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”, 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 of deviation of their sampling rates from an optimality model - depends on the sequence length and the distribution of payoffs, while we find null effects for a variety of other methodological alternatives. We fit several Bayesian models of bias and a heuristic model and found participants’ biased sampling rates were best explained if participants perceived option sampling as costly. We therefore propose a new theoretical viewpoint for the human solution to full information problems, which may have broad implications for this diverse class of decision problems, including even for real-world sequential decision scenarios.

Introduction

Often in everyday life, decisions have to be made regarding options presented in sequence, like when attempting to find the best deal on a certain product or service. 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. When shopping, if one encounters a limited-time offer, should one accept it when it is available or miss it and wait for a better one? If searching for a new flat, should one accept the especially nice one they are viewing now, or risk losing it by spending time viewing more flats? The 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. There are many types of optimal stopping problem, 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, 2000), 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 systematic biases. Moreover, these models may be fashioned into theoretical models of accurate human performance, and can be parameterised to identify where within these computations biases arise. Research into the computational mechanisms that might predict when humans are accurate or biased in these tasks is especially important because of the ubiquity of these decision problems in real life and across scientific fields. Furthermore, optimal stopping tasks have potential practical applications, such as in cognitive behavioural therapy in anxiety disorders (Cardinale et al., 2021), or could be used as a general measure of problem-solving ability and psychometric intelligence (Lee et al., 2005).

Numerous versions of optimal stopping tasks abound, many of which are associated with rather different computational performance benchmarks. This longstanding diversity complicates direct comparisons between studies (for an early review, see; Freeman, 1983). Nevertheless, most optimal stopping problems (e.g., searching for a flat in a competitive market) share basic features. An 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). We focus here on full information problems. This version of optimal stopping problem arguably most closely resembles real-world decision problems and perhaps pose the most computationally difficult challenge for an agent. Here, the agent knows the value of each option (e.g., how nice the currently-viewed flat is) and is not restricted to knowing only its relative rank (as in some other well-studied problems, such as the secretary problem, described below). Full information problems also can incorporate flexible payoff schemes: an agent may be rewarded to some degree by any choice, and is not restricted to be rewarded only for choosing the highest-ranked option (as in the secretary problem). Full information optimal stopping problems allow agents to harness their prior belief about the probability distribution that is generating their decision options (i.e., the generating distribution). For example, an agent can use knowledge of the housing market to prospectively compute the probability that an even nicer flat might be sampled if the current one is refused. To solve full information problems successfully, the decision maker must use knowledge of option values, their generating distribution and the reward values of their choices to solve a difficult computational problem: balancing the potential of improving on the current option against the prospective risk of losing high-ranking options if too many options are sampled (Furl et al., 2019).

Although full information problems incorporate elements from real-world decision problems, until relatively recently, most studies of human decisions on optimal stopping problems focused on a different and simpler problem known as the secretary problem. This focus likely arose, in part, because human biases have been detected for secretary problems by comparing participant sampling behaviour with an especially simple and elegantly-formulated mathematical rule for computing the optimal sampling rate (Ferguson, 1989). Unfortunately, this optimal solution involves some unrealistic restrictive assumptions that render it inapplicable to full information problems. This assumptions include, among several others, that the agent cannot use any prior belief about the option generating distribution, knows only the relative ranks of options but not their actual values and is rewarded only for choosing the top-ranked option. One common finding that emerged from research on the secretary problem is that, even when one or more of its assumptions are relaxed, participants tend to undersample, that is, sample less than is optimal (Seale & Rapoport, 1997, 2000; Sonnemans, 2000; Zwick et al., 2003). It has been proposed that the optimality model for the secretary problem could be modified to explain participants’ undersampling bias. It has been further claimed that the robust performance of such an undersampling “cut-off model” might be applicable even to cases where secretary problem assumptions are violated (Todd & Miller, 1999). The current study will test whether this cut-off model indeed can generalise as a theory beyond the secretary problem, to explain bias in economic full information problems.

Although full information problems are more realistic, in the sense that they lack the many restrictive assumptions of the secretary problem, they are not associated with so simple a rule for deriving optimal performance for comparison with human performance and detection of bias. Instead, for full information problems, an algorithm is needed to derive optimal performance, which uses a belief about the generating distribution to search the possible futures in which further options are sampled, to compute an expected reward value for continued sampling, which can be compared against the reward value of the current option (Costa & Averbeck, 2015; Gilbert & Mosteller, 1966). The behaviour of these optimality models, in the guise of an “ideal observer”, can be compared to human behaviour to detect bias. Moreover, as with the aforementioned secretary problem optimality rule, we can also parametrise the full information problem model to theoretically model sources of bias in human performance.

To date, the sampling behaviour of such ideal observers has been compared to that of human participants to identify bias in several (relatively uncomplicated) versions of full information problems. These results suggest that, as in secretary problems, participants have been observed to sample too few options, frequently in economic scenarios (Baumann et al., 2020; Cardinale et al., 2021; Costa & Averbeck, 2015). Our studies reported herein are inspired by the paradigm used by Costa and Averbeck (2015), who found that participants undersampled when compared to a Bayesian version of the optimality model. This bias was reported across economic scenarios including buying a subway ticket, a television, and a diamond ring. Although Costa and Averbeck (2015) and the other studies that show undersampling bias all used numbers to communicate option values, numerous different stimuli are used across the literature to indicate the value of an option (e.g., Baumann et al., 2020; Cardinale et al., 2021; Costa & Averbeck, 2015; Furl et al., 2019; Goldstein et al., 2020; Guan & Stokes, 2020). However, a recent slate of studies on full information tasks has shown results that show oversampling rather than undersampling (Furl et al., 2019; van de Wouw et al., 2022). These new studies employed a number of new methods, which were introduced primarily to adapt the full information problem to picture-based (rather than number-based) domains, including searches for the most attractive face, food or holiday destination, where option values are inherently subjective and personalised. It is possible that any of these methods (which we will discuss in greater detail in the Methods sections herein) could have given rise to oversampling instead of undersampling.

What exactly are the circumstances under do people show this undersampling bias on economic tasks? And what computational mechanisms give rise to this bias? Our current paper directly investigated these questions by systematically manipulating several of the new methods that are associated with studies that show oversampling, while keeping an economic scenario where participants make decisions about smartphone contract prices. Although participants did increase their sampling for sequences containing more options, generally participants maintained a relatively constant number of samples before decision, across most methods manipulations. In contrast, the ideal observer’s sampling behaviour was sensitive not just to sequence length but also to the payoff scheme (i.e., which ranks of choices were rewarded). This suggests that the stimulus domain *per se* (number- versus picture-based), combined with the choice of payoff scheme, may play a role in determining the size of undersampling bias. On a computational level, we found that participants’ sampling decisions on the economic task were best fit by a Bayesian model with a parameterised cost to sample, compared to a number of other computational models that could theoretically explain bias. The proclivity that we observed for participants’ to sample only within a restricted range, regardless of methods features, may therefore arise because participants’ perceive that sampling is an intrinsically costly activity and so they limit how many options they are willing to sample before committing to one option.

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 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), derived from the ratings measured during phase 1. In phase1, participants also had the opportunity to learn the “generating” distribution of option values and thereby establish prior 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 fix the models’ prior on its generating distribution of option values (See *Ideal observer optimality model* section below for more information).

Next, in the optimal stopping task, participants engaged 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 conditions in all studies except Study 3, which compared 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, 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, as an optimality benchmark, 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 the 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 in an ideal observer 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 participants’ sampling was also dependent on payoff scheme. 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 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. 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 prices 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 estimated parameters for all models (OV and SV) would be on the same scales. We implemented the ideal observer objective values model (IO OV) procedure 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 OV 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 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 access to subjective values of options in some conditions, due to the presence of the initial rating phase (Pilot full, Study 1 full condition, Study 1 ratings condition, Study 2 and both sequence length conditions of Study 3). 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 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, 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.

Theoretical models

We implemented the ideal observer model described above to assess the degree to which humans undersample, depending on whether they optimise their choices according to the objective (IO OV) or subjective values (IO SV) of the prices. By definition, the parameter values of an ideal observer model is 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, 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, 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 (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 action values from each model to probabilities using Softmax and then negative log likelihoods were summed 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). 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 S3). The two free parameters per model 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 configured and recovered parameters (Figure S1), although strong correlations were observed for some models (e.g., r = .9 for the cost to sample model, which is the model that will form the basis of our main conclusions). We also found 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). 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 best model would then show the (statistically) lowest BIC mean value. 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)* 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 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. 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 versions 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, but assume 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 Ideal Observer Optimality Model section above). In such a model, participants would undersample if they intrinsically perceived sampling as costly and so adopted a negatively-valued *C,* and would oversample if they perceived sampling as rewarding as so adopted 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 *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 to compute utilities . Negative values of this parameter can 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 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). 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 consideration by the ideal observer. We fixed the logistic slope to equal .2 (on the basis of successful exploratory 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 (i.e., 50, as all option values were normalised from 1 to 100) 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 the *Ideal Observer Optimality Model* section of the *General* *Methods* 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. As in the BV model, the starting value of the midpoint parameter was initialised at 50.

Pilot Studies Methods

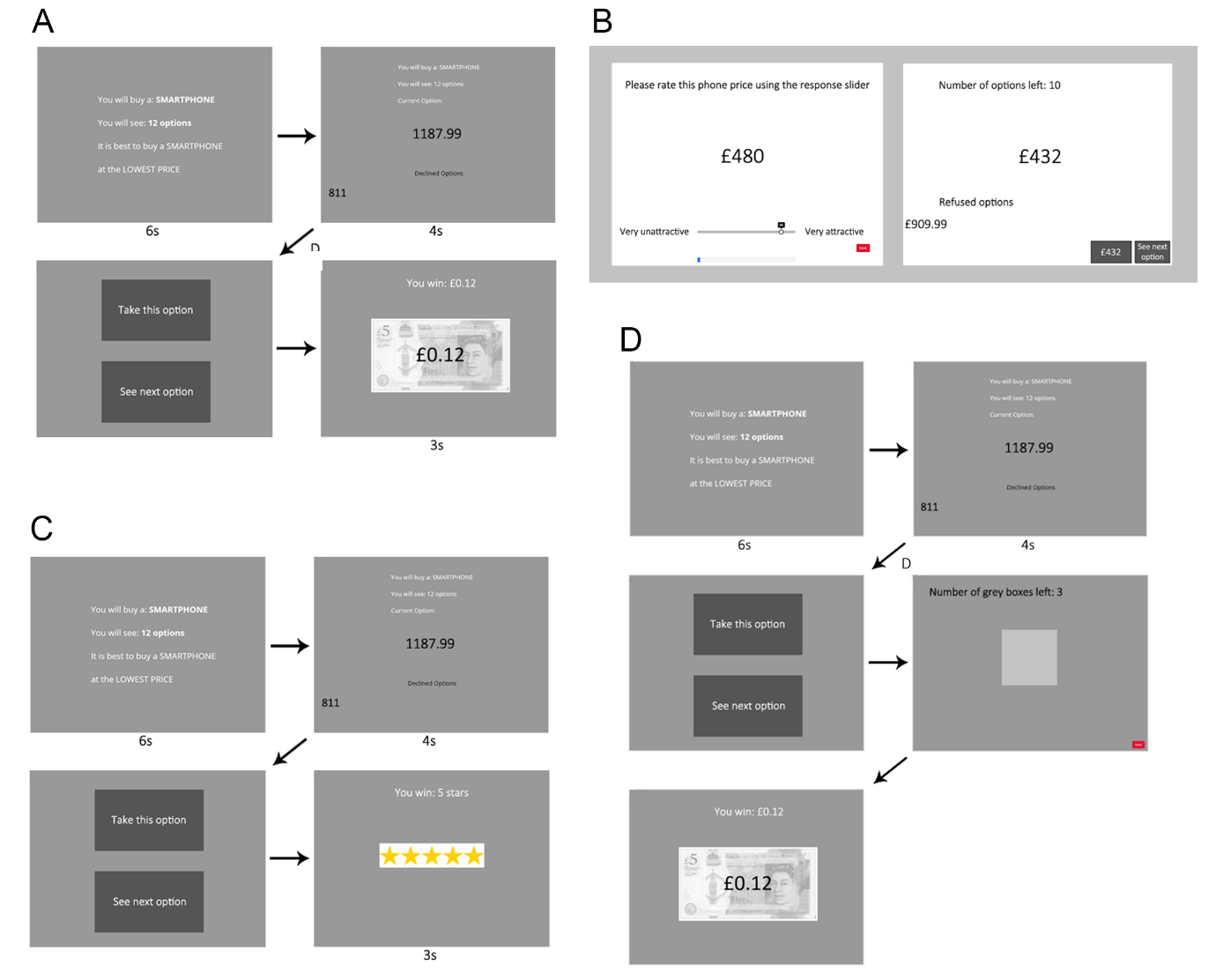
Participants

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 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 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. We thus 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, using seven 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 each used an order of option values that was fixed in advance, so a given sequence’s option values and their order within the sequence was identical for every participant (and corresponding model), 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 also 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.

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). Pilot full was named “full”, as it possessed the full complement of additional methods taken from 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, 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 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.

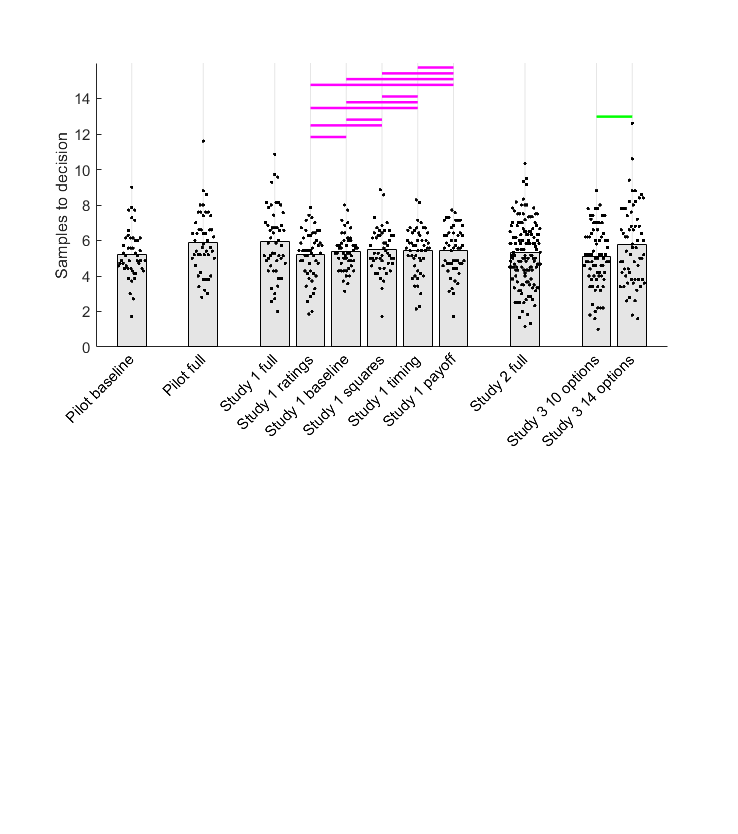
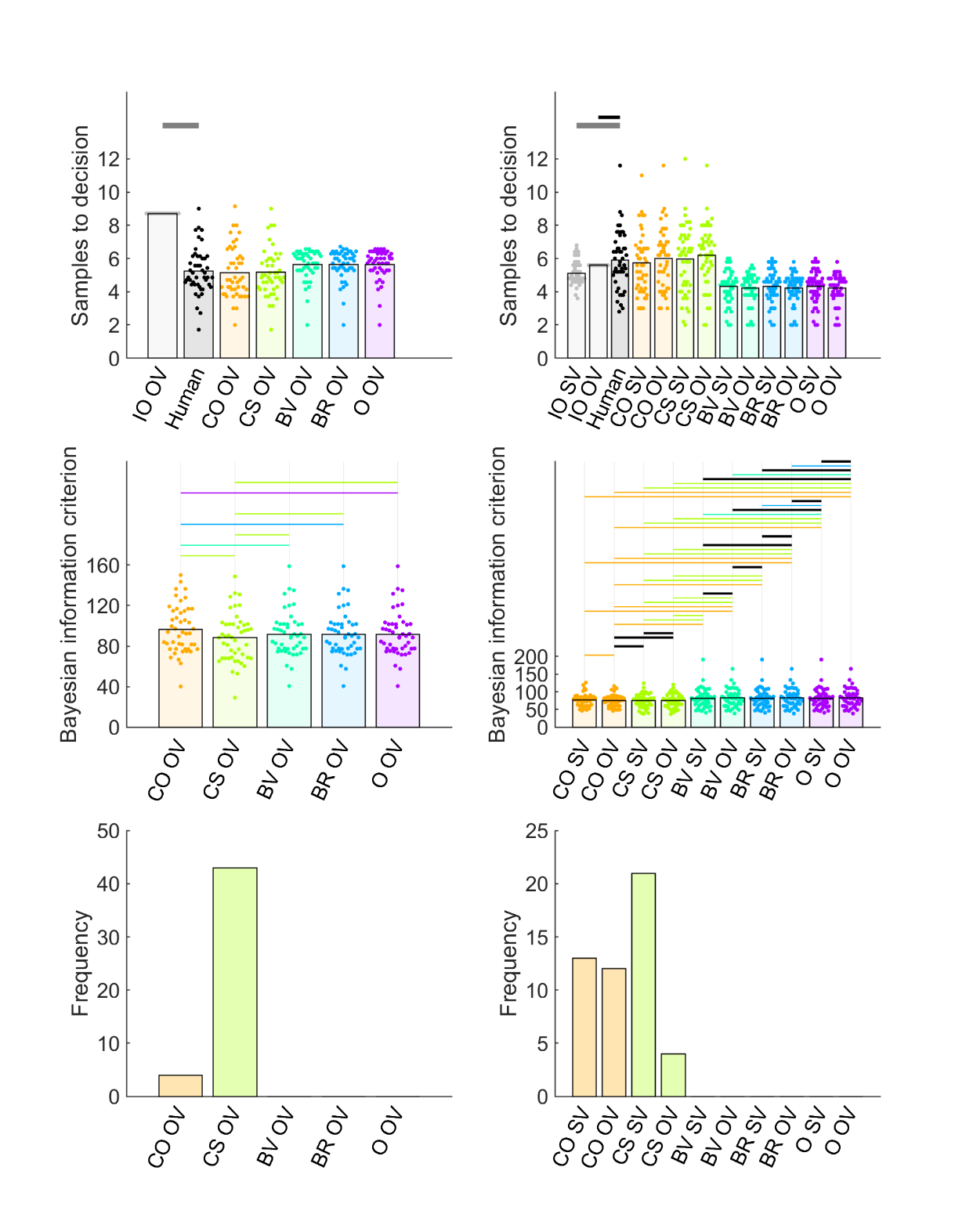


Figure 3 shows results from the comparison of human participants’ sampling (reproduced from Figure 2) with sampling of the ideal observer and theoretical models. As expected, we successfully replicated undersampling in the Pilot baseline condition (Figure 3, left column), where participants sampled fewer options than the ideal observer (Cohen’s *d* = -2.52). Our implementation of the OV version of the ideal observer (i.e., IO OV) is comparable to that of previous studies showing undersampling (Baumann et al., 2020; Costa & Averbeck, 2015).

Figure 3. Model comparison for Pilot baseline (left column) and Pilot full (right column). First and second rows show individual participants as points and bars show their mean values 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). Human participant data are the same as in Figure 2. 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 black when *BF01* > 3 or in the colour corresponding to the better model when *BF10* > 3. For Pilot baseline, the abundant light green lines suggest that CO is the superior model. For Pilot full, 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, which corroborate the results in the second row. 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.



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. Nevertheless, the CS model mostly closely approximated the human participants’ exact mean samples (Figure 3, left column, first row), and, moreover, CS was the best predictor of individual participants’ sampling rates (Figure S5). Model comparison using BIC scores provided strong evidence that the CS model better fit participants than other models. Bayesian pairwise tests (Figure 3, left column, second row), showed that CS had significantly lower (better) BIC values than any other model. And, CS was the best fitting model in almost every individual participant (Figure 3, left column, third row).

For Pilot full, in contrast to Pilot baseline, there was no clear undersampling bias. Instead, the ideal observer sampled at a reduced rate, compared to Pilot baseline, while participants sampled at approximately the same rate (Compare light grey bars across first row of Figure 3). Participants’ sampling was consequently statistically equivalent to optimal sampling, when compared to IO OV (Cohen’s *d* = .17), the same optimality standard used in Pilot baseline. Moreover, when compared to IO SV, there was even a weak (Cohen’s *d* = .45) oversampling effect (Figure 3, top right). It is likely that the IO models reduced their sampling rates in Pilot full, compared to IO sampling in Pilot baseline, because of 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. 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.) change how IO models implement their computations and so cannot change the IO model’s sampling rate. Moreover, neither OV or SV versions of the IO in Pilot full sampled as much as the IO OV model Pilot baseline, nor did IO OV produce undersampling in Pilot full as it did in Pilot baseline, suggesting that the use of subjective versus objective values when modelling is not sufficient to change model behaviour enough eliminate the undersampling bias either. The differing reward payoff functions remain as a possibility to explain why the IO model changes its sampling rate between Pilot baseline and Pilot full (while participants only minimally change their sampling rates), leading to undersampling only in Pilot baseline and not in Pilot full. We will test this possibility more rigorously in Study 1.

The weight of the evidence in Pilot full suggests that, as in Pilot baseline, CS is the theoretical model that best explains participants’ variability in bias. As in Pilot baseline, both OV and SV versions of CS and CO best reproduce participants’ sampling rates (Figure three, top right), although CO fails to reproduce participants’ performance in terms of ranks achieved (Figure S4 in Supplementary Materials). Moreover, individual participant sampling rates simulated by the fitted CS OV and CS SV models better predict participants’ empirical sampling rates better than any other model (Figure S5 in Supplementary Materials). Model comparison using BIC scores (Figure 3, middle right) suggested that the OV and SV versions of both CO and CS produced better average model fits than other models, while differing little from each other. However, when considering the frequency of participants that best-fitted each model (right panel in the third row of Figure 3), CS SV was the best-fitting for more participants than any other model. However, both CO models (CO OV and CO SV in sum) fit approximately as many participants as both CS models together. It is thus possible that some participants in this sample used a CO heuristic, although this model seems less adept at accurately predict participant samples and ranks than the CS model. To disentangle CO and CS contributions to the full condition, we will test replications of this full condition in Studies 1, 2 and 3. To foreshadow these reports here, these studies will agree that the CS model is most predictive of participants’ decisions, while CO models show some irregularities in this regard.

In summary, the optimality IO model sampled more for Pilot baseline than for Pilot full, leading to evidence for undersampling in Pilot baseline but no evidence for undersampling in 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 pilot studies presumably arose due to the differences in their reward payoffs. Both studies showed some evidence that the theoretical CS model well-fit participants’ choices.

Study 1

The paradigm design that we adapted to use in Pilot baseline were taken from Costa & Averbeck (2015) and resulted in findings of undersampling like that study. However, we adapted many of the design features for 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 a study that showed *over*sampling (Furl et al., 2019) and we found that this design eliminated the undersampling bias by changing the sampling rate of the IO optimality model. This pattern raises a distinct possibility that at least one of these methods differences might affect 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 (1) sampling performance of participants and (2) of the OV and SV versions of the IO model. Our first hypothesis was, because participants’ samples to decision were so similar between Pilot baseline and Pilot full, that these methods differences would not substantially change participants’ number of samples to decision in Study 1. Because of the possibility that we might need to interpret null effects (where participants sample at equal rates in different methods conditions), we implemented Bayesian tests using null (equal means) models (Figure 2). Our second hypothesis was, 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 or whether real money was used for incentivization, we do not expect the presence or absence of these task features to affect IO sampling behaviour. Our third hypothesis was, because both IO SV and IO OV both sampled less in Pilot full than IO OV sampled in Pilot baseline, we hypothesise that the use of objective (OV) or subjective values (SV) when computing the IO will have relatively little effect on whether participants undersample. If all of the above hypotheses hold, 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 continue to be 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), 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 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 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. 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. 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 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). 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 Pilot 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 option values in each sequence were fixed in advance, and then these sequences were presented in random order. The condition is “full” in the sense that it collectively implements the methodological features taken from Furl et al. (2019) and shown in Table 1. These features included the two-phase task structure, which was implemented using the same methods as Pilot full and Study 1 ratings condition. The mean (over participants) Pearson’s correlation coefficient between the two ratings for each price collected in the first phase was .87. Other task features included screen timings that are self-paced (as in Study 1 timing), grey squares, which replace the remaining options in the sequence and which participants must page through when a choice is made (as in Study 1 squares). Participants also received no rank-dependent monetary bonus (as in Study 1 payoff). However, Pilot full was in the only condition in Study 1 in which participants were not also instructed to obtain one of the top-three-ranked options in each sequence. Instead, as in Pilot full, participants were instructed to choose the best price possible. 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

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” 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 five comparisons with the full condition, which were statistically inconclusive. Cohen’s *d* values for these comparisons are visualised in Figure S7 in the Supplementary Materials.

We next compared participants’ number of samples against those of the IO optimality models, to evaluate decision bias. 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. Thus, as we predicted, neither participant performance nor IO performance nor the undersampling bias appears sensitive to the presence or absence of grey squares, self-advanced timing or extrinsic monetary reward.

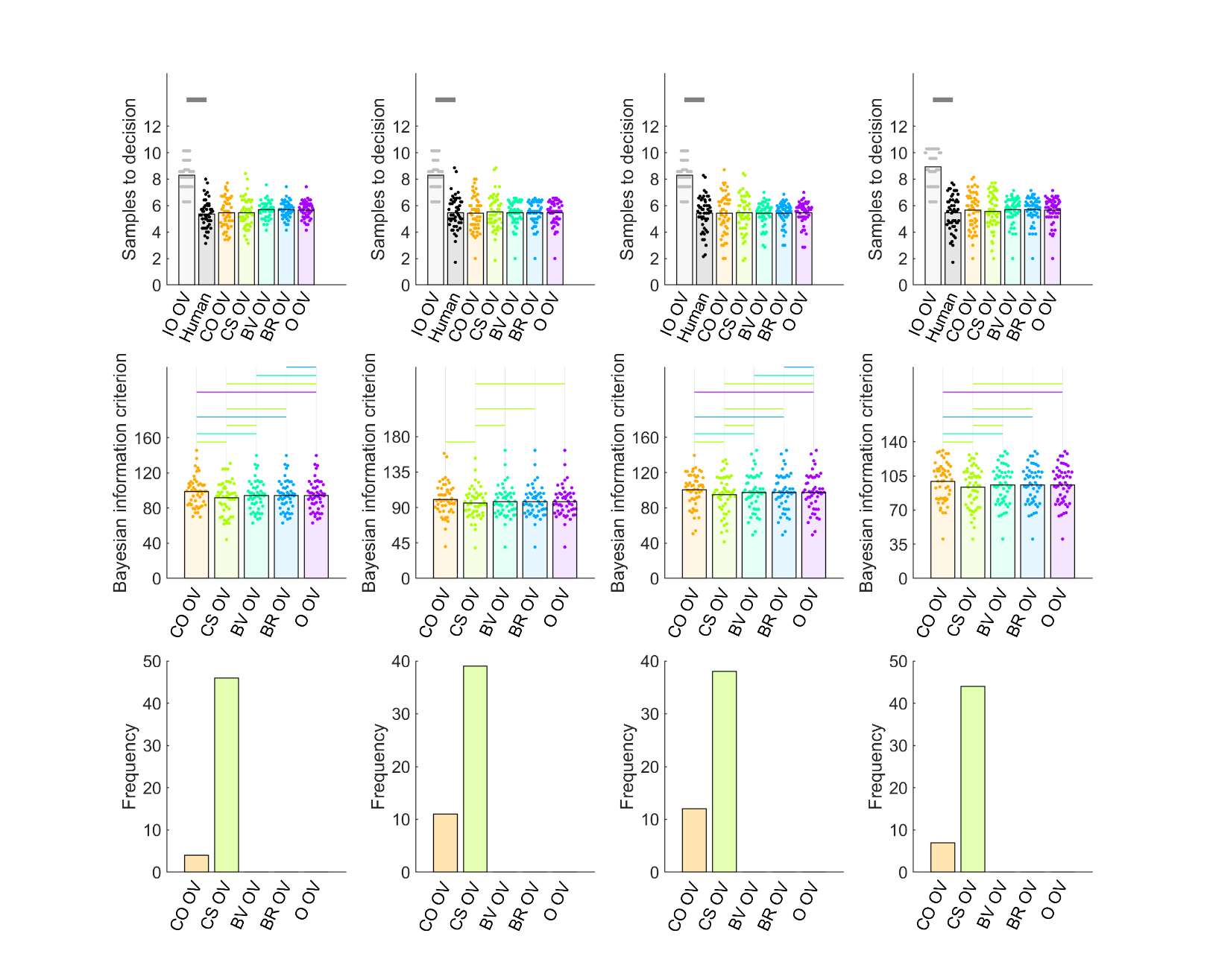


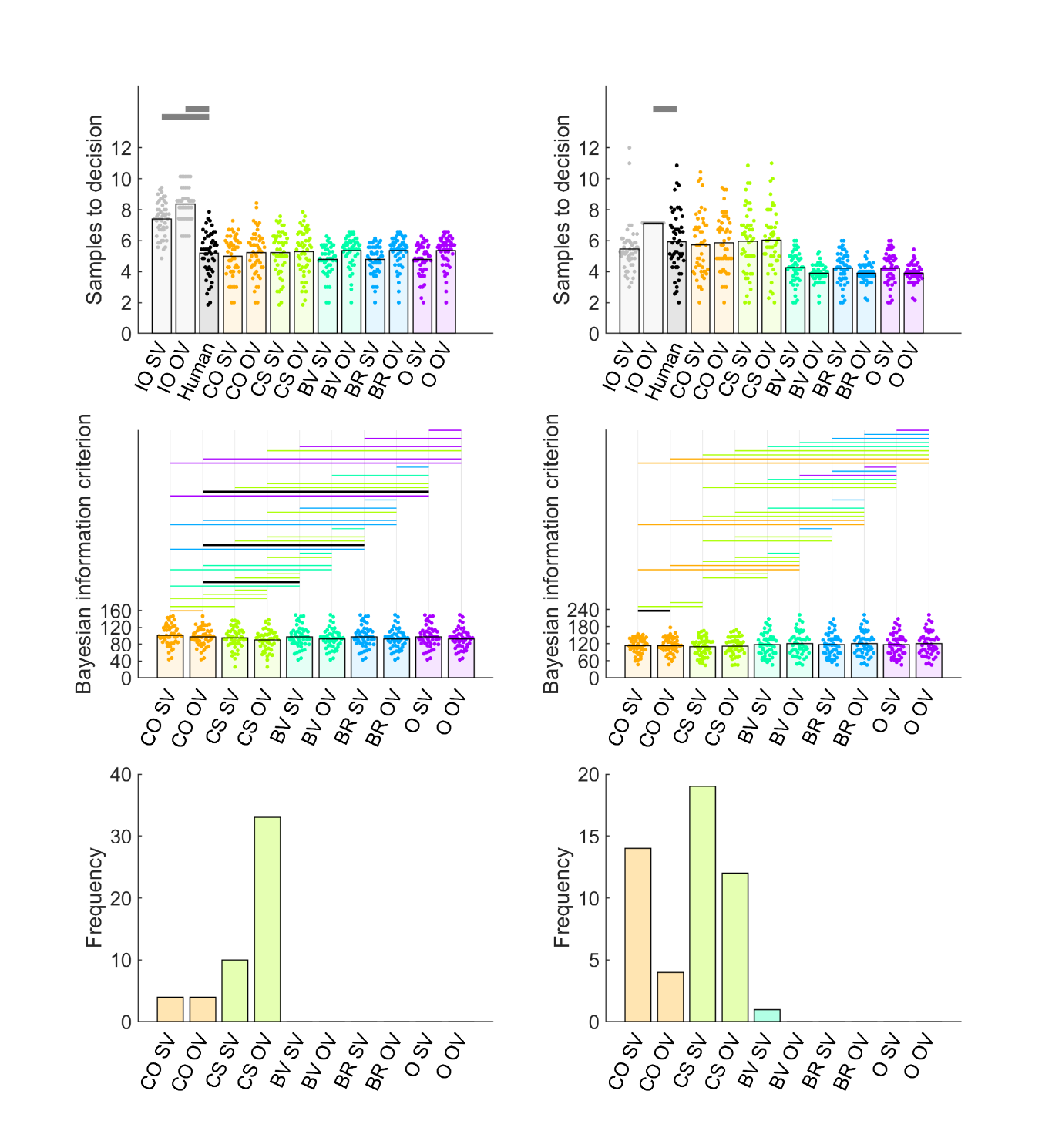
Figure 4. Model comparison for (columns from left to right): Study 1 baseline, squares, timing and payoff conditions. First and second rows show individual participants as points and bars show their mean values. In the first row, horizontal lines above human and IO samples data indicate in thin black when *BF01* > 3 (moderate evidence for equal means) or in thick grey when *BF10* > 3 (moderate evidence for different means). Human participant data are the same as in Figure 2. 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 black when *BF01* > 3 or in the colour corresponding to the better model when *BF10* > 3. The abundant light green lines suggest that CS outperforms other models. The third row shows that the count of participants for which CS was best-fitting is 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 first row of Figure 5 shows, for the two conditions with an initial rating phase (ratings and full), Bayesian test results comparing participant sampling with model sampling by IO SV and IO OV. Sampling bias in the rating condition (left column) showed undersampling for both IO OV (Cohen’s *d* = -1.72) and IO SV (Cohen’s *d* = -1.20). However, the participant versus ideal observer comparison for the full condition more closely resembled the results of Pilot full (right column), in which the ideal observer did not sample as much as in the other conditions. Even though IO OV sampled about one option less on average in full than in ratings, participants still statistically undersampled in full compared to IO OV (Cohen’s *d* = -0.61). In contrast, the difference in sampling between participants and IO SV was small enough to be statistically inconclusive (Cohen’s *d* = 0.19).

The one remaining methodological difference between the IO models in Study 1 full and those of the remaining conditions (baseline, squares, timing, payoff, rating), which 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). Concomitantly, human participants sample at more or less the same rate regardless of the payoff scheme or any of the other methods features that we tested.

Note that the results of Pilot full and Study 1 full are inconsistent enough that they collectively can’t clearly indicate the direction of bias (if any). The number of samples for IO models is much more similar to that of human participants in these conditions and the IO sampling rate means appear to vary in the vicinity of those of the participants from study to study. We will resolve this issue and more precisely measure bias in Study 2 by implementing a full condition with some improved design elements and a statistically better-powered sample size.

Figure 5. Model comparison for Study 1 rating condition (left column) and full condition (right column). First and second rows show individual participants as points and bars show their mean values. In the top row, horizontal lines above human and IO samples data indicate in black when *BF01* > 3 (moderate evidence for equal means) or in grey when *BF10* > 3 (moderate evidence for different means). Human participant data are the same as in Figure 2. 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 black when *BF01* > 3 or in the colour corresponding to the better model when *BF10* > 3. 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 better 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 produced similar results. Only CS models closely approximated participants’ mean number of samples to decision (Figures 4 and 5, first row). These CS models closely approximated participants’ mean rank of chosen price, with especially poor performance on the part of CO models (Figures S8 and S13, first row). 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 S9-S12, S14, S15). CS OV and CS SV models always best-fit participants’ behaviour (Figure 5, second and third rows), resulting in better BIC values than other models on average (BF10 > 3) and best-fitting many more participants than any other models. The results in Study 1 full were similar, though the analysis of the number of participants who were best-fit by each model (Figure 4, third row, right panel) showed somewhat more contribution of the CO heuristic than for the other Study 1 conditions.

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 most methods details. Indeed, they are not even programmed with information about whether there are grey squares, etc.) and so could not have shown such effects. However the IO model is programmed with the payoff scheme and duly appears to sample more in conditions 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). There was relatively little consistent differences between the behaviour of the SV and OV versions of either the IO or theoretical models.

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 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 a direct assessment of bias in this condition difficult to determine with high precision. 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 (in Pilot full and Study 1 full) to 151 in Study 2 full. Also, 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 Prolific participant recruitment platform Prolific.

Procedures.

The study was developed in Javascript jsPsych 7.3.1 (de Leeuw et al., 2023), which was used to reproduce 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. The mean (over participants) Pearson’s correlation coefficient between the two ratings for each price was .85. Then participants 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 was 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 (Figures 1 and 6). From Figure 6, which shows Bayesian pairwise test results comparing participants’ sampling that that of the two ideal observers (threshold 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). These are somewhat different results from those observed in Pilot full and Study 1 full, which used a more limited set of sequences and a third of the participants. It is also noteworthy that Study 2 found no evidence for *over*sampling, either when comparing to IO SV or IO OV. This suggests that economic-based tasks cannot (reliably) produce oversampling even when copying the same methods and ideal observer model as used in picture-based tasks, which do show oversampling (Furl et al., 2019; van de Wouw et al., 2022). We conclude that oversampling in those studies must not arise specifically from any of the methods details we consider here but rather from the content domain of these studies, which used pictures instead of prices (e.g., attractive faces, foods, holiday destinations).

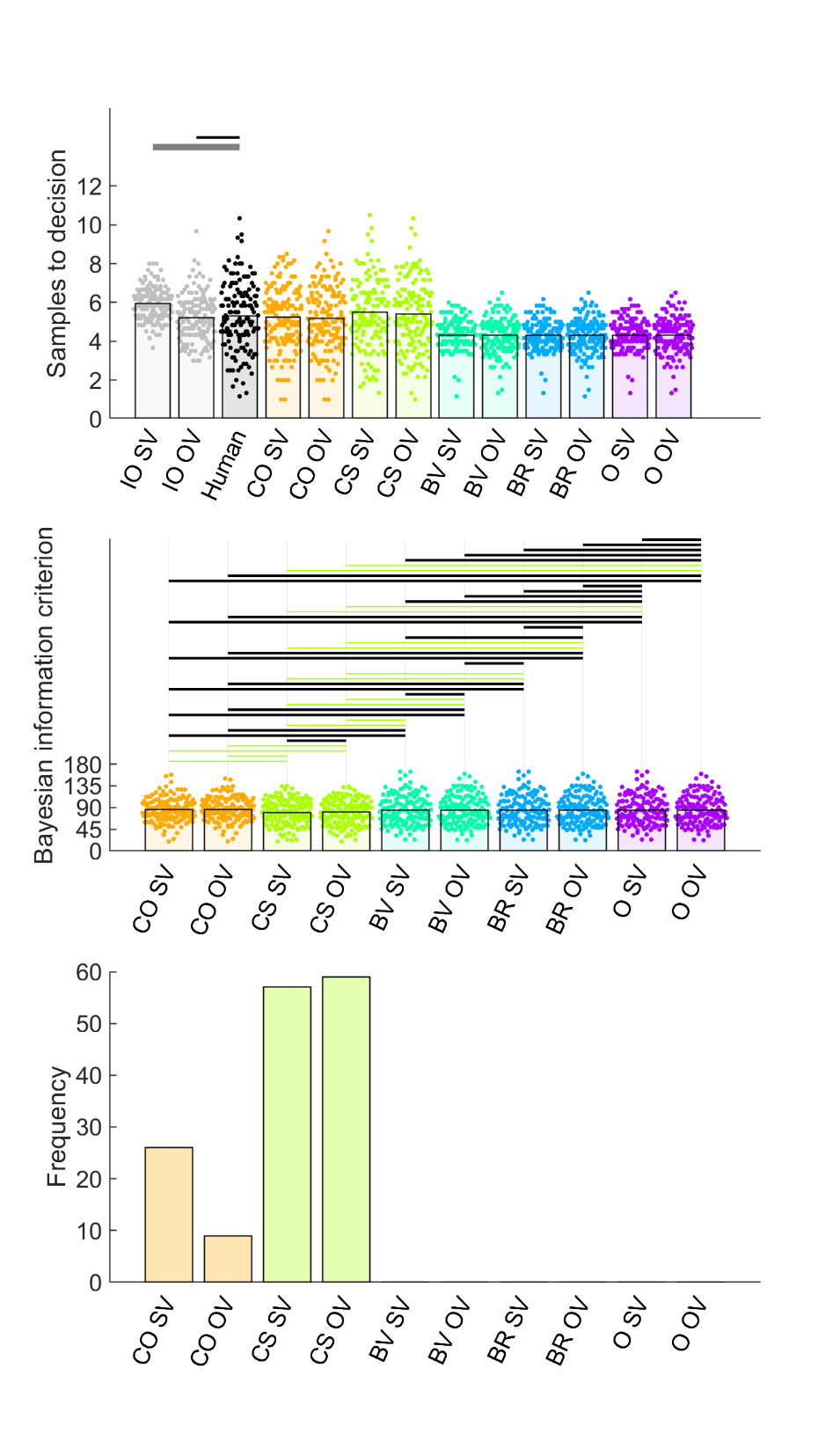


Figure 6. Model comparison for Study 2 full. First and second rows show individual participants as points and bars show their mean values. In the first row, human and IO samples are indicated by thin black horizontal lines when *BF01* > 3 (moderate evidence for equal means) or thick grey lines when *BF10* > 3 (moderate evidence for different means). Human participant data are the same as in Figure 2. 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 CS was the best-fitting model for the most participants. 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 variability in participant bias that we observe within the economic / price domain. Study 2 confirms the results we found throughout our studies reported herein, with the best evidence favouring CS as the best model for the average participant. The SV and OV CS models both: (1) reasonably reproduced participants’ mean number of samples to decision (Figure 6, first row) and ranks (Figure S6); (2) simulated sampling behaviour that was more highly correlated with individual participants’ number of samples than any other model (Figure S17); (3) showed statistically better BIC values than every other model (Figure 6, second row); and (4) best-fit more individual participants than another model.

Study 3

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 increase how much they sample under this incentivisation scheme. Our theoretical modelling fitting results so far suggest that participants’ sampling is controlled by a perceived intrinsic cost to sample, which would indeed limit participants from increasing how much they sample, relative to an ideal observer, which operates under the ground truth of no sampling cost and so is freer to increase its sample rate.

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 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. 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 hypothesise that we will replicate Costa & Averbeck (2015): longer sequences will lead participants to increase their number of samples but will also exacerbate their undersampling bias. Moreover, we 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 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.

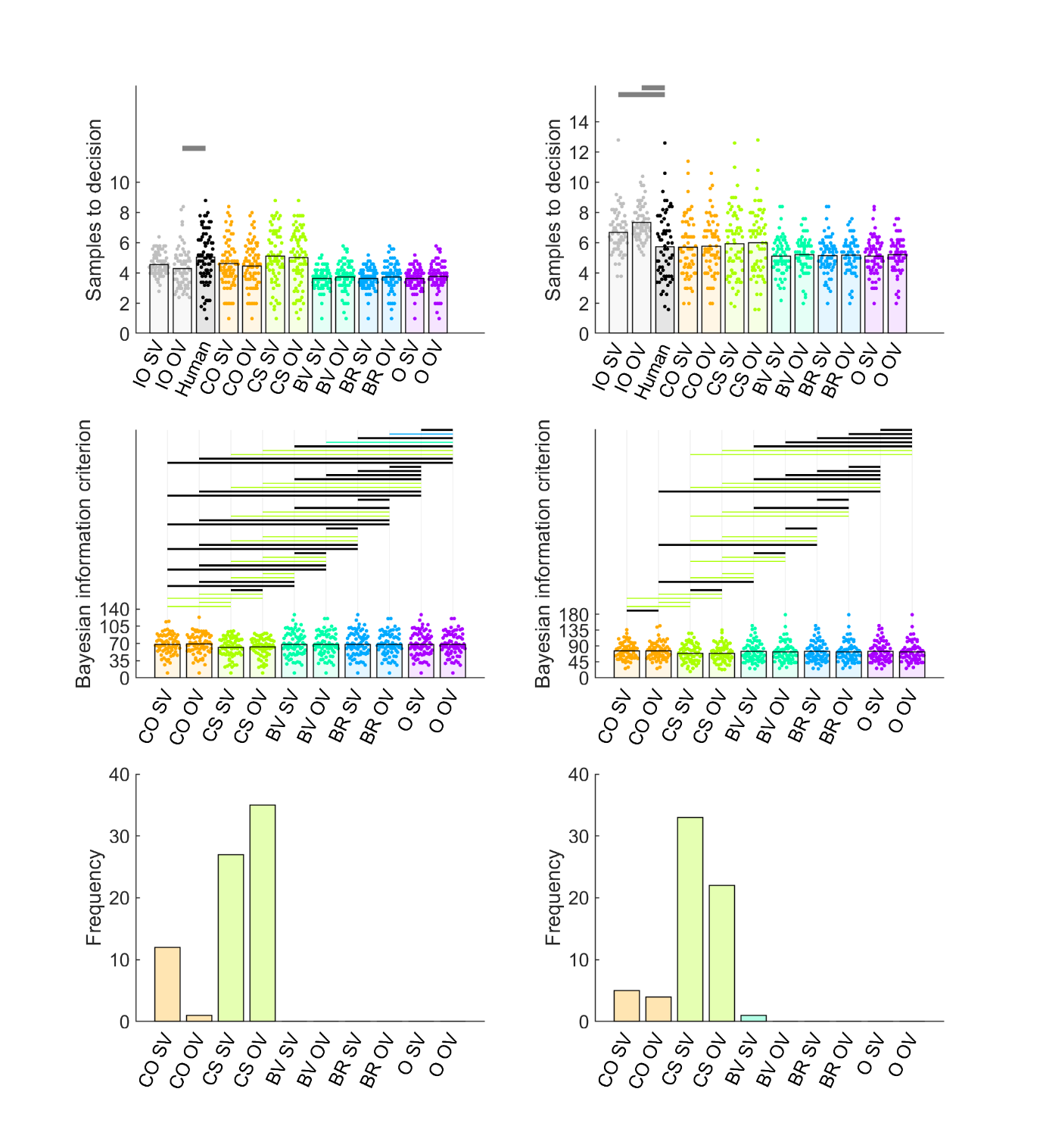
Study 3 Results.

Figure 2 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 in Figure 7 suggest that, at 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 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’ ranks (Figure S18, first row) and the two CS models outperformed all other models when predicting individual participant sampling (Figures S19 and S20). The CS models exhibit 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 using 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, which discourages them from increasing their sampling, and can lead to increasing undersampling bias, as sequences lengthen.

Figure 7. Model comparison for Study 3 10 options condition (left column) and 14 options condition (right column). First and second rows show individual participants as points and bars show their mean values. In the first row, human and IO samples are indicated by black horizontal lines when *BF01* > 3 (moderate evidence for equal means) or grey lines when *BF10* > 3 (moderate evidence for different means). Human participant data are the same as in Figure 2. 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.



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 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.

The undersampling bias was nearly eliminated in the “full” condition, a finding we were able to replicate (Pilot full, Study 1 full, Study 2, Study 3). Yet this is not because participants changed their behaviour much in this full condition, compared to a variety of other conditions in Study 1, which manipulated the presence or absence of various task features from the full condition (van de Wouw et al., 2022). Instead, the IO model sampled more than participants did in every condition except for the full condition, where it reduced its sampling rate to become more like participants’ (lower) sampling rate. This full condition was the only condition where participants were instructed to simply maximise the option value of their choices, as opposed to obtain one of the top-three ranked options. Indeed, 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 values in the IO model. Thus, we must conclude that only the reward payoff scheme used in the full condition could have reduced the IO model’s sampling rate and modulated undersampling bias. Indeed, because the IO model is willing to increase its sampling rate to the one appropriate for its payoff scheme, and participants are not so willing, 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, 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 subjectively regard increased sampling as intrinsically costly. Our model comparison also eliminated several other competing theoretical accounts for participants’ biased sampling.

Whence does this does this suboptimal cost to sample arise? Although this is still an open question for future research, the data we report allows us to reject some possibilities. Our studies imposed no extrinsic penalty, monetary or otherwise, for extra sampling. Attempting to save time by stopping searches earlier is unlikely to fully explain participants’ behaviour either. Some conditions tested herein used atask feature designed to prevent participants from shortening the study by sampling less: the remaining options after choice in each sequence were replaced with grey squares, which participants were required to page through to finish each sequence, even after choosing an option. Moreover, some conditions also manipulated the screen timing: fixed screen timings add several seconds extra time cost per sample, compared to a condition with fixed screen timing, where a participant can quickly sample again at the tap of a key without waiting. Neither the grey squares nor the screen timing manipulations affected participants’ sampling rates (nor that of the IO model), null findings which are bolstered here by evidence from Bayesian null effects tests. Working memory is not directly a cause either. Participants had all rejected options available on the screen in front of them and so were prevented from forgetting rejected options.

Nevertheless, there are still open explanations, about which we may speculate. Although participants cannot “forget” rejected option values, they may nevertheless suffer limited bandwidth for integrating new evidence from samples. That is, participants might find that performing predictive computations based on more than four or five options to be effortful, even if the values themselves are not forgotten. Moreover, inspection of the ranks that participants achieved with their choices (a measure of their choice accuracy), shown in the first rows of Figures S4, S8, S13, S16 and S18 suggests that participants’ choices closely approximated performance of IO, even though participants exerted less effort to sample. Others (Todd & Miller, 1999) have argued that there is adaptive value for participants to reduce sampling to suboptimal levels if they can maintain “satisficing” levels of choice outcome performance, with less exertion or intrinsic sample cost – a form of heuristic decision making. Such might be the case here.

Future research may wish to ascertain how and why participants adopt certain intrinsic cost to sample values by fitting and comparing theoretical computational models in paradigms in which participants oversample, rather than undersample. Although we were unable to reliably induce participants to oversample in the present work (and instead 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). Thus, we hypothesise that a CS model might explain sampling bias in paradigms using images to present option values if participants find sampling through numbers to be intrinsically costly, but searching through images intrinsically rewarding, like the supposedly “addictive” activity of sequentially accepting or rejecting images on Tinder, a popular mobile application for dating. If it is participants’ attitudes towards their experiences of the options themselves that controls their sampling bias (and thereby the reward value of their choices), then this would have profound implications for how people engage with sequential searches in real-world contexts. For example, sellers in some domains may choose how they present shopping options to consumers to encourage fast yet satisficing searches.

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 a number of other parameterised versions of this model (Supplementary Text A and Figures S1-S3). 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’ behaviour 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 (e.g., the cost to sample) 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. Similarly, our model comparison agree that variable-threshold models (i.e., the CS OV and CS SV models) better fit participant choices compared to the relatively rigid thresholds of the CO OV heuristic.

In summary, we show that, participants’ sampling behaviour on optimal stopping tasks is relatively insensitive to most methodological manipulations. In contrast, the ideal observer (which reflects optimal performance) is relatively more sensitive than participants to at least the payoff scheme and the sequence length, 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 perceive that continued sampling can become intrinsically aversive.

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%3c391::AID-BDM359%3e3.0.CO;2-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

For each theoretical model (CO, CS, BV, BR and O), we simulated choice data for four levels of the key theoretical parameter (Figure S1, x axis). We selected the range of simulated key theoretical parameter values to produce sampling rates to span the range of possible samples to decision – that is, between approximately one and 12 samples (Figures S2 and S3). For each of these parameter levels, we simulated 50 participants, similar to the smallest sample size collected herein.

We presented option values to every configured model to derived simulated take option or sample again choices associated with each configured parameter value. For each simulated participant, we created an option-generating distribution of 90 option values, randomly-generated from a Gaussian distribution with mean 40 and standard deviation of 20. These values were chosen to approximate the mean and standard deviation of the price and empirical participants ratings distribution (both scaled to range between 1 and 100) derived from Study 2. A small amount of jitter (sampled from a Gaussian with mean 5 and standard deviation 5) was added to the mean and standard deviation for each simulated participant, to make generating distributions personal to each simulated participant, as in the empirically-collected data. The ensuring distribution was truncated within the range of 1 to 100. We populated five sequences per simulated participant with twelve options per sequence from these participant-specific generating distributions.

We fitted 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).

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. We compared configured parameter values (horizontal axes) for simulated participants (each shown as an individual scatter point) against parameter values estimated (vertical axes) after fitting models to decisions simulated using configured parameter values (horizontal axes). 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. 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.

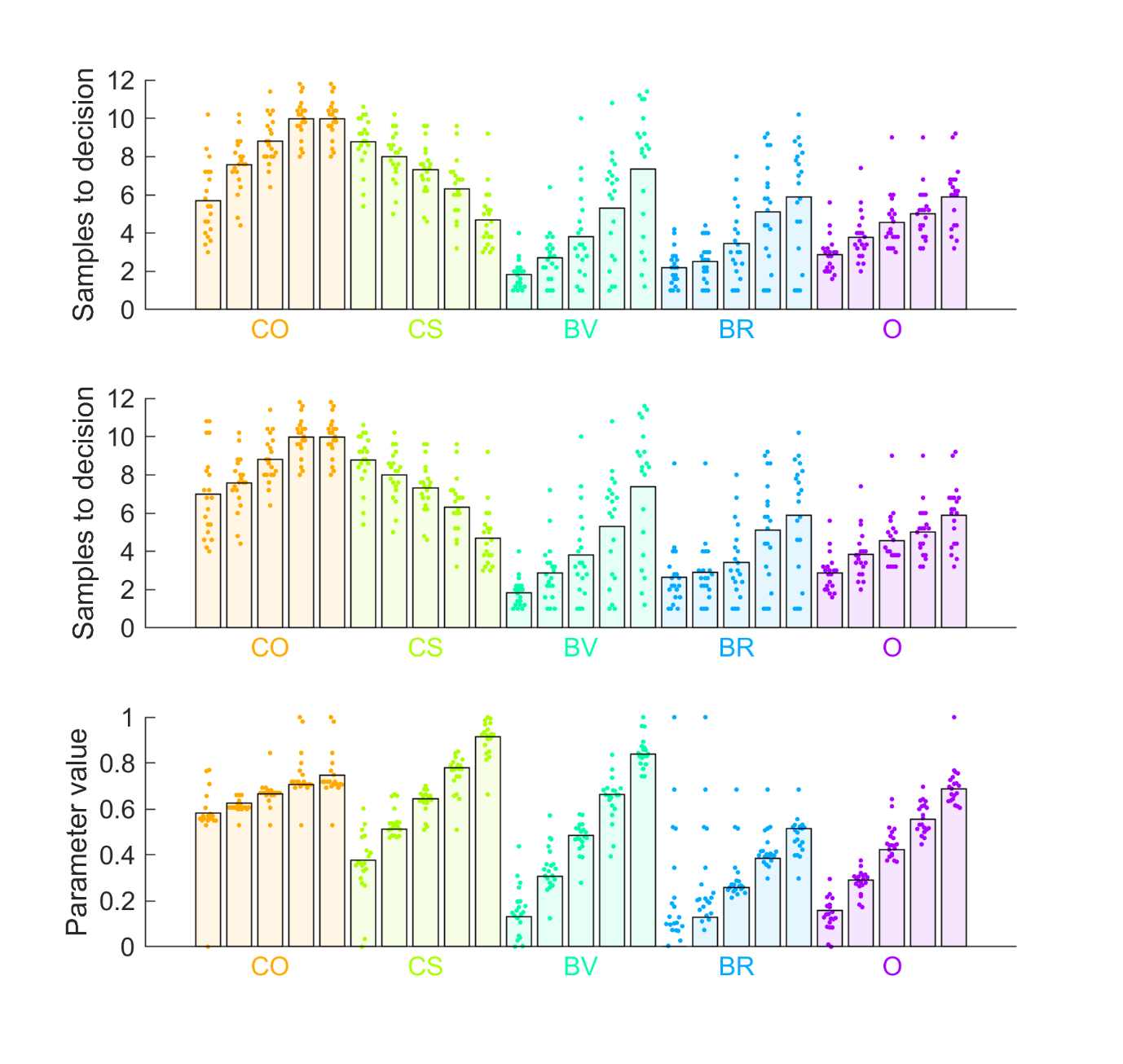
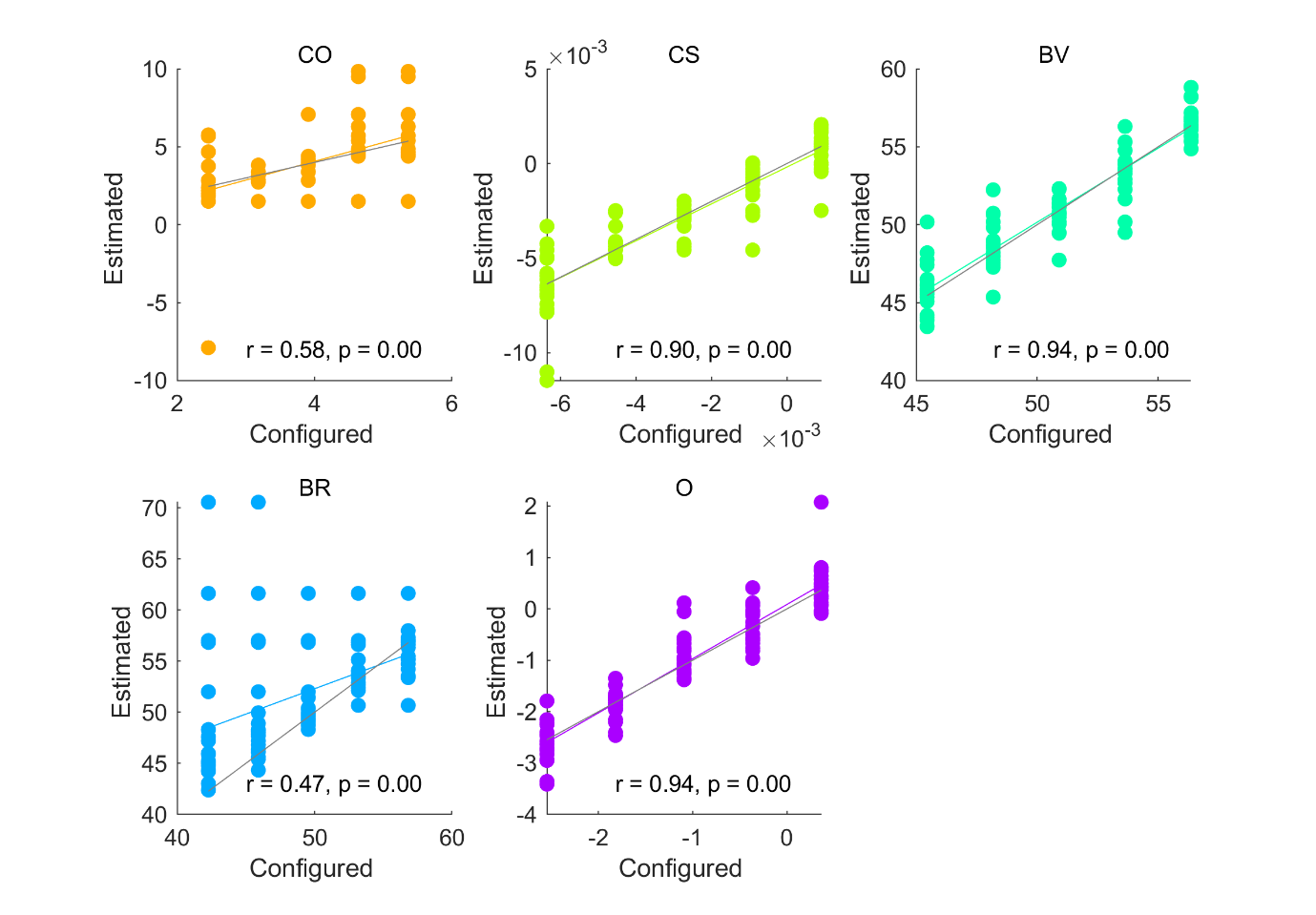


Figure S2. (Top panel) Mean sampling rates (bars) for simulated participants (points), for each configured parameter value for each model. Note that varying configured parameter values leads to systematic increase or decrease in simulated sampling rates. (Middle panel) Models were fitted to the data 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). One extreme outlier was removed from CO model plotting to improve visibility (See caption, Figure S1). Parameter values were normalised to be between 0 and 1 for each model to facilitate plotting of all model parameters on the same scale. Abbreviations: CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism.

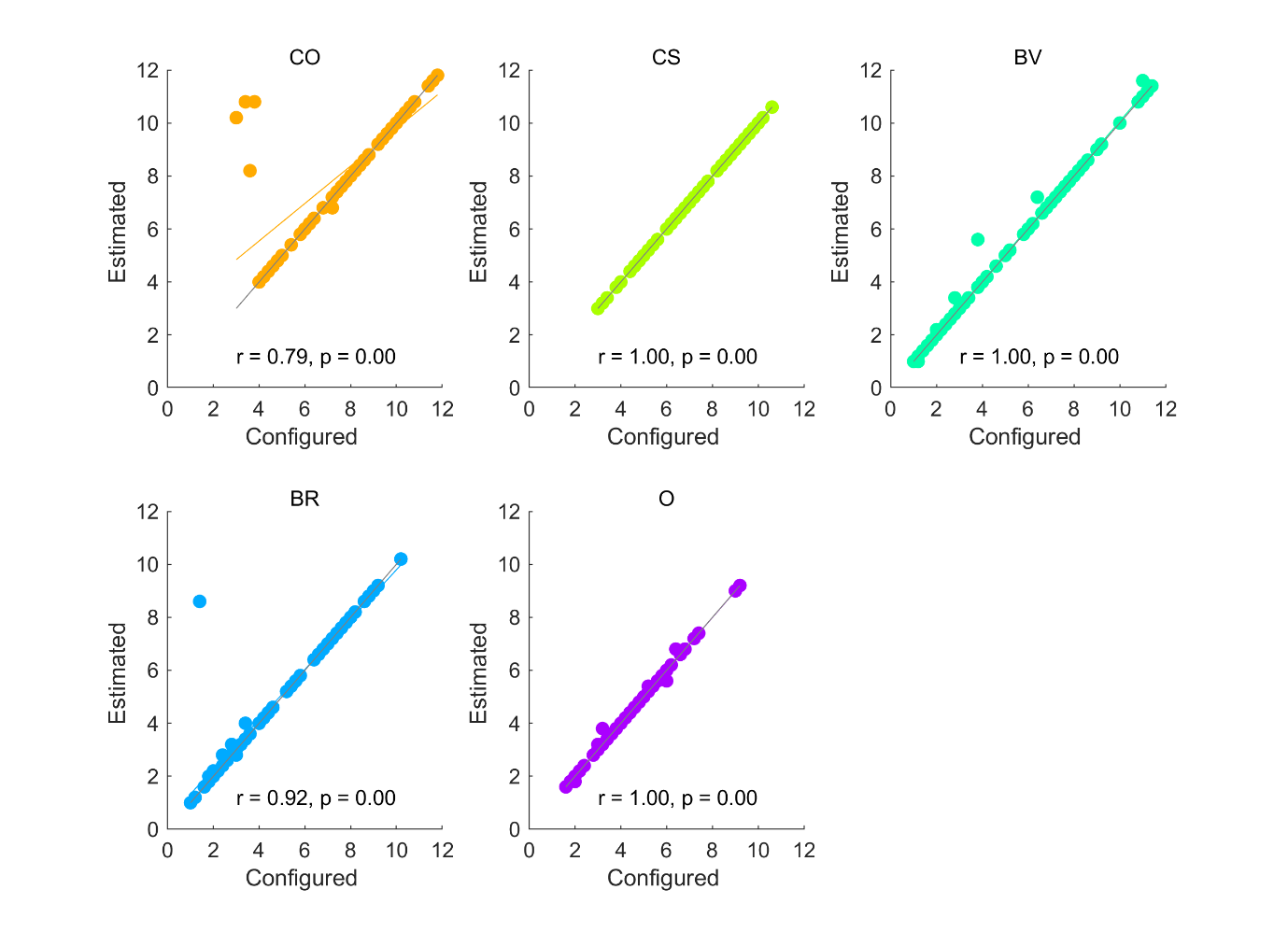


Figure S3. Sampling rates simulated using configured parameters (horizontal axis) are plotted against sampling rates computed from recovered (estimated) parameters. Recovered parameters are highly suitable for reproducing the sampling choices that they are intended to model. The grey diagonal indicates 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.

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

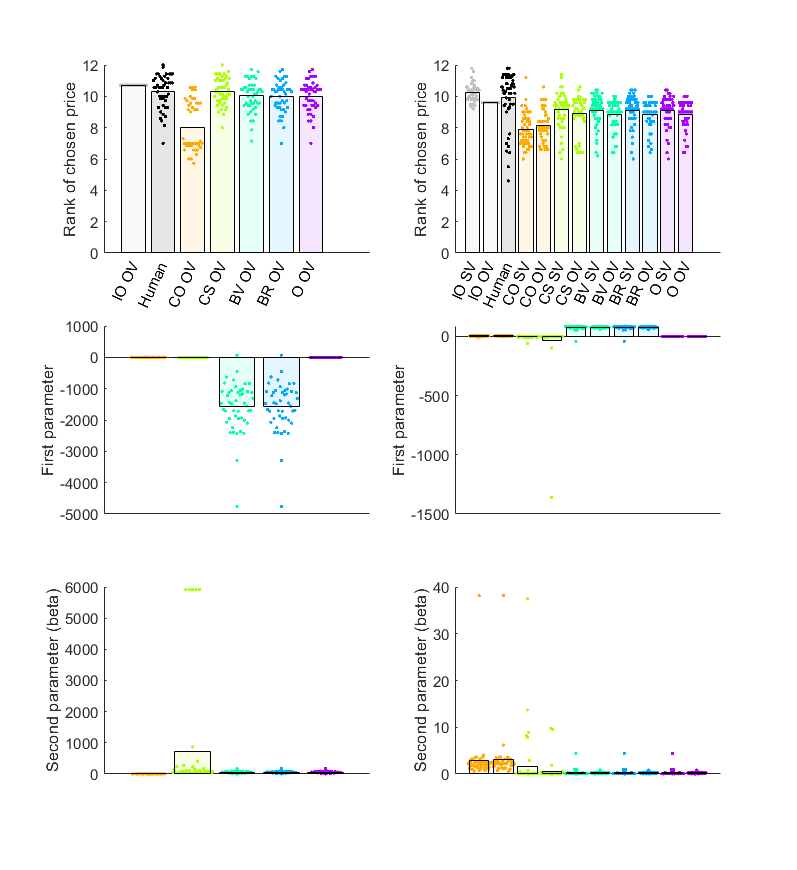


Figure S5. Linear relationships between human participant sampling in Pilot baseline versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured lines represent regression lines, with corresponding *R2* printed on plot. 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 lines represent regression lines, with corresponding *R2* printed on plot. The 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.

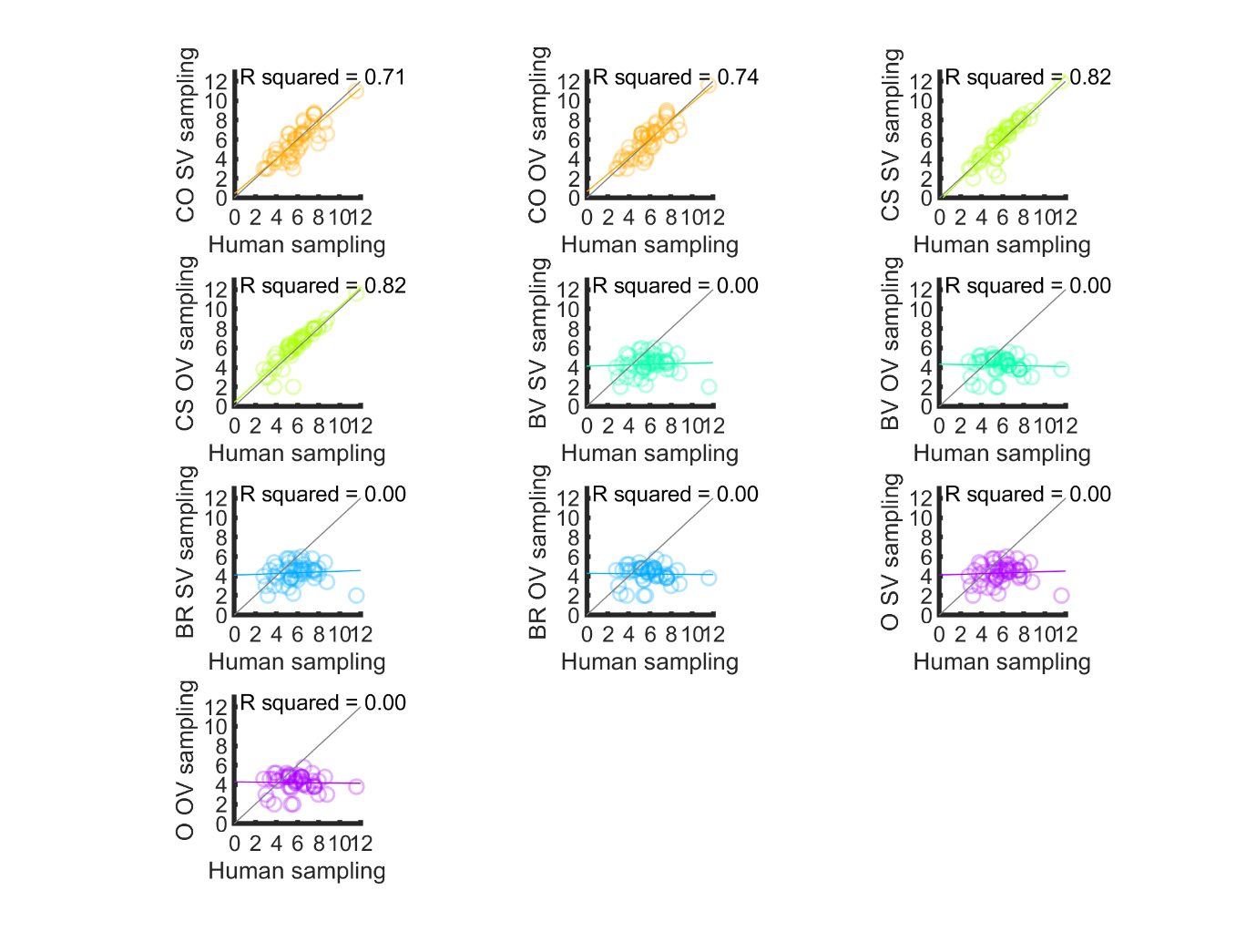


Figure S7. Cohen’s d effect sizes for pairwise comparisons in Study 1.

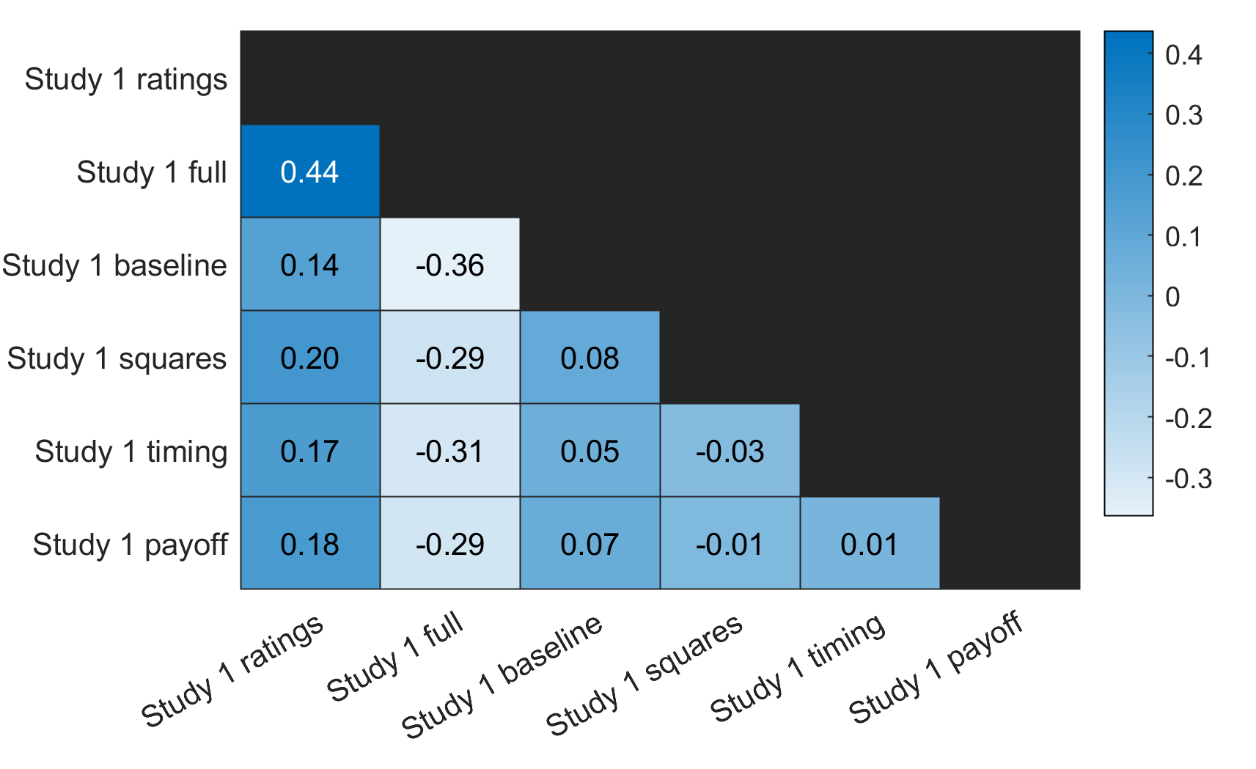


Figure S8. Model comparison for (columns from left to right): Study 1 baseline, squares, timing and payoff conditions. Top and middle rows show individual participants as points and bars show their mean values. The top row shows ranks of chosen prices. The second row plots the “first” or theoretical 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 S9. 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.

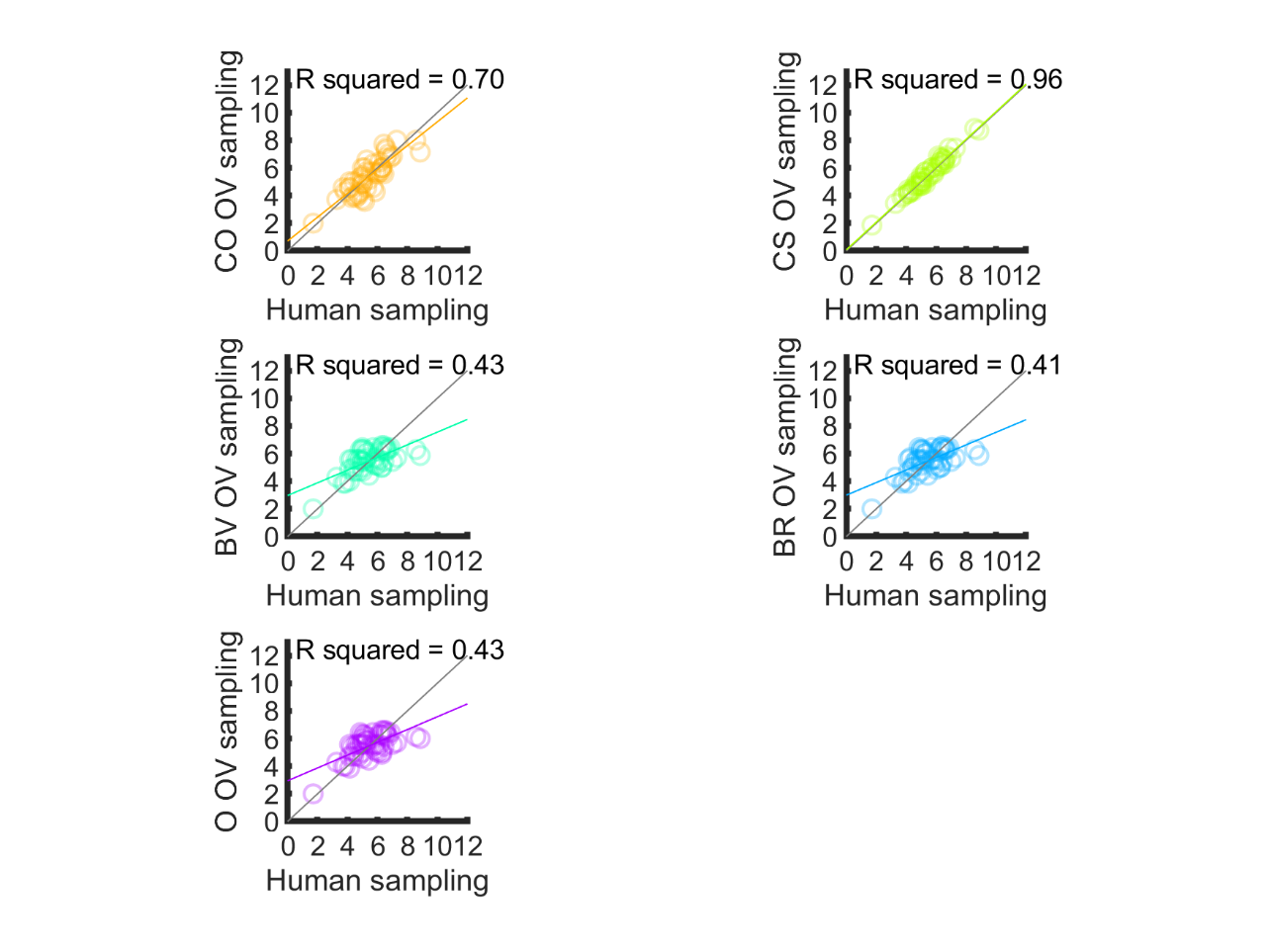
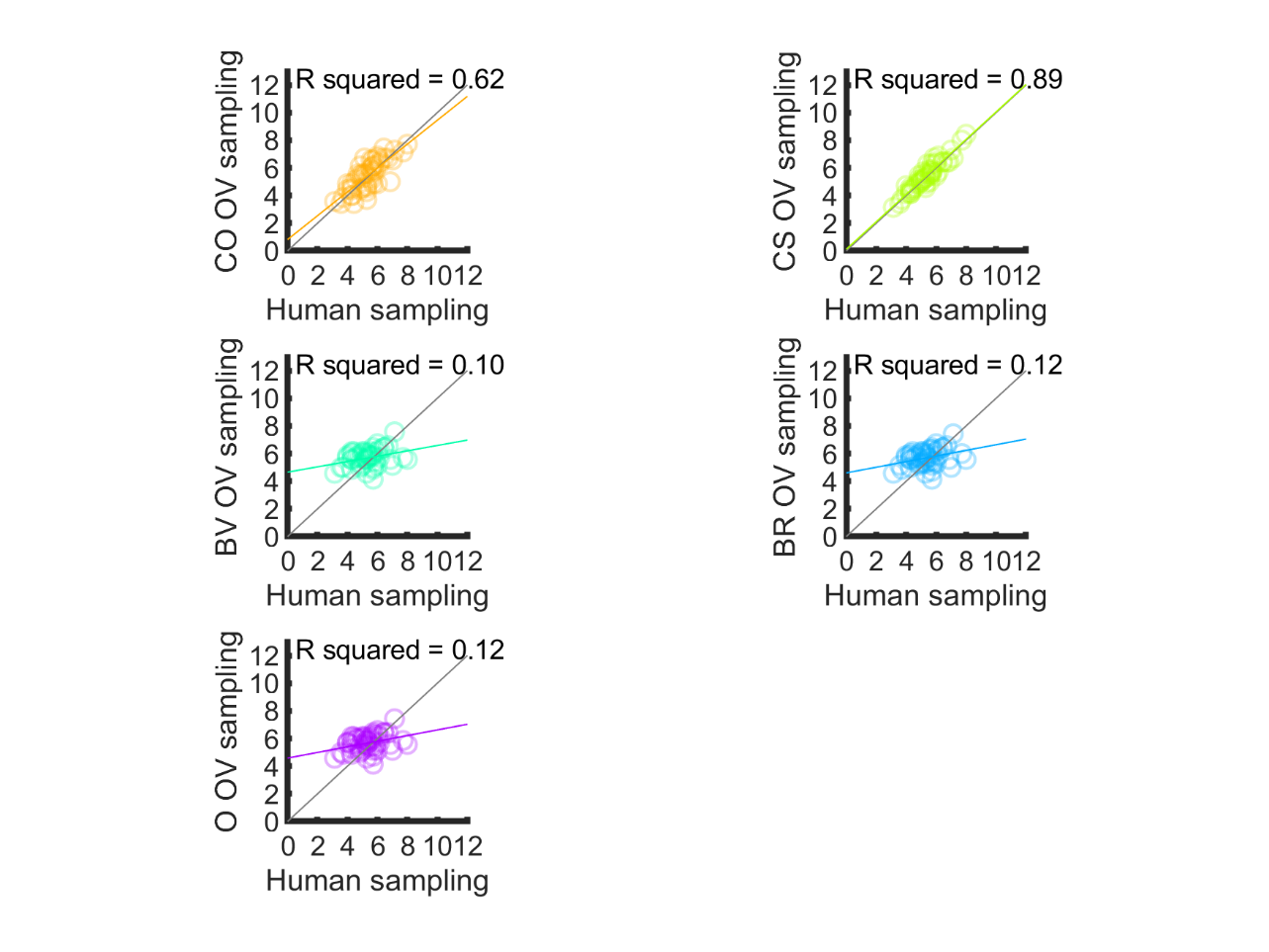


Figure S10. Linear relationships between human participants’ sampling in Study 1 squares condition versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured line represents the regression line, with corresponding *R2*printed 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.

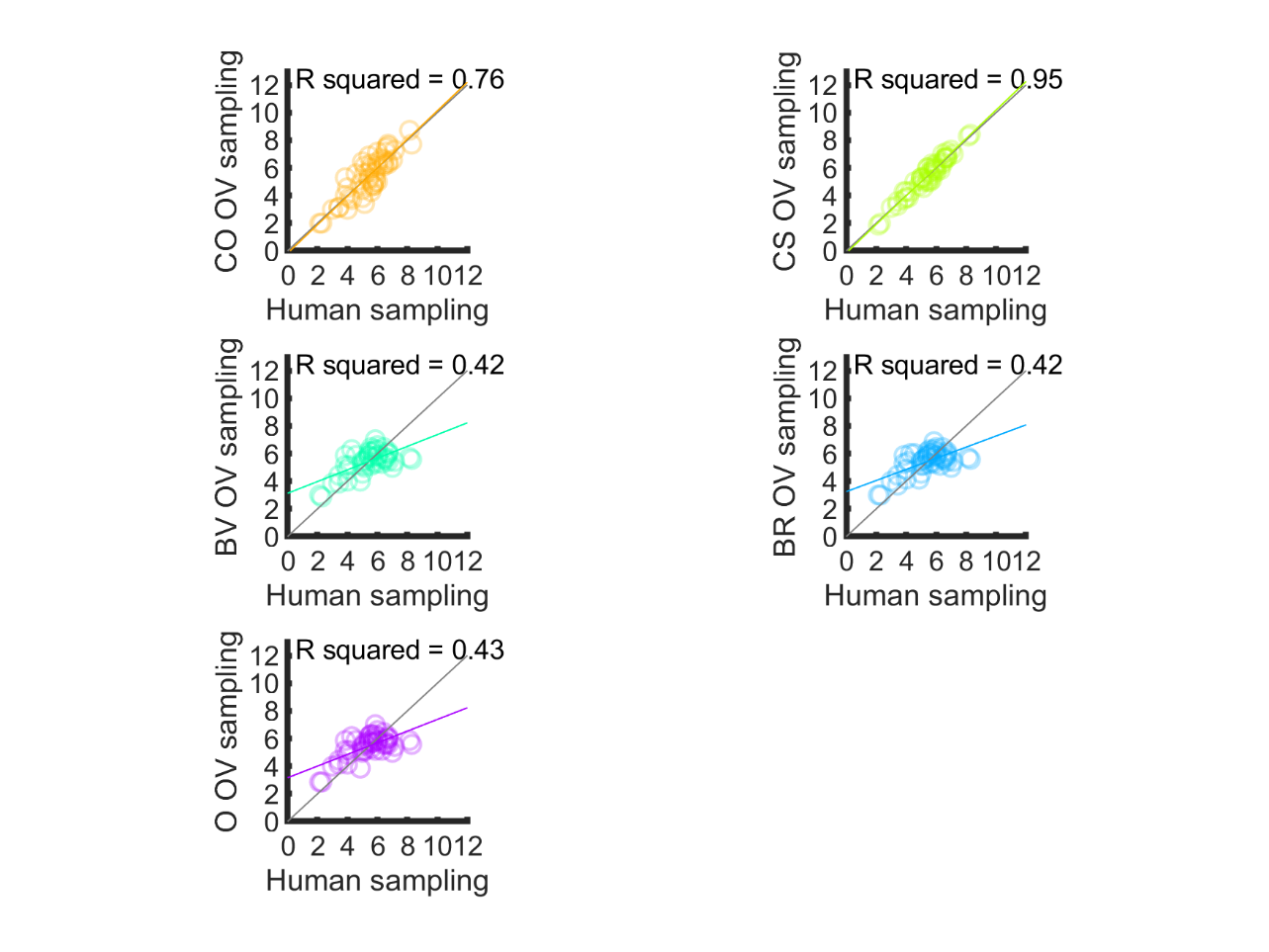


Figure S11. Linear relationships between human participants’ sampling in the Study 1 timing condition versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured line represents the regression line, with corresponding *R2* 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.

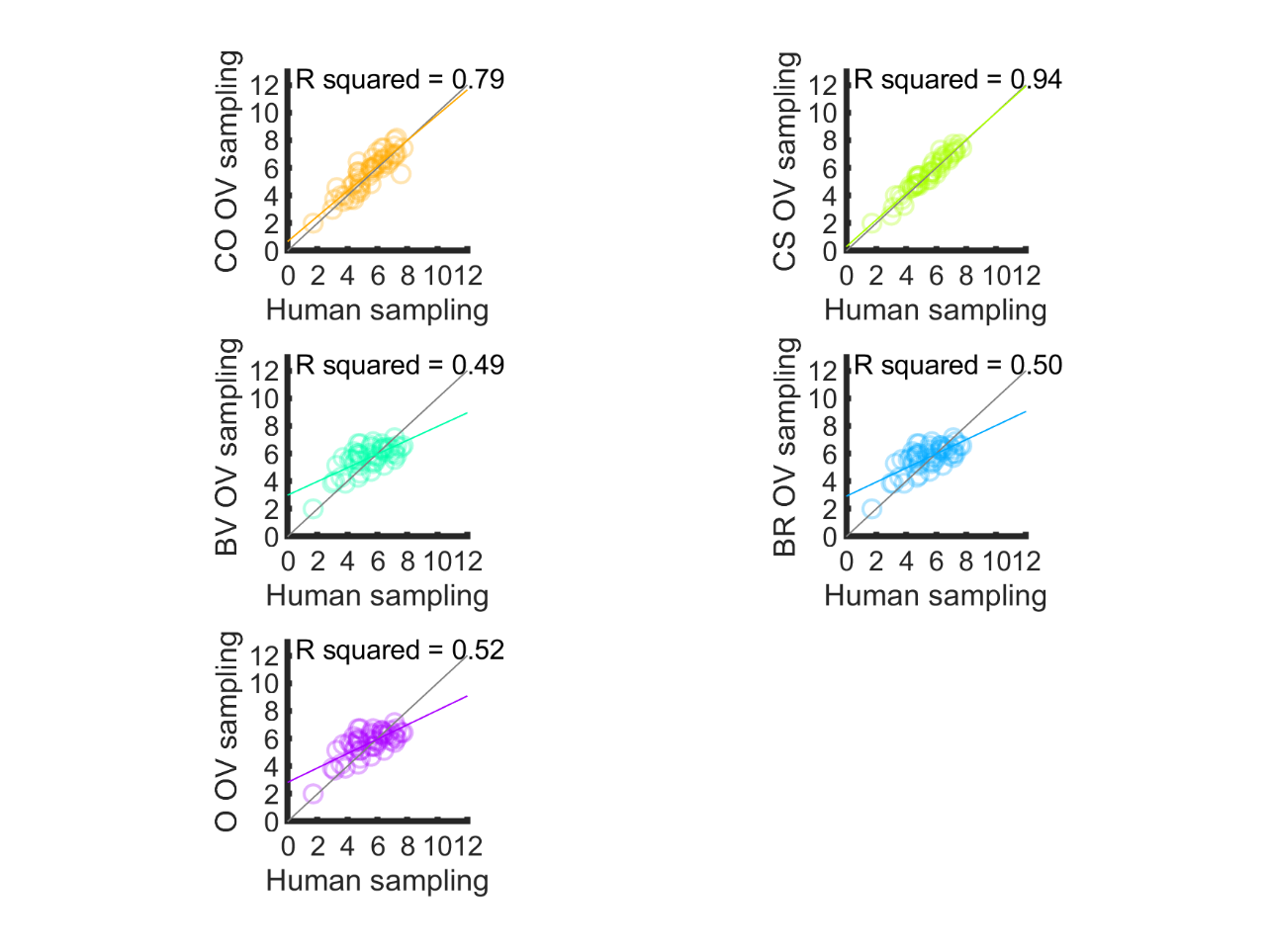


Figure S12. Linear relationships between human participants’ sampling in the Study 1 payoff condition versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured line represents the regression line, with corresponding *R2*printed 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.

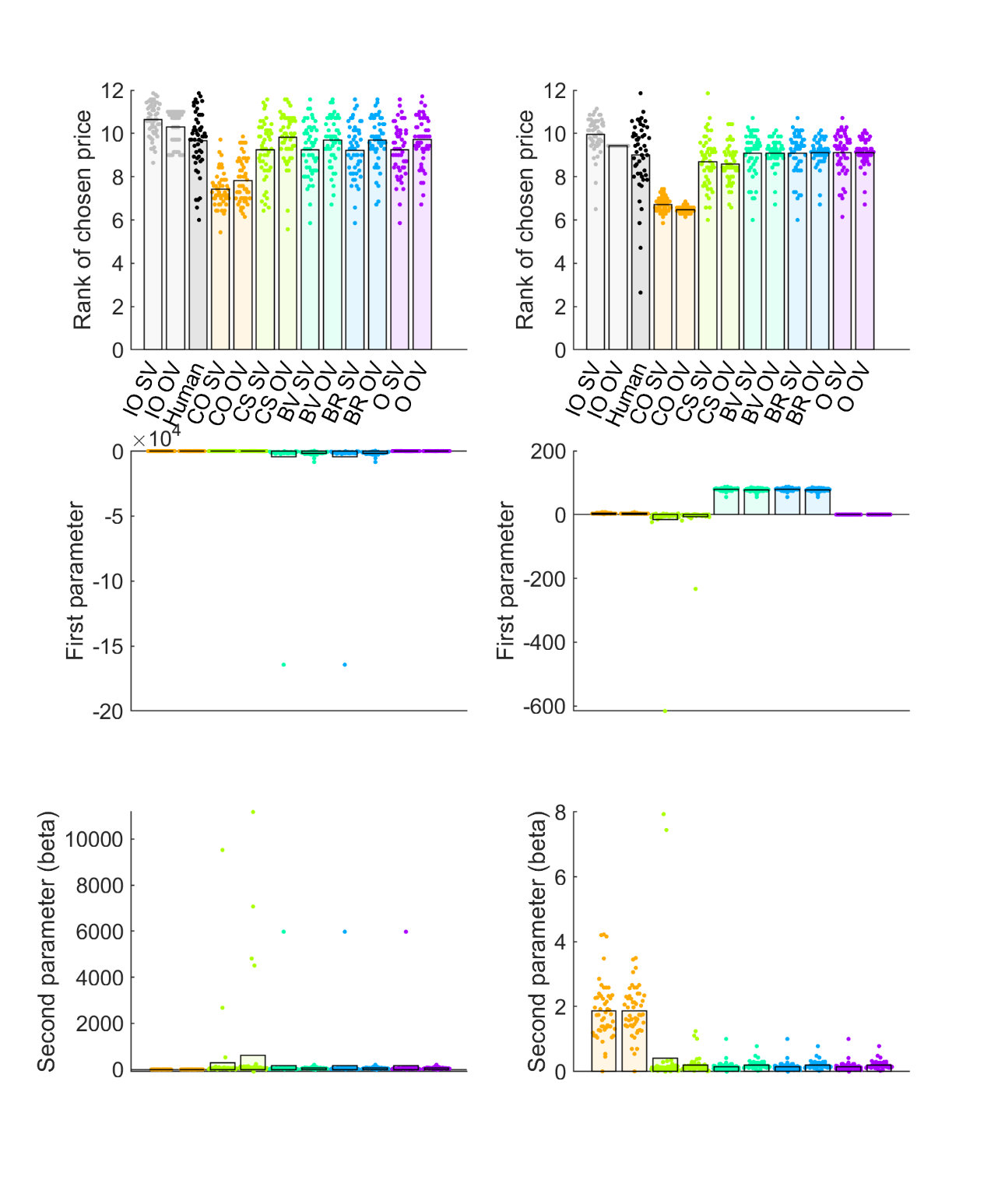


Figure S13. Model comparison for Study 1 ratings (left column) and full (right column) conditions. Top and middle rows show individual participants as points and bars show their mean values. The top row shows ranks of chosen prices. The second row plots the “first” or theoretical parameter values, estimated for each fitted model. The third row shows the “second”, or inverse temperature parameter beta, estimated for each fitted model. Abbreviations: IO = ideal observer, CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values, SV = subjective values.

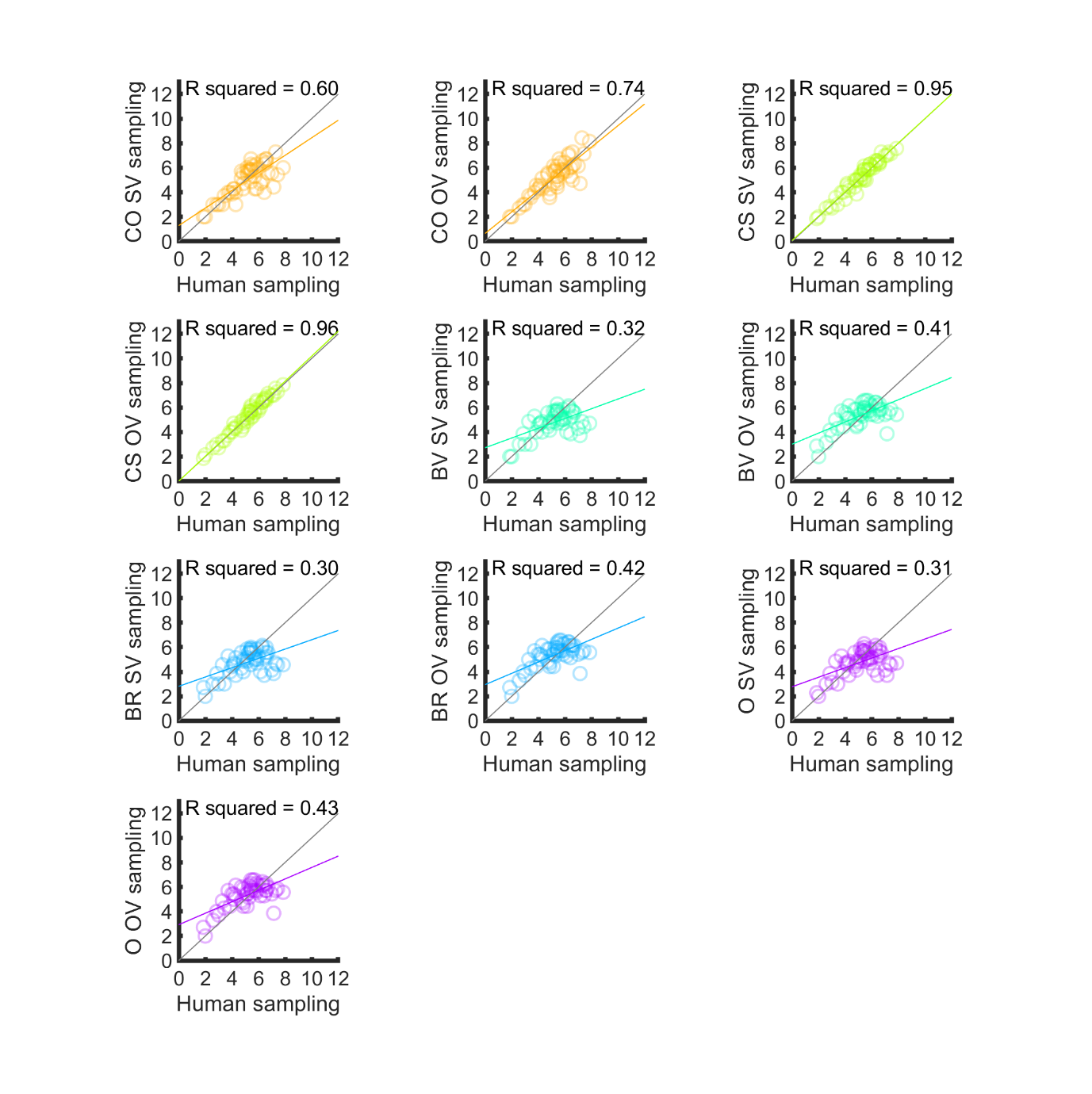


Figure S14. Linear relationships between human participants’ sampling in the Study 1 rating condition versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured line represents the regression line, with corresponding *R2* 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.

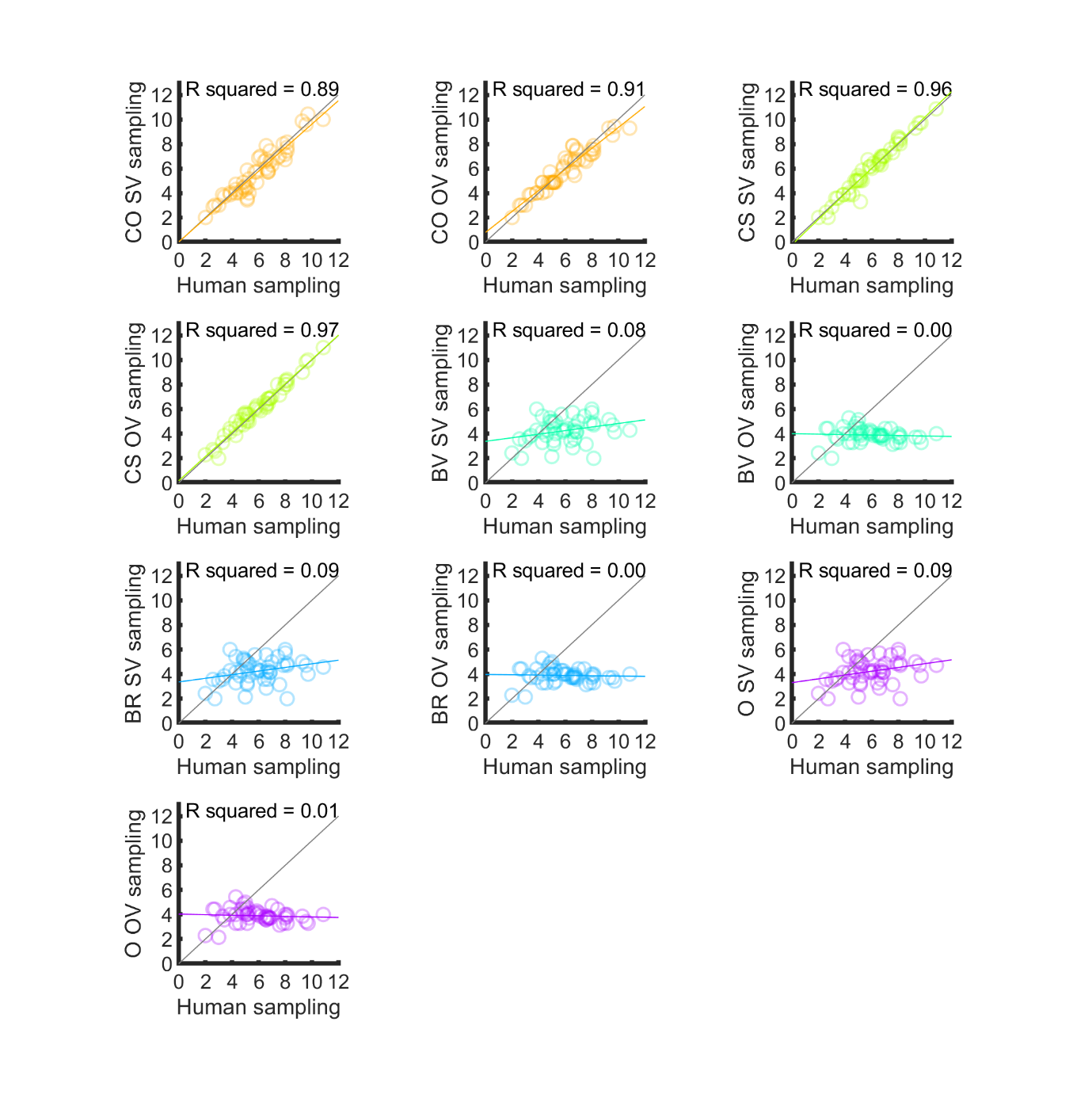


Figure S15. Linear relationships between human participants’ sampling in the Study 1 full condition versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured line represents the regression line, with corresponding *R2* printed on plot. The CS and CO models both predicts human data with high accuracy. Abbreviations: CO = cut-off, CS = Cost to Sample, BV = Biased Values, BR = Biased Reward, O = optimism, OV = objective values, SV = subjective values.

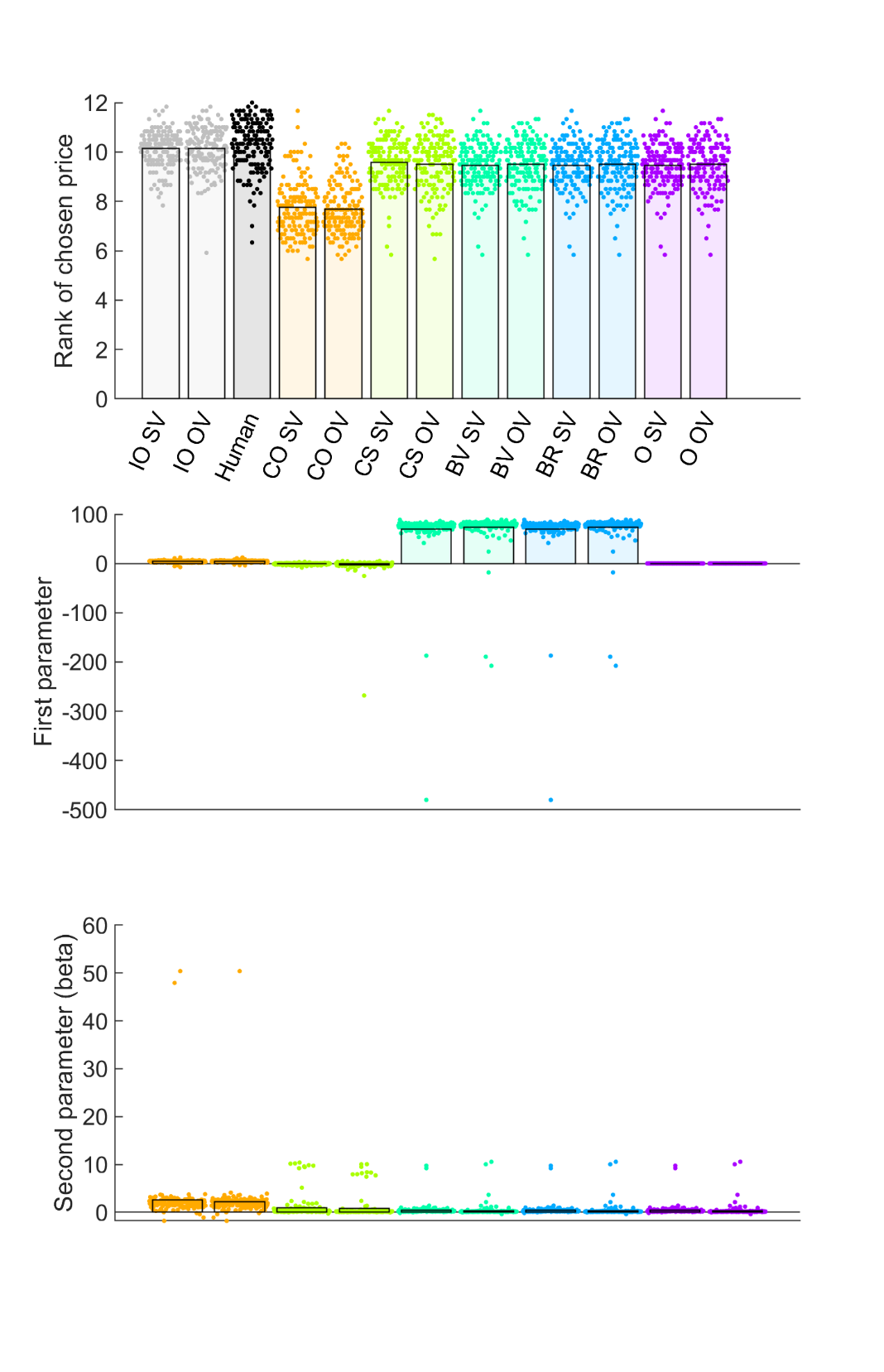


Figure S16. Model comparison for Study 2. Top and middle rows show individual participants as points and bars show their mean values. The top row shows ranks of chosen prices. The second row plots the “first” or theoretical parameter values, estimated for each fitted model. The third row shows the “second”, 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.

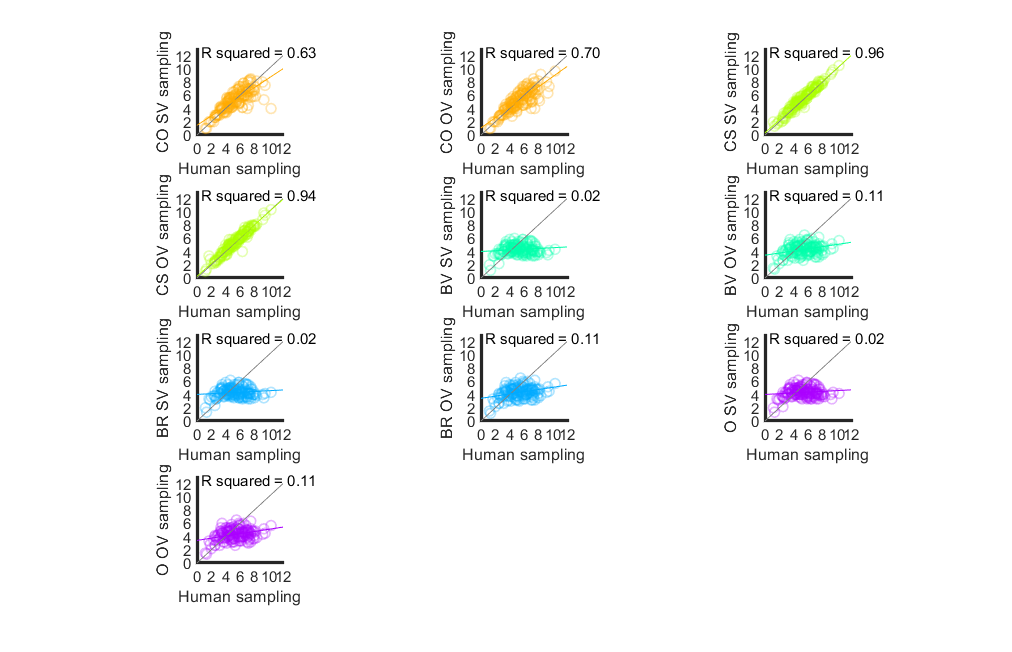


Figure S17. Linear relationships between human participants’ sampling in Study 2, versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured line represents the regression line, with corresponding *R2* printed on plot. 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.

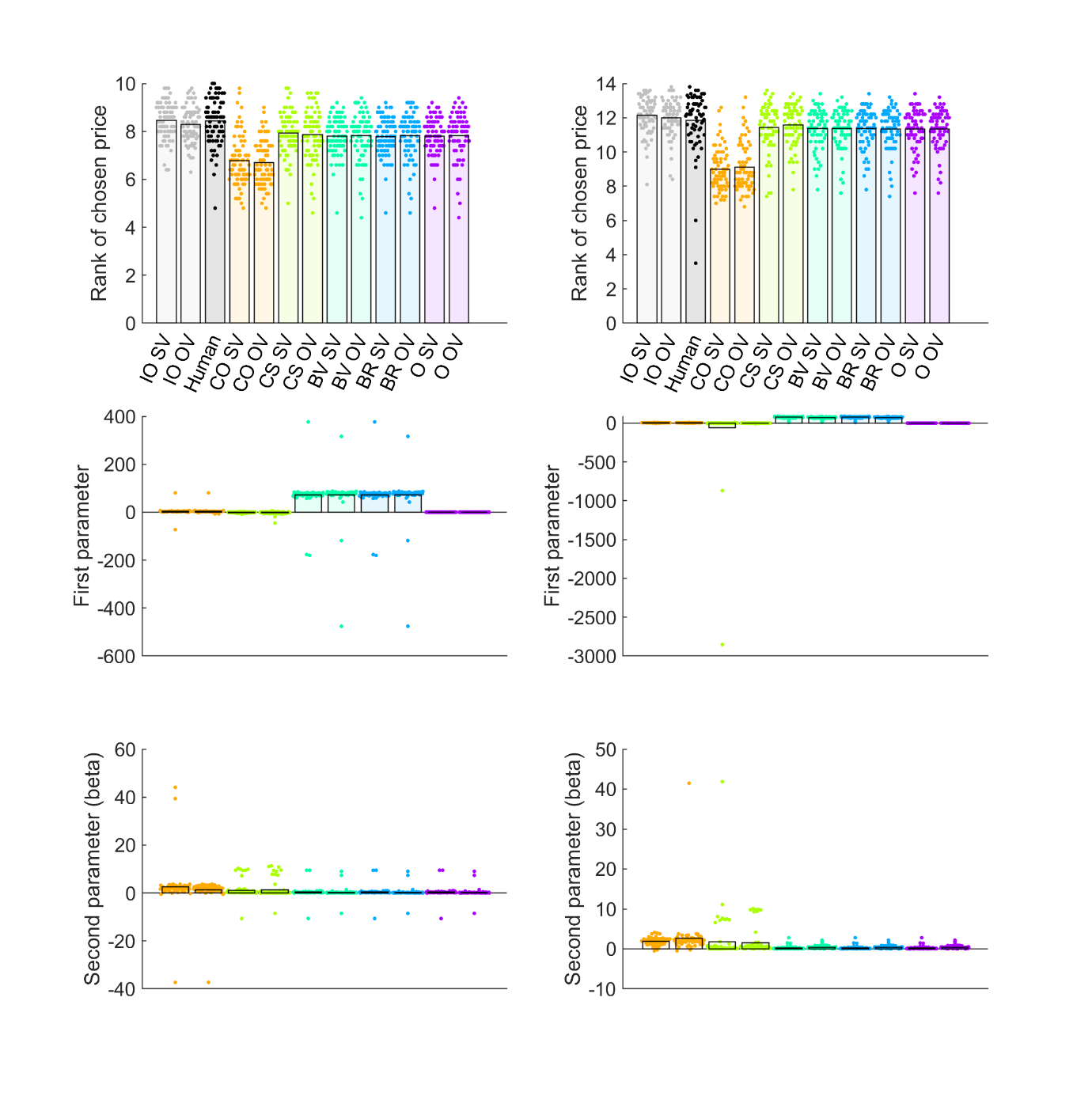


Figure S18. Model comparison for Study 3 10 options (left column) and 14 options (right column) conditions. Top and middle rows show individual participants as points and bars show their mean values. The top row shows ranks of chosen prices. The second row plots the “first” or theoretical parameter values, 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 S19. 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.

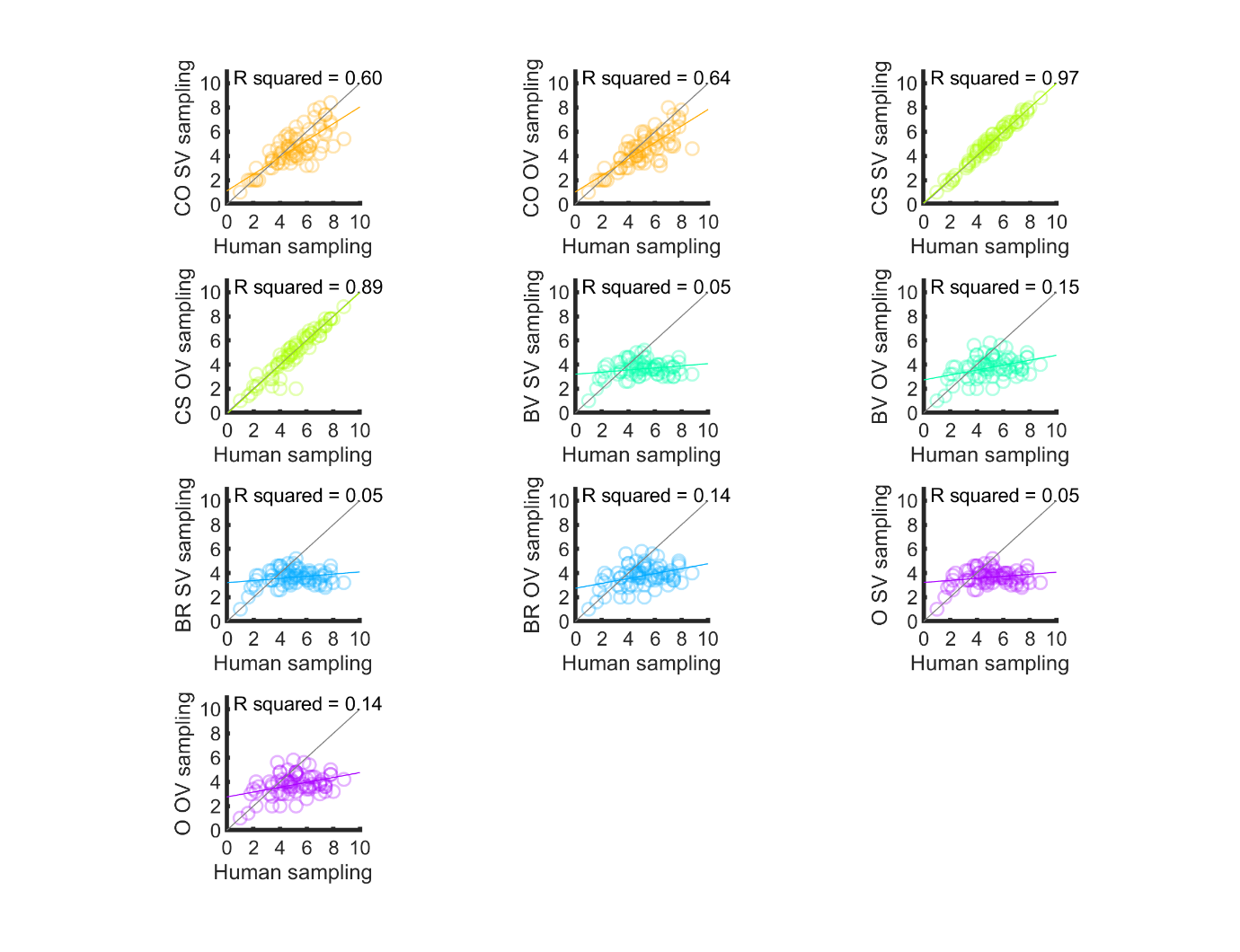


Figure S20. Linear relationships between human participants’ sampling in Study 3 14 options condition, versus sampling of corresponding models. The grey diagonal represents where participant and model numbers of samples are equal. The coloured line represents the regression line, with corresponding *R2* 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.

