highest ranked choices, compared to the only model that used a payoff scheme that rewarded all choices commensurate with the chosen option value (Study 1 full).

Figure 5. Model comparison for Study 1 rating (left column) and full (right column) conditions. Points in the first and second rows show data corresponding to individual participants, while bars show their mean values. Human participant data are reproduced from Figure 2. In the top row, grey horizontal lines above human and IO samples data indicate when *BF10* > 3 (moderate evidence for different means). Human and IO sampling never showed *BF01* > 3 (moderate evidence for equal means). The second row shows BIC values (lower values indicate better model fit) for participants (points) and their mean values (bars). Horizontal lines are shown in the colour corresponding to the better-fitting model when *BF10* > 3 or in black when *BF01* > 3. In the rating condition, BP OV fits better than any other model (blue horizontal lines). In the full condition, the result is more ambiguous, as only BP OV significantly better fits behaviour than any other model. The third row demonstrates that, in both conditions, BP models best fitted the most participants. Abbreviations: IO = ideal observer, CS = cost to sample, CO = cut-off, BP = biased prior, OV = objective values, SV = subjective values.



Lastly, we evaluated computational theoretical models that could explain biases in individual participants. All the conditions produced similar results. All four conditions in Figure 4 replicate unambiguous evidence favouring the BP model, based on both statistical tests on individual participant BIC values (middle row) and frequencies of best-fit participants (lower row). For Study 1 ratings (Figure 5, left), BP OV dominated all other models. For Study 1 full, statistical comparisons of BIC values were somewhat ambiguous (Figure 5, middle right), though both BP models best fit the most participants (Figure 5, lower right). Overall, the evidence collectively favours BP models as the most common explanation of participants’ choices.

Study 2 Introduction

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 in 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 a direct assessment of bias in this condition difficult to determine with high precision. Our second study aimed to obtain a higher quality estimate of participant sampling bias in the full condition by overcoming a number of limitations of our previous full condition designs. We increased the target sample size from approximately 50 (in Pilot full and Study 1 full) to 151 in Study 2 full. Additionally, we generated a new set of sequence option values for every participant, whereas in Pilot full and Study 1 full, all participants engaged with sequences that were fixed in advance.

Study 2 Methods

Participants

One hundred fifty one participants based in the UK enrolled, using the participant recruitment platform Prolific.

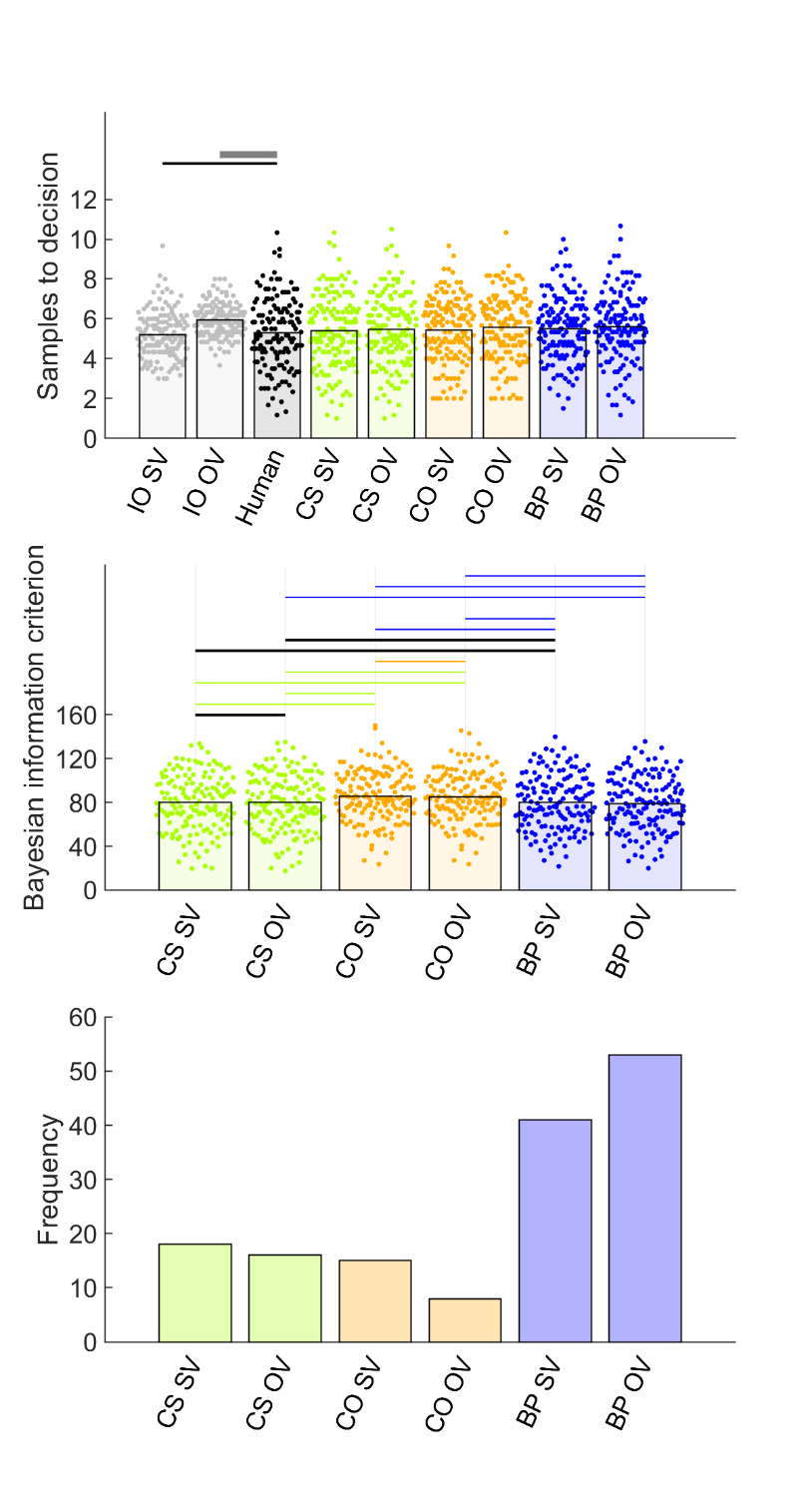
Procedures.

The study was developed in Javascript jsPsych 7.3.1 (de Leeuw et al., 2023). In Phase 1, participants rated 90 prices (the same smartphone prices used in our three studies reported above) two-times each, with all stimuli presented in one random sequence. Prices appeared above a 1 to 100 scale, and participants indicated the “attractiveness” of each price via mouse click on the scale. The mean (over participants) Pearson’s correlation coefficient between the two ratings for each price was .85. Next, participants performed an optimal stopping task with six sequences of 12 price option values, randomly sampled without replacement from the 90 prices. The study implemented self-paced screen timing. There were no grey squares. Instead, upon choice, the paradigm proceeded directly to the feedback screen. The feedback screen appeared 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 the previous studies reported herein (Figure 2). From Figure 6, which shows Bayesian pairwise test results comparing participants’ sampling to that of the two ideal observers (*BF01* > 3, moderate evidence for null model), it can be seen that participants sample statistically equivalently to IO OV (Cohen’s *d* = .05) and undersample compared to IO SV (Cohen’s *d* = -0.32). Study 2 also replicated the model-fitting results we found throughout our studies reported herein, with the evidence favouring the two BP models as outperforming other models, especially in terms of the number of best-fit participants (Figure 6, lower row).

Figure 6. Model comparison for Study 2. Points in the first and second rows show data corresponding to individual participants, while bars show their mean values. Human participant data are reproduced from Figure 2. In the first row, human and IO samples are demarcated by thin black horizontal lines when *BF01* > 3 (moderate evidence for equal means) or thick grey lines when *BF10* > 3 (moderate evidence for different means). 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. BP OV fits behaviour better than any model except CS SV. The third row demonstrates that the two BP models best fit the most participants. Abbreviations: IO = ideal observer, CS = cost to sample, CO = cut-off, BP = biased prior, OV = objective values, SV = subjective values.



Study 3 Introduction

Figure 2 suggests that participants are loath to change how much they sample. They are not sensitive to the presence or absence of the various methods features listed in Table 1. And, even though rewarding only the top-three options leads the IO model to increase the number of options it samples, participants do not correspondingly increase how much they sample under this incentivisation scheme. The goal of Study 3 was to ensure that our implementation of the optimal stopping task was methodologically viable and that it is in practice possible to experimentally modulate how much participants sample at least to some degree. 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 length. Participants in their study appeared reluctant to increase how much they sampled, whereas the ideal observer increased its sampling rate to adapt to the longer sequence lengths without constraint – a pattern that appears consistent with the reluctance with which participants increase their sampling rates in our studies reported herein. In our third study, we replicated this effect of sequence length on participants’ average number of samples, using sequence lengths of 10 and 14 options. We also took the opportunity to further replicate and bolster our conclusion that a biased prior is a worthy explanation of participants’ performance, using the two more datasets Study 3 provides.

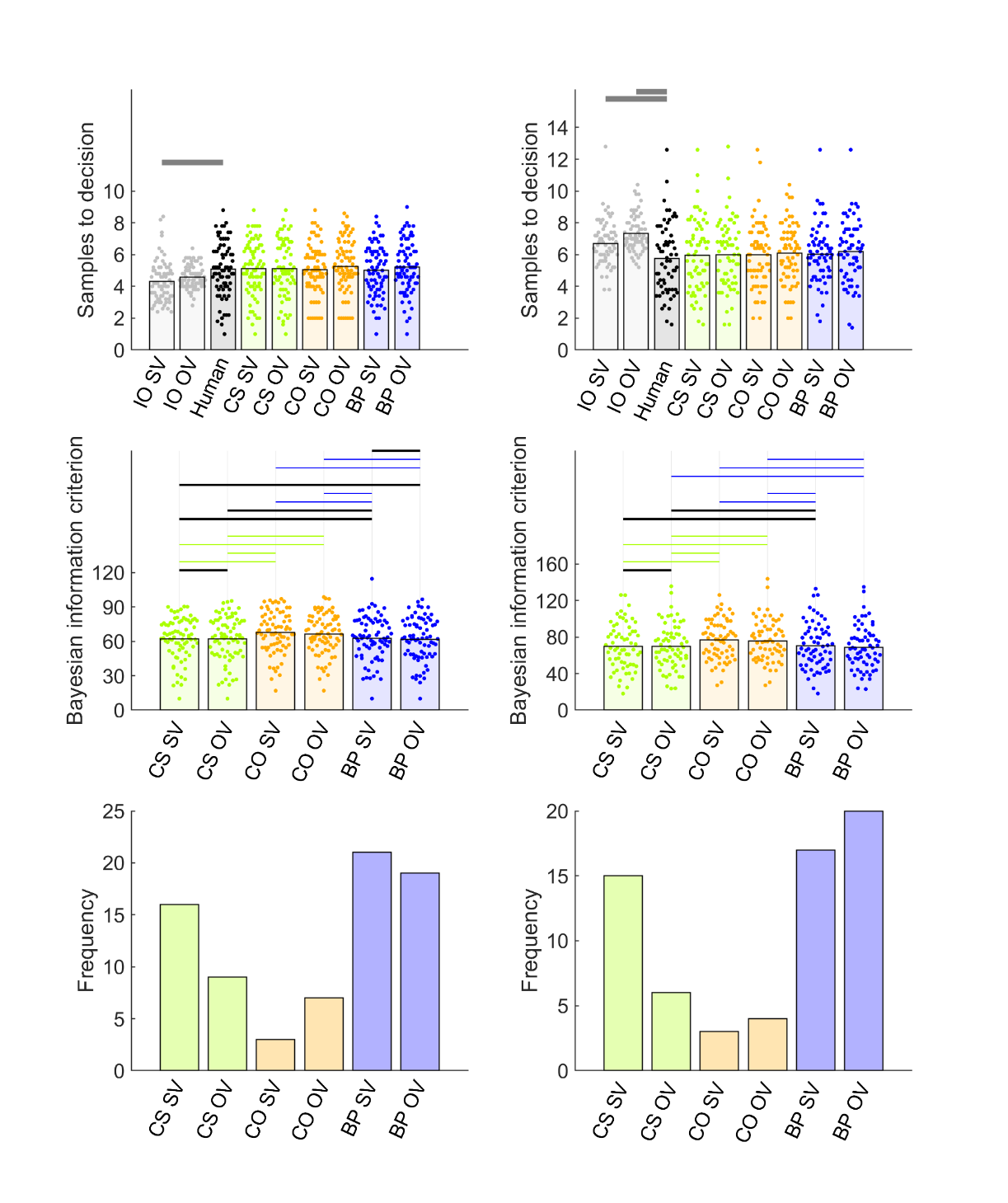
Study 3 Methods

The preregistration of Study 3 can be found at <https://osf.io/vcf7u>. We enrolled 140 participants from the UK using Prolific, where half the participants 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 of the optimal stopping phase of the study. The averages (over participants) of the Pearson’s *r* values computed between the two phase 1 ratings to each price were .88 for the 10 option condition and .84 for the 14 option condition.

Study 3 Results and Discussion

Figure 2 (rightmost bars) shows that our hypothesis was confirmed: participants sampled significantly more for longer sequences, replicating the findings in Costa & Averbeck (2015). As in Costa & Averbeck, undersampling was not observed for the shorter sequence length. In our study, the Bayesian tests (shown in Figure 7) suggest that at a sequence lengths of 10 options, participants slightly *overs*ampled (rather than undersampled) compared to IO OV (Cohen’s *d* = 0.33), while the difference with IO SV remained inconclusive (Cohen’s *d* = 0.26). In contrast, participants showed clearer evidence for an undersampling bias at sequence lengths of 14, as they sampled statistically *less than* both IO OV (Cohen’s *d* = -.63) and IO SV (Cohen’s *d* = -0.44). Our model-fitting also confirmed our hypothesis that participants’ sampling biases could be explained best by a BP model, though the CS model clearly made a stronger contribution in Study 3 than Study 2. In summary, participants can and will change their sampling behaviour to some degree in some contexts. However, at least on tasks using the economic domain that we studied here, participants’ number of samples are “held in place” by a largely pessimistic expectation about the quality of upcoming samples, which discourages them from increasing their sampling to the optimal degree, which leads to increasing undersampling bias as sequences lengthen.

Figure 7. Model comparison for Study 3 10 options (left column) and 14 options (right column) conditions. Points in the first and second rows show data corresponding to individual participants, while bars show their mean values. Human participant sampling data are reproduced from Figure 2 and show a significant effect of sequence length (*P* < 0.05). In the first row, human and IO bars are demarcated by grey lines when *BF10* > 3 (moderate evidence for different means). Human and IO never showed *BF01* > 3 (moderate evidence for equal means). 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 third row demonstrates that BP best fit the most participants’ data, though CS models also account for some participants. Abbreviations: IO = ideal observer, CS = cost to sample, CO = cut-off, BP = biased prior, OV = objective values, SV = subjective values.



General Discussion

In our pilot studies, we first established that we could replicate an undersampling bias (Baumann et al., 2020; Cardinale et al., 2021; Costa & Averbeck, 2015) by adapting a previous implementation of an economic full-information problem (Costa & Averbeck, 2015). In addition to this replication, we also tested novel task variables (used in other studies like van de Wouw et al., 2022) that hypothetically might modulate undersampling bias. We were able to modulate the size of the undersampling bias in two ways: by manipulating the payoff scheme and by manipulating sequence length.

Multiple times, we implemented a so-called “full” condition, which was the only condition where participants simply maximised the option value of their choices, rather than obtain one of the top-three ranked options. Every condition except the full condition replicated a robust participant undersampling compared to both IO OV and IO SV (where it was available in the Study 1 ratings condition). In contrast, in full conditions (except for sequence length 14), undersampling was inconsistent at best when compared to IO OV and was completely eliminated (or even showed oversampling) when compared to IO SV. This robust undersampling in non-full conditions paired with reduced undersampling in full conditions was not because participants changed their behaviour much. It was because both of the IO models lowered their sampling rates in these full conditions.

The IO models seem to reduce their sampling rates in the full condition because of its payoff scheme. The full condition implemented a variety of task methods not present in Costa & Averbeck (2015), though we were able to experimentally eliminate as alternative possible causes the other task features in the full condition, including screen timing, grey squares, extrinsic monetary reward, the presence of a first rating phase and the use of subjective or objective values in the IO model. Thus, we must conclude that, because the IO model is willing to increase its sampling rate to the one appropriate for its payoff scheme (when the top three ranks are rewarded), and participants are not so willing (as they tend to always sample at nearly the same rate), the size of the undersampling bias is correspondingly modulated by the payoff scheme.

We observed a similar phenomenon in Study 3 for sequence length. Although both participants and the IO model increased their sampling rates for longer sequences (14 options compared to 10), the IO model showed a greater sampling increase for longer sequences than participants did, and thus the undersampling bias correspondingly increased – a finding replicated from Costa and Averbeck (2015). It appears that, while sometimes participants can, albeit reluctantly, increase their sampling, they generally prefer to limit how much they sample, even when it is optimal to increase sampling rate more than they do.

Crucially, we were able to theoretically explain, in terms of a computational mechanism, participants’ sampling bias. Our model fits suggest that participants’ reluctance to increase sampling rates when it is optimal to do so arises because participants expect future option values to be lower on average then they will be. In the BP model, we added a constant to the mean of the prior “generating” distribution of option values. In cases where undersampling occurs, this constant appears to be negative (See, for example, Figure S6, middle row) – reflecting a pessimistic expectation. These parameter values become slightly positive in many full conditions, in which undersampling appears reduced (e.g., Figure S8). Participants of course still make errors in these full conditions and our data suggests that mis-specified prior expectations may account for these suboptimal decisions as well. We should note, however, that the CS model – in which participants perceive sampling itself to be costly or rewarding – well-fit a substantial number of participants and therefore may well have influenced suboptimal decisions in many of our participants as well.

Our study alone cannot explain whence this biased prior arises and this remains an open question for future research. It is possible that participants, in economic contexts, might adopt a “safe” or conservative strategy (i.e., response bias) that protects against getting stuck with an especially poor outcome. Indeed, inspection of the ranks that participants achieved with their choices (a measure of their choice accuracy), shown in the first rows of Figures S4, S6, S7, S8 and S9 suggests that the quality of participants’ choices closely approximated those of the IO’s choices, despite their suboptimally low sampling rates. Consequently, one can adopt a pessimistic stance that protects from the uncertainty of a poor outcome and still “satisfice”; that is, perform at near-optimal levels.

The BP model appeared to garner replicated evidence across datasets whether participants had the opportunity to learn the prior distribution from a preceding ratings task (e.g., Study 1 ratings condition) or not (Figure 4). One possibility is that participants develop from the outside world a pre-conceived idea of the distribution of outcomes and new learning within the task (either from the ratings phase or from the sequence options themselves) does not succeed in overwriting this preconception. Another possibility is that participants may learn the prior to some degree from the option values as they experience one sequence after another (Goldstein et al., 2020). However, we did not find learning effects across sequences here, consistent with previous reports of studies on full-information problems (Lee, 2006). Nor is it clear why this strategy would lead to a pessimistic prior and undersampling. Baumann et al. (2020) included a different approach from ours to using a learning phase prior to the optimal stopping task to ensure that participants were acquainted with the generating distribution. As in Lee and Courey (2020), participants learned abstract mathematical density functions. Based on these, participants drew histograms of distributions, on which they received feedback to ensure their understanding. According to Goldstein and Rothschild (2014), such a graphical elicitation technique can lead to rather accurate representations of probability distributions in participants. This approach is not likely to be especially ecologically-valid, however. Another appealing explanation for participants’ apparently biased prior is that participants did not treat our task as a full information problem and did not use any prior distribution. Indeed, the CO heuristic derives from a “prior-free” mathematical solution to the secretary problem, which gives optimal performance assuming that participants have no knowledge of the prior distribution. Nevertheless, the CO heuristic did not perform well in our model comparison by contrast to either of the models based on the full information problem solution (CS and BP). More research into how participants learn option value distributions would be useful. We hypothesise that a biased prior might persist, regardless of how participants are exposed to the prior, though more study is needed to generalize beyond our study.

Although we were unable to reliably induce participants to oversample in the present work and instead we identified variables that modulate the size of undersampling bias, others like van de Wouw et al. (2022) have demonstrated and replicated oversampling bias. Their work, rather than presenting options as numeric prices as we did here, communicated option values using images, such as the attractiveness of faces, foods and holiday destinations. Our manipulations of task features in Study 1 have already tested and rejected other task differences used in their paradigm that might give rise to oversampling (e.g., grey squares, timing, etc), leaving the pictorial stimulus domains as the most likely instigator of oversampling in van de Wouw et al. (2022). It is possible that a biased (i.e., overly optimistic) prior might account for oversampling in image-based contexts as well as undersampling in number-based, economic domains.

Our model comparison is the first time theoretical models that specify the computations humans use to solve full information problems have been compared so comprehensively. Costa and Averbeck (2015) introduced the parameterised cost to sample model that we consider here and fitted that model to participants’ sampling choice on an economic full information task. However that study did not perform a model comparison with alternative models. Moreover, our current study for the first time provides a comprehensive parameter recovery analysis for this model and introduces and tests a number of other similar theoretical models. Our work also builds on the approach recently taken by Baumann et al. (2020), who compared the CO OV model we consider here with “threshold models” (Lee, 2006). Although these threshold models are useful tools for directly estimating participants’ choice thresholds at each sequence position from participants’ behavioural data, we took a different approach for our model comparison. Our approach was to compare models that are “computational” in the sense that they specify the computations that participants might theoretically be using to accurately solve the task, including specification of how participants compute their decision thresholds. In the parameterised Costa & Averbeck (2015) models we considered, the action value for sampling again (See Methods) acts as the effective decision threshold, which varies over trials depending on the perceived prospect of sampling a better option value, and which the value of the current option needs to exceed before the model will commit to a choice. Using these models, there is no need to explicitly parameterise the threshold, as it arises naturally from the computations within the model. Moreover, we obtain the added capability of parameterising bias terms and then simulating how these bias terms influence the computation of thresholds, which cannot be done using threshold models, at least as they have been implemented in the past. Nevertheless, our results largely agree with a key finding from the model comparison in Baumann et al., who showed that models that change their decision threshold across samples better fit participants’ data than the CO OV model, in which the decision threshold is established after the cut off sequence position and henceforth remains fixed.

In summary, we show that participants’ sampling behaviour on optimal stopping tasks is relatively insensitive to many methodological manipulations. In contrast, the ideal observer (which reflects optimal performance) is relatively more sensitive than participants to at least payoff schemes and sequence lengths, such that these two factors can modulate the degree of undersampling bias. We explain participants’ sampling behaviour using a theoretical model by which participants implement optimal Bayesian computations to solve the task accurately, but a systematic undersampling bias develops when participants mis-predict the quality of upcoming sampling, based on a biased belief in the distribution of possible outcomes.

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, 388–407. <https://doi.org/10.3758/s13428-019-01237-x>

Averbeck, B. B. (2015). Theory of choice in bandit, information sampling and foraging tasks. *PLoS Computational Biology 11*, e1004164. <https://doi.org/10.1371/journal.pcbi.1004164>

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>

Castellano, S., Cadeddu, G., & Cermelli, P. (2012). Computational mate choice: Theory and empirical evidence. *Behavioural Processes, 90,* 261–277. <https://doi.org/10.1016/j.beproc.2012.02.010>

Castellano, S., & Cermelli, P. (2011). Sampling and assessment accuracy in mate choice: A random-walk model of information processing in mating decision. *Journal of Theoretical Biology, 274,* 161–169. <https://doi.org/10.1016/j.jtbi.2011.01.001>

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>

Freeman, P. R. (1983). The secretary problem and its extensions: A review. *International Statistical Review / Revue Internationale de Statistique 51*, 189–206. <https://doi.org/10.2307/1402748>

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>

Goldstein, D. G., McAfee, R. P., Suri, S., & Wright, J. R. (2020). Learning When to Stop Searching. *Management Science 66,* 1375–1394. <https://doi.org/10.1287/mnsc.2018.3245>

Goldstein, D. G., & Rothschild, D. (2014). Lay understanding of probability distributions. *Judgment and Decision Making 9*, 1–14. <https://doi.org/10.1287/mnsc.2018.3245>

Guan, M., & Lee, M. D. (2018). The effect of goals and environments on human performance in optimal stopping problems. *Decision 5,* 339–361. <https://doi.org/10.1037/dec0000081>

Guan, M., & Stokes, R. (2020). A cognitive modeling analysis of risk in sequential choice tasks. *Judgment and Decision Making 15,* 823–850.

Kolling, N., Scholl, J., Chekroud, A., Trier, H. A., & Rushworth, M. F. S. (2018). Prospection, perseverance, and insight in sequential behavior. *Neuron 99,* 1069–1082. <https://doi.org/10.1016/j.neuron.2018.08.018>

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>

Lee, M. D., & Courey, K. A. (2020). Modeling Optimal Stopping in Changing Environments: A Case Study in Mate Selection. *Computational Brain & Behavior 4*, 1–17. <https://doi.org/10.1007/s42113-020-00085-9>

Lee, M. D., O’Connor, T. A., & Welsh, M. B. (2005). Decision-making on the full information secretary problem. *Proceedings of the Twenty-Sixth Conference of the Cognitive Science Society*, 819–824.

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>

Scholl, J., Trier, H. A., Rushworth, M. F., & Kolling, N. (2022). The effect of apathy and compulsivity on planning and stopping in sequential decision-making*. PLoS Biology 20*, e3001566. <https://doi.org/10.1371/journal.pbio.3001566>

Seale, D. A., & Rapoport, A. (1997). Sequential decision making with relative ranks: An experimental investigation of the "secretary problem". *Organizational Behavior and Human Decision Processes, 69,* 221–236. <https://doi.org/10.1006/obhd.1997.2683>

Seale, D. A., & Rapoport, A. (2000). Optimal stopping behavior with relative ranks: The secretary problem with unknown population size. *Journal of Behavioral Decision Making, 13,* 391–411. [https://doi.org/10.1002/1099-0771(200010/12)13:4<391::AID-BDM359>3.0.CO;2-I](https://doi.org/10.1002/1099-0771(200010/12)13:4%253C391::AID-BDM359%253E3.0.CO%3B2-I)

Sonnemans, J. (2000). Decisions and strategies in a sequential search experiment. *Journal of Economic Psychology 21,* 91–102. <https://doi.org/10.1016/S0167-4870(99)00038-0>

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. <https://doi.org/10.1017/S1930297500003557>

Zwick, R., Rapoport, A., Lo, A. K. C., & Muthukrishnan, A. V. (2003). Consumer sequential search: Not enough or too much? *Marketing Science 22*, 503–519. <https://doi.org/10.1287/mksc.22.4.503.24909>

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 25 participants (our empirical studies recruited at least approximately 50 participants) with only five sequences per participant and twelve 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 BP, 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 also formulated two further theoretical models, which did not perform so well during parameter recovery. These are described below.

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). This threshold is the key theoretical parameter in 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 are then submitted to the ideal observer model, rendering optimal performance, albeit on option values eligible for consideration only when above the threshold. 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 25 simulated participants (each shown as an individual scatter point) per simulated parameter level. 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.

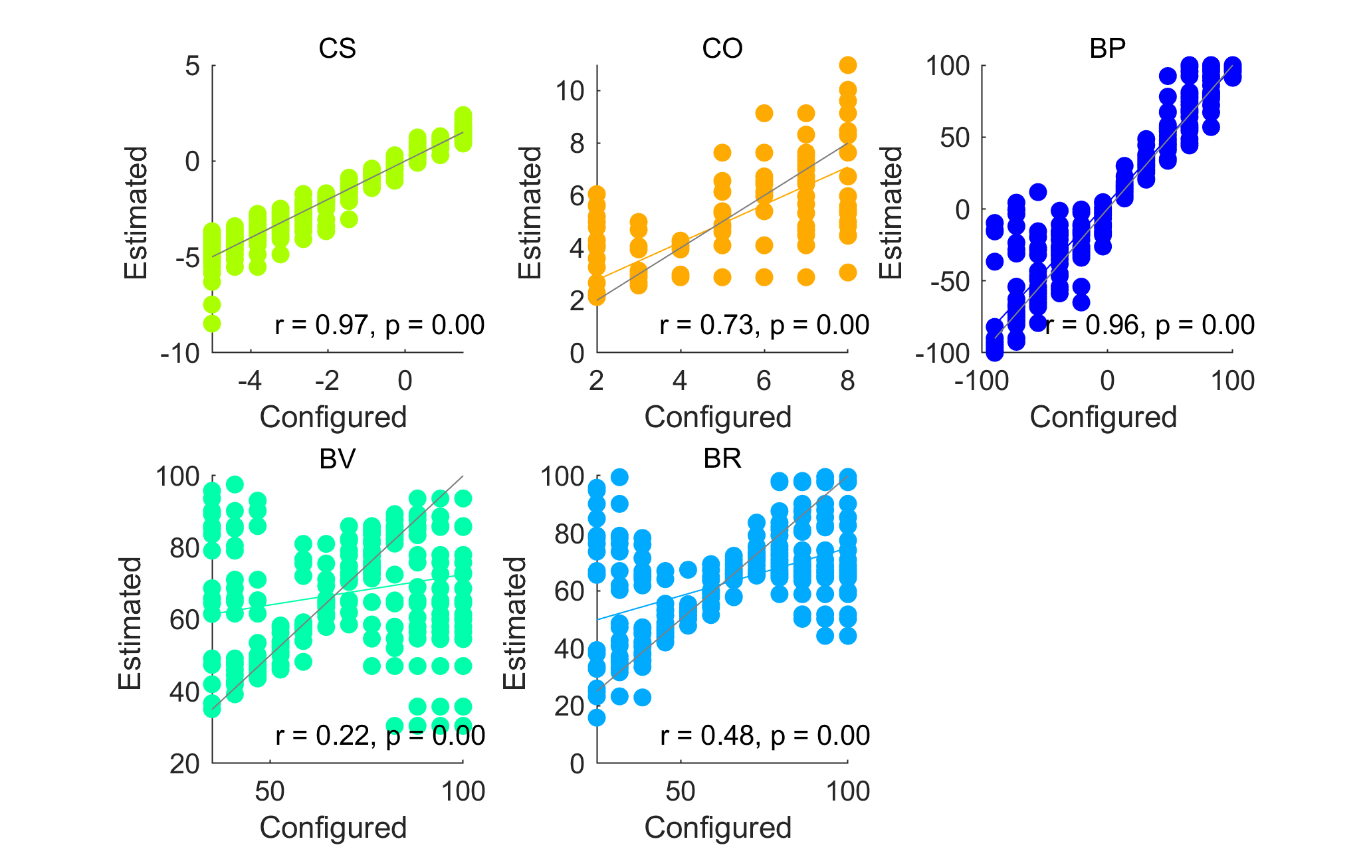


Figure S2. Top panel: We plot sampling rates for individual simulated participants (points) and their mean values (bars) for each configured parameter level in the parameter recovery analysis. Systematically varying configured parameter values successfully increases or decreases simulated sampling rates for CS, CO and BP. 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.

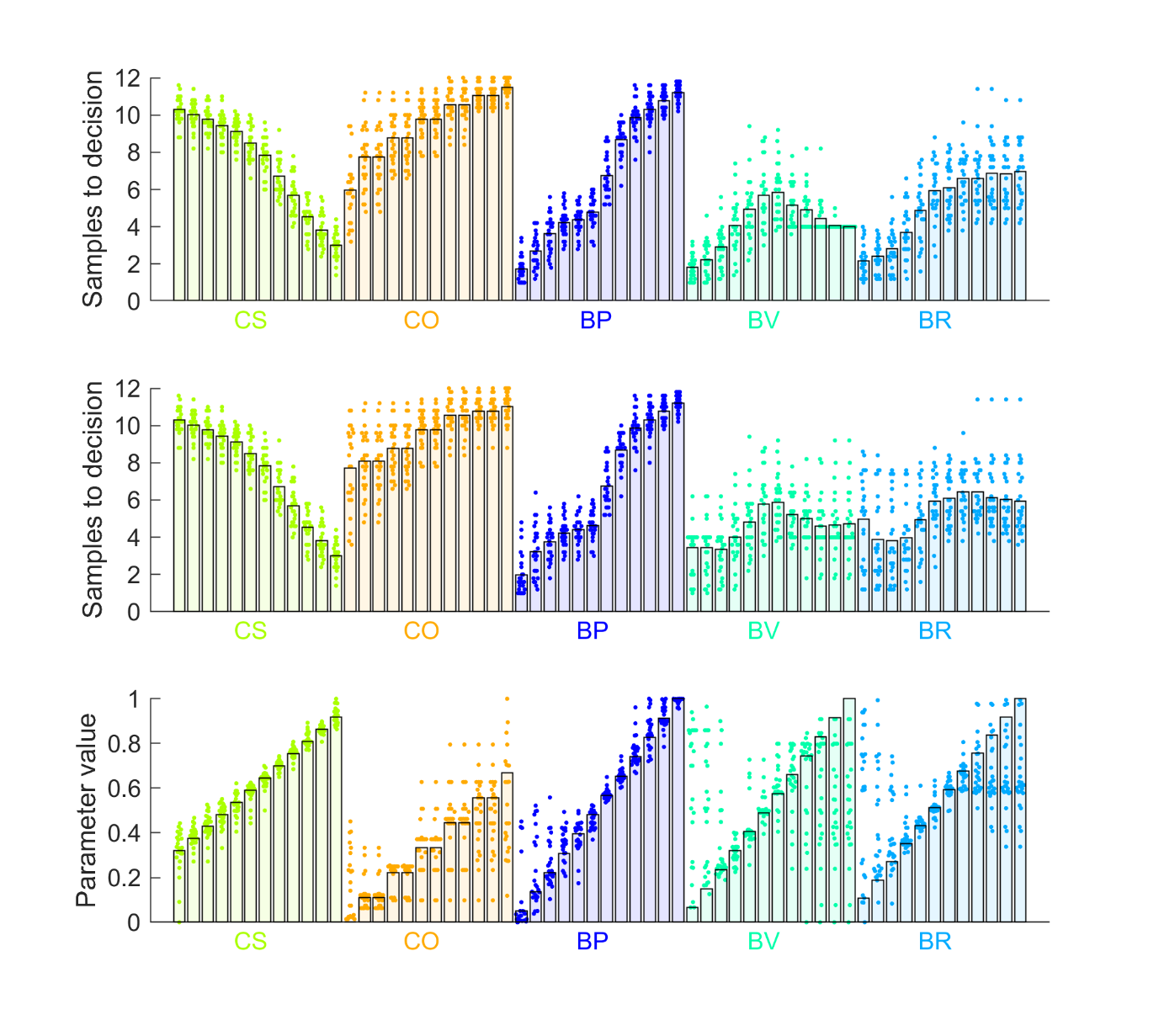


Figure S3. Sampling rates simulated using configured parameters (horizontal axis) are plotted against sampling rates computed from estimated parameters. The grey diagonal indicates when sampling rates based on configured and estimated parameters would be exactly equal. The coloured line indicates the 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.

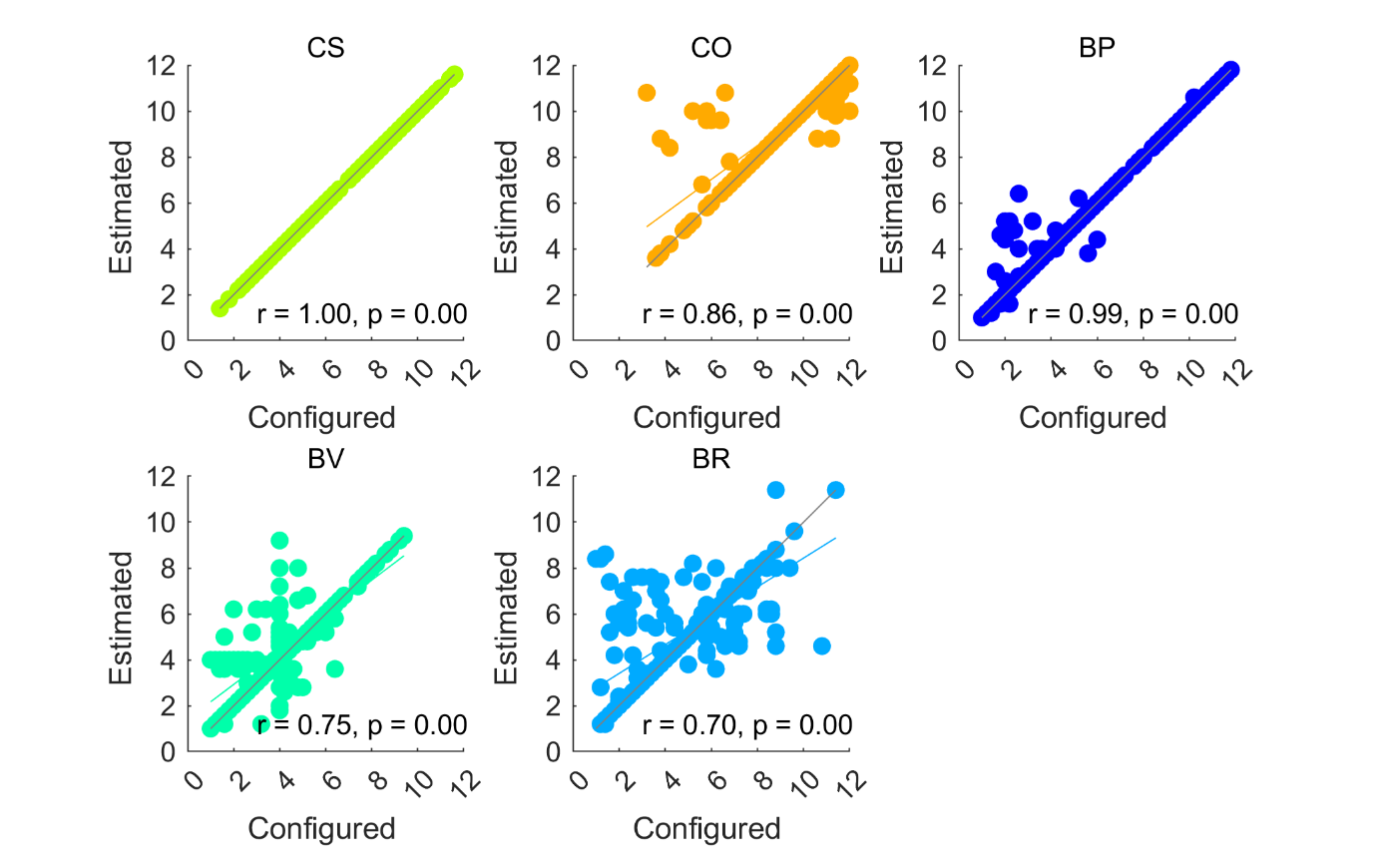


Figure S4. Model comparison for Pilot baseline (left column) and Pilot full (right column). Top and middle rows show individual participants as points and bars show their mean values. The top row shows the ranks of chosen items. The second row plots the “first” or key theoretical parameter values, estimated 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.

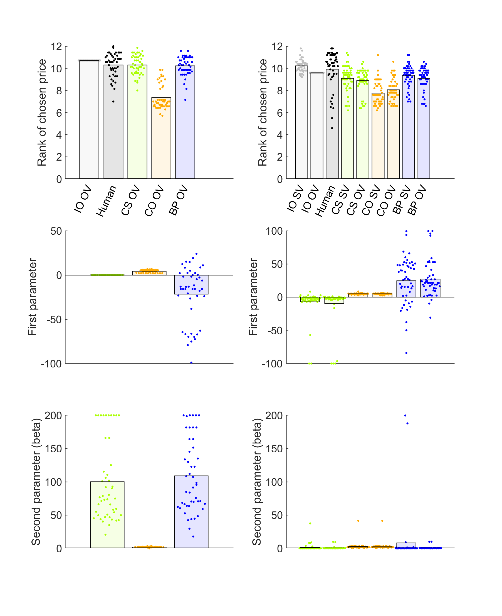


Figure S5. Cohen’s *d* effect sizes for pairwise comparisons of participants’ sampling rates in Study 1.

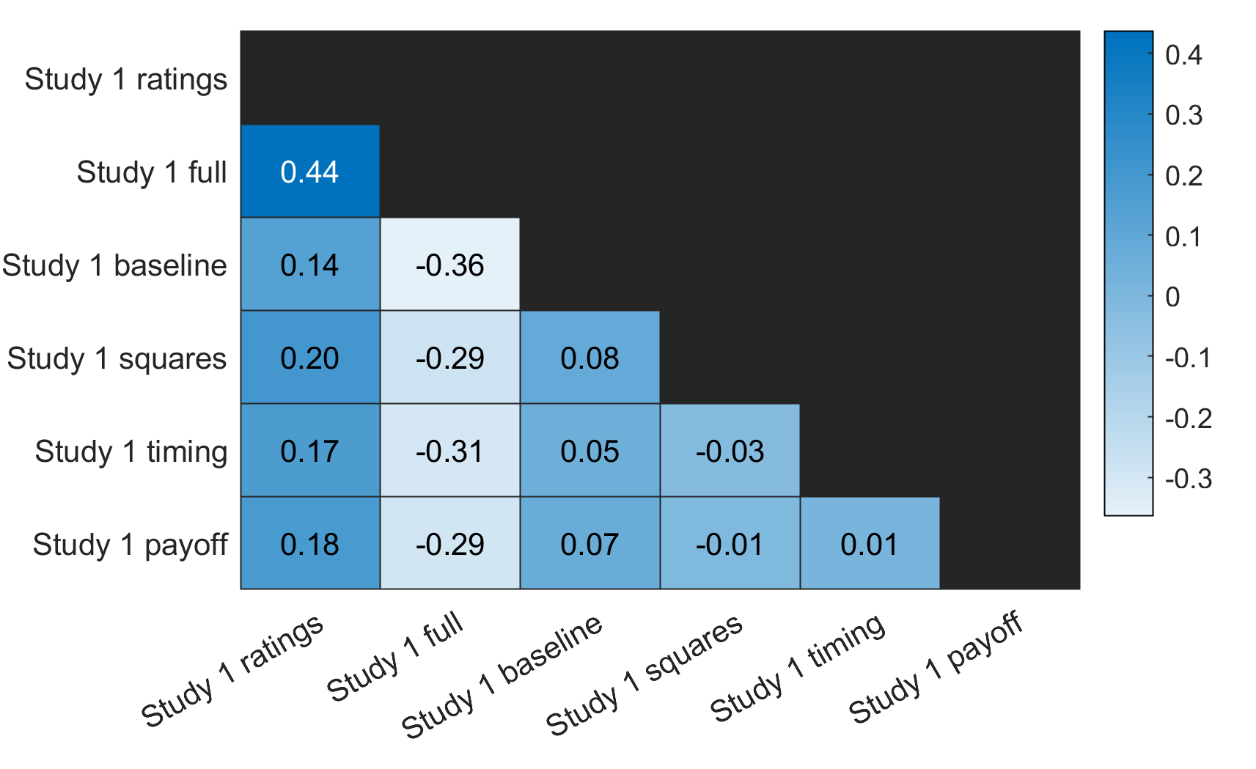


Figure S6. Model comparison for (columns from left to right): Study 1 baseline, squares, timing and payoff conditions. Top and middle rows show individual participants as points and bars show their mean values. The top row shows ranks of chosen prices. The second row plots the “first” or theoretical 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.

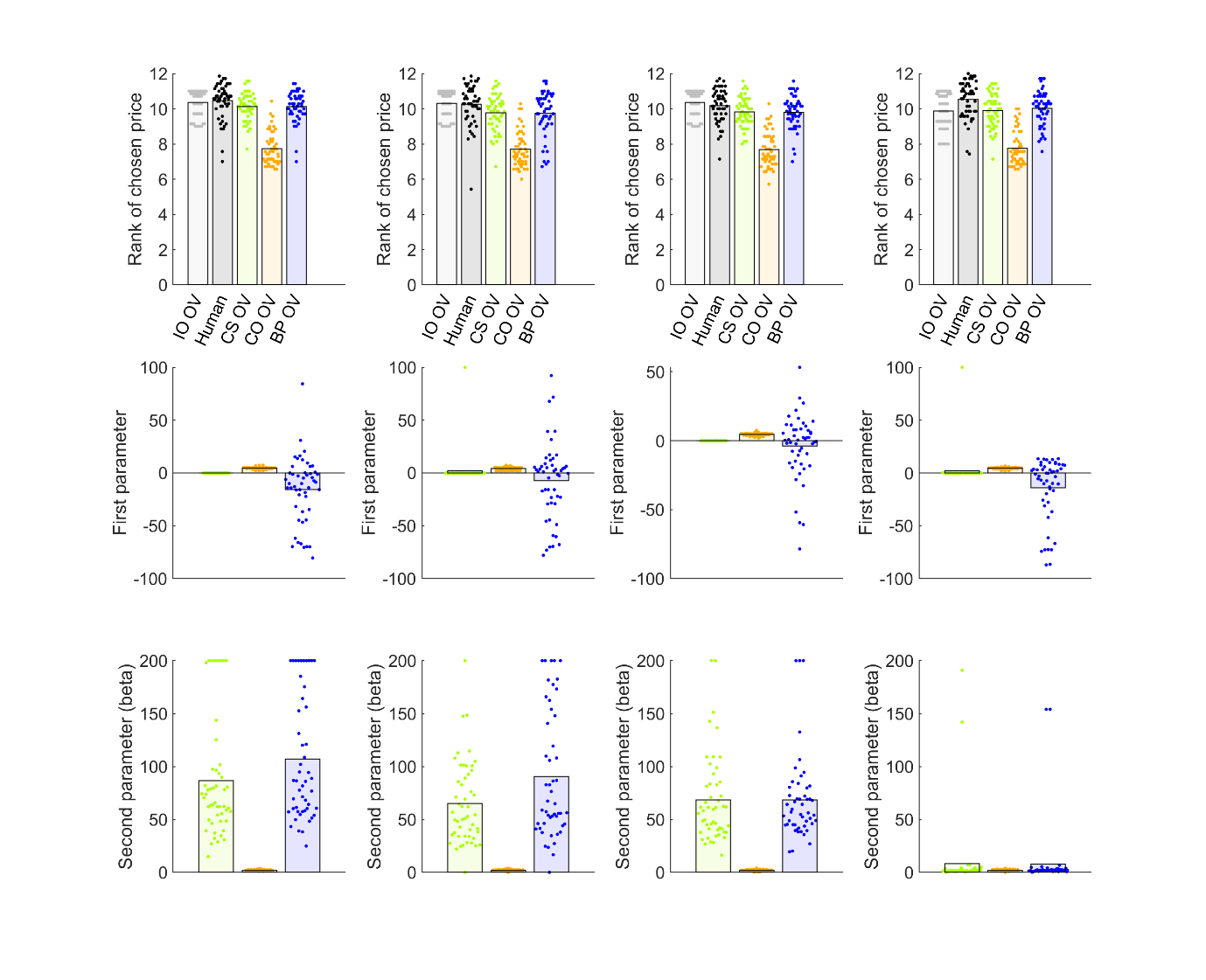


Figure S7. Model comparison for Study 1 ratings (left column) and full (right column) conditions. Top and middle rows show individual participants as points and bars show their mean values. The top row shows ranks of chosen prices. The second row plots the “first” or theoretical parameter values, estimated 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.

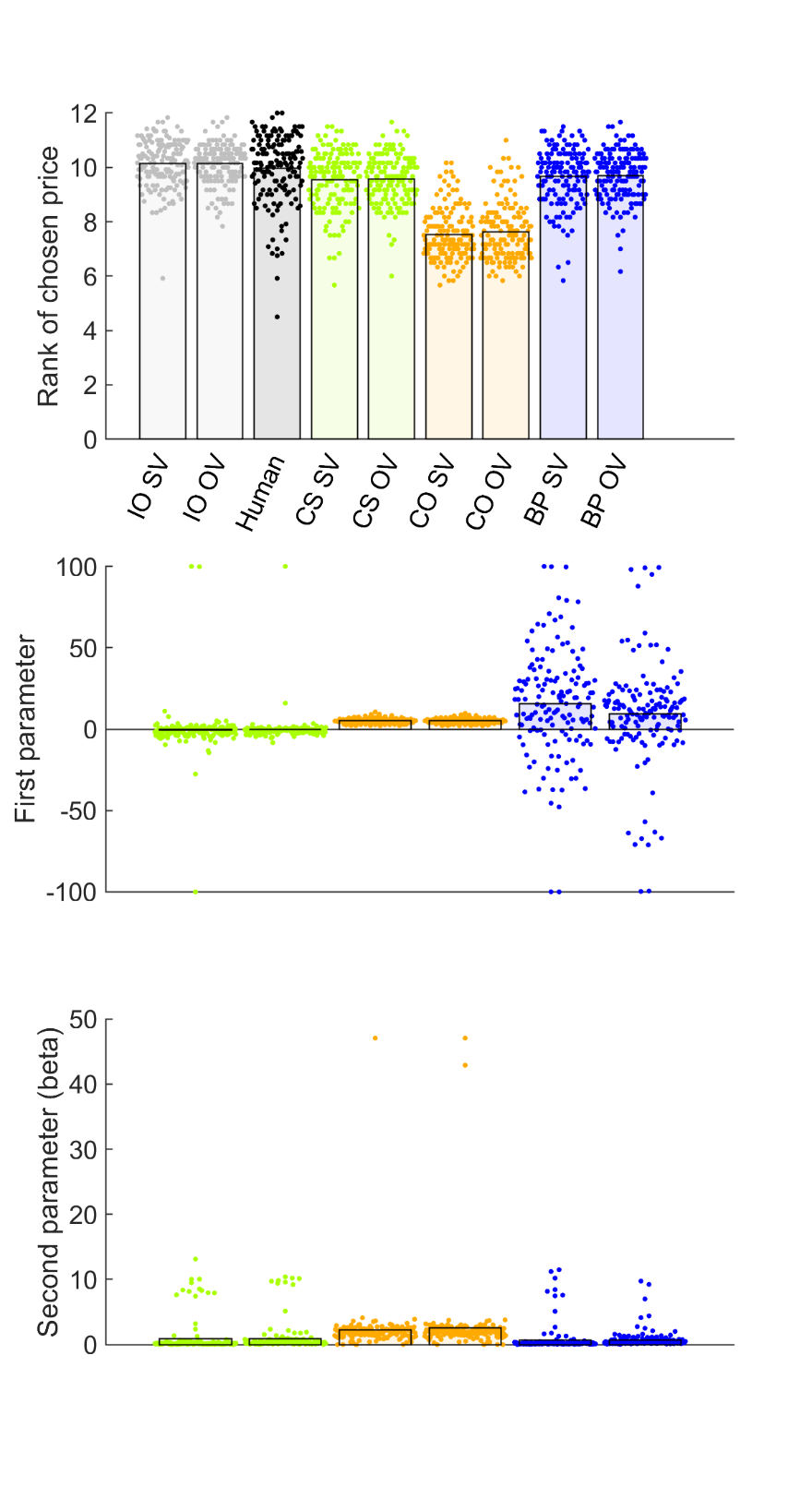
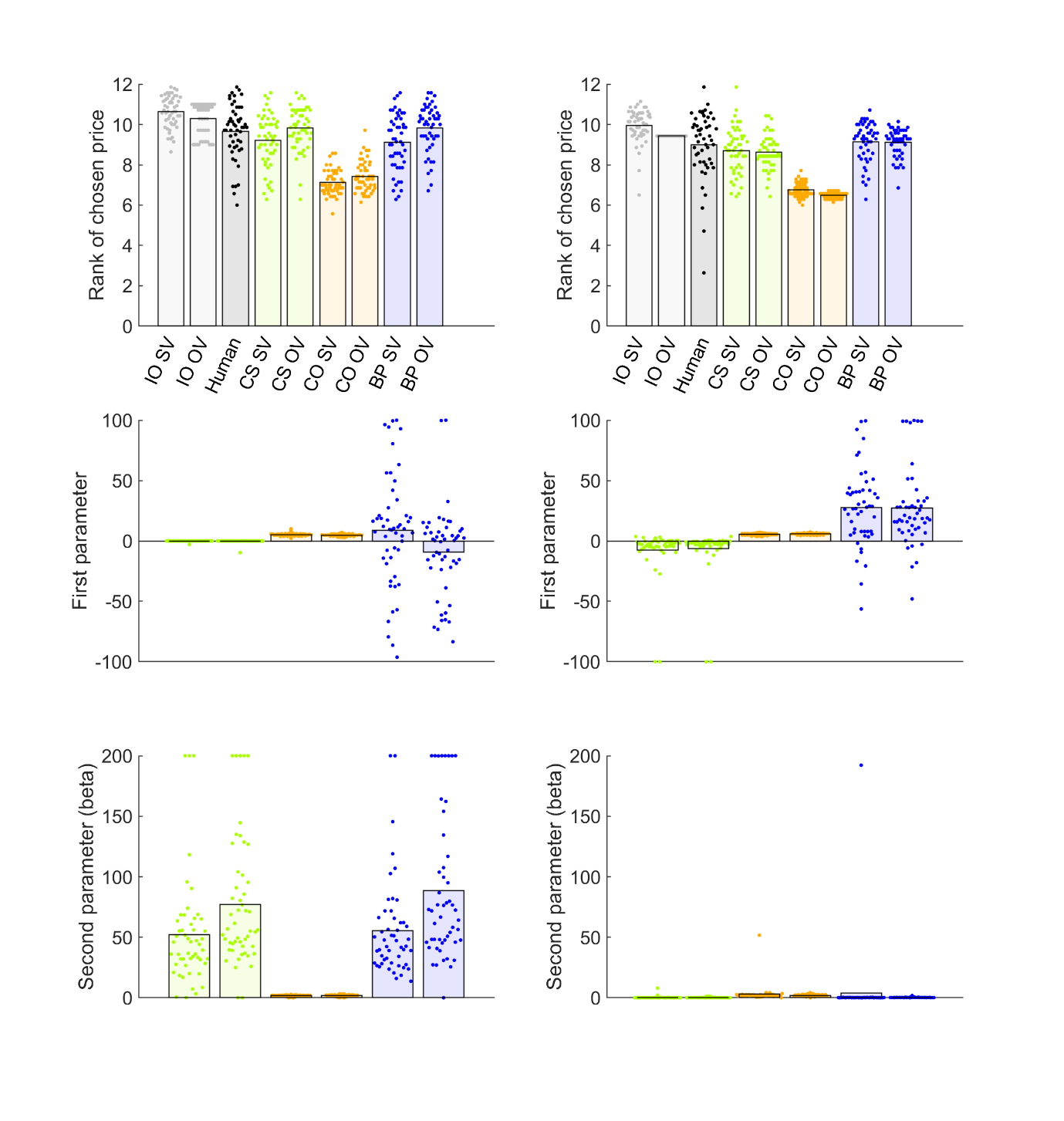


Figure S8. Model comparison for Study 2. Top and middle rows show individual participants as points and bars show their mean values. The top row shows ranks of chosen prices. The second row plots the “first” or theoretical parameter values, estimated for each fitted model. The third row shows the “second”, 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.

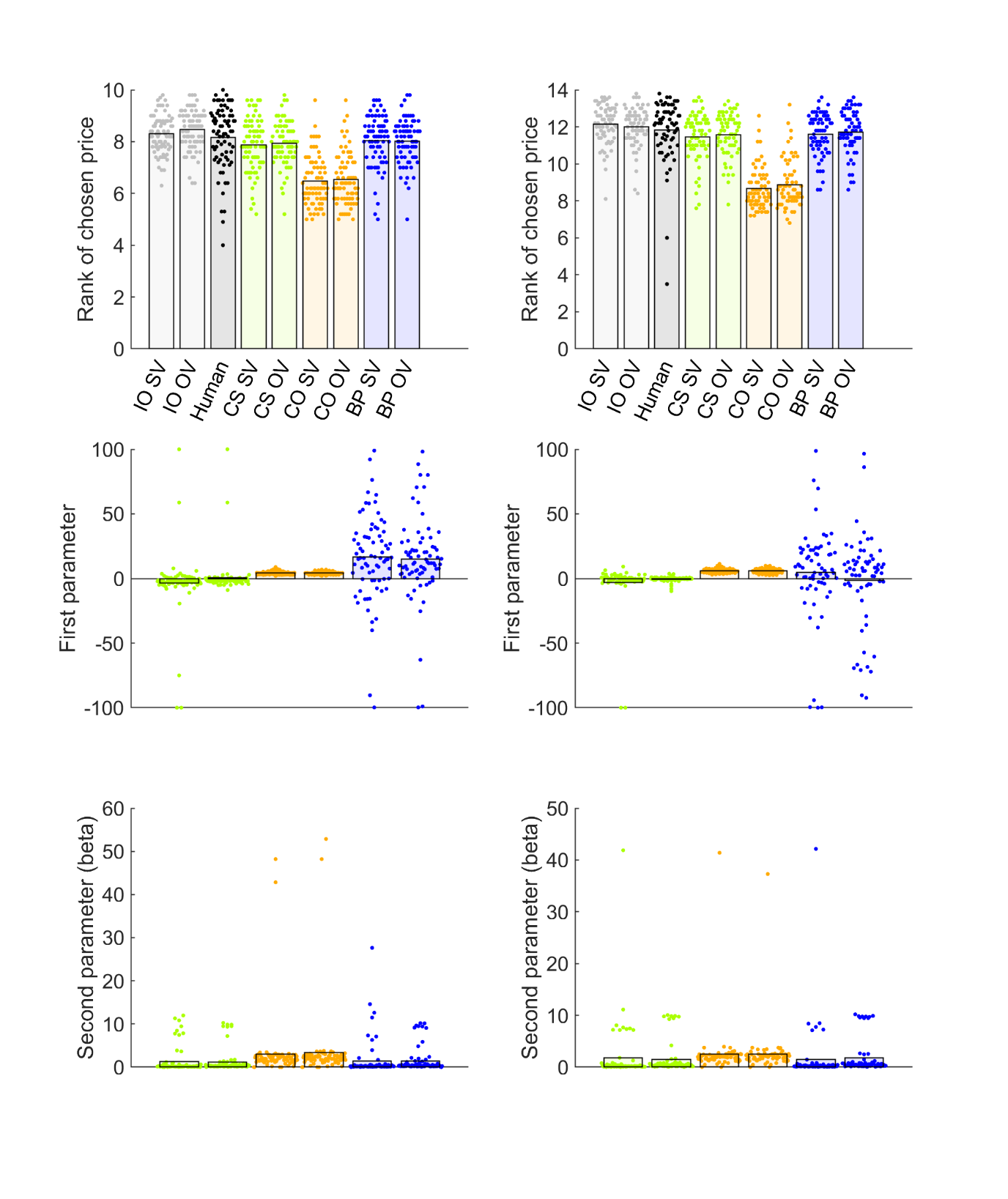


Figure S9. Model comparison for Study 3 10 options (left column) and 14 options (right column) conditions. Top and middle rows show individual participants as points and bars show their mean values. The top row shows ranks of chosen prices. The second row plots the “first” or theoretical parameter values, 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.