Attentional Model for Framing-Time Pressure (FTP) Study

Ian D. Roberts

May 21, 2020

Contents

1 Basic Model

This document describes the model by breaking it down into its more basic components. In a nutshell, the model simulates a decider who starts with initial beliefs about what the choice option values might be and then attempts to learn the values of the options while attempting to minimize choice uncertainty. Furthermore, the decider will adapt its attentional search process when choice time is limited by re-weighting the factors that direct attention. Simulations suggest that the model is able to produce the sorts of shifts in attention and choice probabilities that we have observed behaviorally.

2 Model Components

2.1 Subjective valuation of alternatives

Risky alternatives are subjectively valued in line with prospect theory where the subjective value of an option is the product of the weighted probability and outcome:

$$v(x,p) = w(p)v(x) \tag{1}$$

Probability weighting:

$$w(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{\frac{1}{\gamma}}} \tag{2}$$

Value function:

$$v(x) = \begin{cases} x^\rho & x \geq 0 \\ -\lambda |x|^\rho & x < 0 \end{cases} \tag{3}$$

For these simulations the parameters were set to the following values:

- $\rho = 0.88$
- $\lambda = 1.5$
- $\gamma = 0.61$

2.2 Prior beliefs about alternatives

To generate the risky choice task trials, initial endowments were drawn from a uniform distribution ranging from 20 to 90 points and probabilities of winning the gamble were drawn from one of three truncated normal distributions with

$\mu = 0.28$, 0.42, or 0.56 (all with $\sigma = 0.2$ and bounded between 0.1 and 0.9). The expected value of the gamble and sure alternatives were equivalent on each trial.

Because the values of the risky alternatives presented on each trial are randomized, it is impossible for the decider is to form accurate predictions about what the values might be for any given trial. However, it is possible for the decider to generate more diffuse predictions about the range of possible values that options can take. Here we assume that the decider has a fairly accurate belief about the possible values that each option might take. For each option's prior (sure gain, sure loss, gamble) we create a normally distributed prior with μ equal to the median subjective value and σ^2 equal to the variance in subjective value.

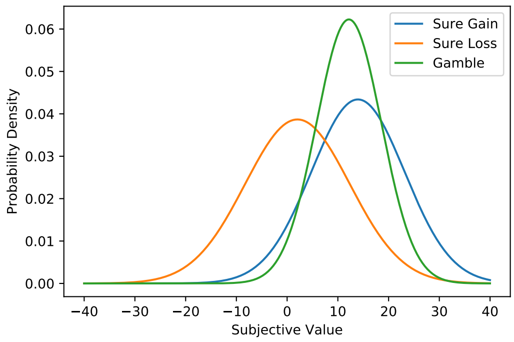


Figure 1: Choice option priors

2.3 Uncertainty reduction

The model posits that the choice process can be generally thought of as an attempt to reduce uncertainty, of which there are two primary types:

1. Uncertainty about the values of the available alternatives
2. Uncertainty about which alternative to choose

These uncertainties can be quantified in terms of entropy. For the decider's current uncertainty about a given alternative, we calculate the entropy in the normally distributed prior[1]

$$H(x) = \frac{1}{2}(1 + \log(2\sigma^2\pi)) \tag{4}$$

[1] https://en.wikipedia.org/wiki/Normal_distribution#Maximum_entropy

For the uncertainty about what to choose, we calculate the entropy across choice alternatives[2]

$$H(x) = -\sum_{i=1}^{n} P(x_i) \log P(x_i) \tag{5}$$

2.4 Belief updating (Kalman filter)

Although the decider begins each trial with a fairly uncertain belief about each option's potential value, she can learn about the values of the alternatives by visually attending to the options. We model this learning process as "sampling" from the alternatives by visually fixating them to gradually update the initial prior beliefs. Updating beliefs is carried out via a Kalman filter. Accordingly, with each sample in time t the expected value (i.e., μ of the belief prior) for option j was updated as a function of the current expected value and the prediction error of the sample:

$$\mu_j(t+1) = \mu_j(t) + K_j(t)(\hat{\mu}_j - \mu_j(t)) \tag{6}$$

where $\hat{\mu}_j$ is the "true" subjective value of the option on that trial and $K_j(t)$ is the Kalman gain which adjusts the learning rate. With a Kalman filter, the learning is adjusted dynamically such that the learning rate is greater when there is greater uncertainty about a value. Thus, the Kalman gain was updated with every sample and depended on the current level of uncertainty about an option's subjective value:

$$K_j(t) = \eta \frac{\sigma_j^2(t)}{\sigma_j^2(t) + \sigma_{\epsilon,j}^2} \tag{7}$$

where η is a "laziness" parameter which allows the Kalman filter to learn at a slower pace when < 1. We assume that different choice options might have higher levels of "noisiness" which can result in slower learning and a lower maximum for the level of certainty about the "true" value that can be achieved. To account for these differences in "noisiness", we include the $\sigma_{\epsilon,j}^2$ term (see Fig. 2 for an illustration of this parameter's effect). The variance of the current belief distribution $\sigma_j^2(t)$ is updated with each sample:

$$\sigma_j^2(t+1) = (1 - K_j(t))\sigma_j^2(t) \tag{8}$$

[2] `https://en.wikipedia.org/wiki/Entropy_(information_theory)#Definition`

For our simulations, we use the following parameter settings:

- $\sigma^2_{\epsilon,gamble} = 20.0$
- $\sigma^2_{\epsilon,sure} = 10.0$
- $\eta = 0.3$

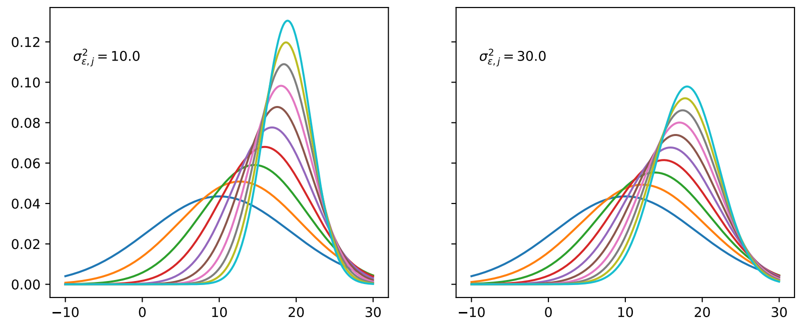


Figure 2: Kalman updating example

2.5 Choice probabilities

At any given point in time during a trial, the decider has some probability of choosing each alternative. These probabilities are determined based on the decider's current beliefs about the subjective values of the alternatives in accordance with signal detection for 2-alternative forced choice. To account for the fact that people tend to choose the alternative that they are currently fixating, there is a response bias towards the currently fixated option, therefore called the "sight bias". Using the differencing approach:

$$\mu_{gamble-sure} = (\mu_{gamble} - \mu_{sure}) + fixation(t)b \tag{9}$$

$$\sigma^2_{gamble-sure} = \sigma^2_{gamble} + \sigma^2_{sure} \tag{10}$$

where b is the sight bias parameter and $fixation(t)$ is an indicator variable with a value of 1 when the gamble is fixated, -1 when the sure option is fixated, and 0 when fixating neither. Choice probabilities are then calculated:

$$p(gamble) = 1 - \Phi(0, \mu_{gamble-sure}, \sigma^2_{gamble-sure}) \tag{11}$$

$$p(sure) = \Phi(0, \mu_{gamble-sure}, \sigma^2_{gamble-sure}) \tag{12}$$

where Φ is the cumulative distribution function of the normal distribution. Fig. 3 illustrates the calculation of choice probabilities. Based on the current beliefs about choice alternative subjective values (left panel), the probabilities of the either alternative having the higher subjective value is calculated by taking the difference between the two belief distributions (middle panel). The density below zero (shaded yellow) is the probability that the sure option has the higher value whereas the density above zero (shaded grey) is the probability that the gamble has the higher value. If the decider is currently fixating an option (say, the gamble), a sight bias is applied that increases the probability of choosing the currently fixated alternative (right panel).

For our simulations we use the following parameter settings:

- $b = 5.0$

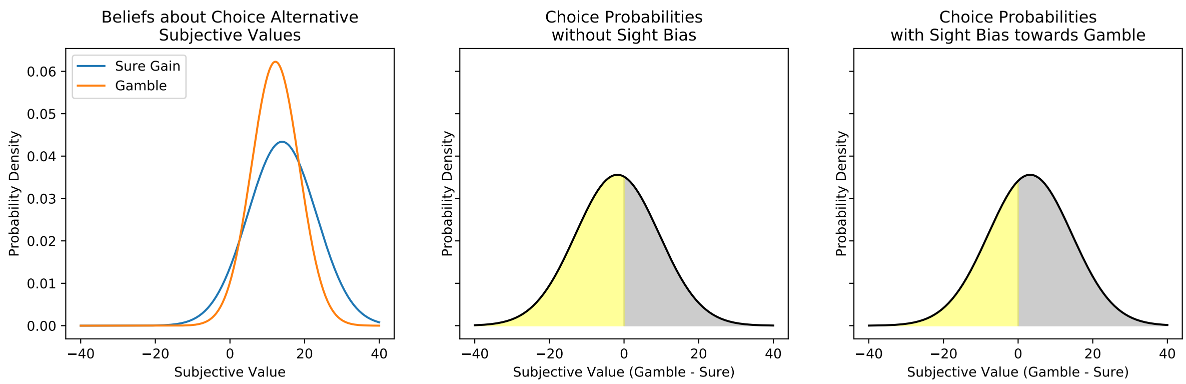


Figure 3: Choice probabilities example

2.6 Time constraints

Sometimes, the time to make a decision is limited. Because we assume the decider has an imperfect, noisy representation of how much time remains, the choice to continue gathering information about the choice alternatives carries the risk that time could run out and the decider would end up with nothing. The chances of this happening before the decider is able to make a choice increases as time progresses. To model the decreasing probability of being able to continue information search safely, we use the following function:

$$T(t, d, d_\epsilon) = \frac{1}{(1 + e^{d_\epsilon (t-d)})} \tag{13}$$

where t is the current time point, d shifts the overall estimate of the deadline, and d_ϵ controls the uncertainty about when time will run out (see Fig. 4).

This function is used to dynamically modulate weights given different attention priorities and information search termination (see below).

For our simulations we use the following values:

- $d = \begin{cases} 5.0 & \text{under time constraints} \\ 25.0 & \text{without time constraints} \end{cases}$

- $d_\epsilon = 1.0$

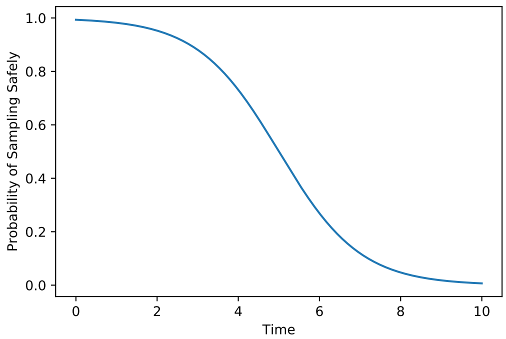


Figure 4: Time pressure

2.7 Attending to alternatives

Three main factors (and their interactions with each other and time pressure) are assumed guide visual attention:

1. Relative uncertainty about subjective values (i.e., the decider tends to attend to the alternatives about which they are most uncertain)

2. Efficiency of belief updating (i.e., the decider tends to attend to the options that can be encoded the most efficiently, primarily when under time pressure)

3. Relative subjective value of the choice alternatives (i.e., the decider tends to attend to the more highly valued option)

To decide to which alternative to attend, the attention value of each alternative is calculated as a weighted combination of the above terms. Because the decision of which alternative to attend to assumes that there will be another sample drawn before a choice will be made, the decider "looks ahead" to what the time pressure will be like in the future. In this way, it is assumed that time pressure influences attentional patterns before it affects the decision to terminate information search (see below) because the decider's attention precedes action and there needs to be enough time remaining to shift attention before deciding (e.g., to look at the most highly valued option before making the final decision). Furthermore, we assume that shifting fixation and drawing a sample costs more time that holding the current fixation and drawing another sample. Therefore, we also apply a "switching cost" when fixating an alternative would entail a fixation shift. To accommodate both attention pressure and switching cost, a future time point is used to calculate pressure:

$$t_j = t + AttendPressure + (1 - fixation_j(t))SwitchCost \tag{14}$$

where *AttendPressure* determines how early the decision pressure starts to influence the attentional process, *SwitchCost* determines how much added cost there is for an attentional shift, and $fixation_j(t)$ is an indicator variable which is equal to 1 if the decider is currently fixating option j and 0 otherwise.

To calculate the weight given to information, efficiency, and subjective value when deciding where to attend, the following terms are calculated. First, the information term:

$$a_{info} = h_{choice} T(t_j) w_{info} h_j \tag{15}$$

Thus, the greater the uncertainty about option j's value (h_j as calculated by 4) the more likely the decider is to attend to that option. A static weight w_{info} is applied to the uncertainty, as well as two dynamic weights: $T(t_j)$ and h_{choice}. First, the $T(t_j)$ weight means that the overall weight given to uncertainty about an option's value decreases as time pressure builds. Second, the h_{choice} weight means that as uncertainty about what to choose reaches 0 so does the weight given to uncertainty about options' values.

Next, there is a term for determining how much weight is given to the efficiency with which information can be gained about an option's value:

$$a_{noise} = -h_{choice}(1 - T(t_j)) w_{noise} \sigma^2_{\epsilon,j} \tag{16}$$

Thus, less weight is given to options as the noisiness of their subjective value representations increases, $\sigma^2_{\epsilon,j}$. A static weight w_{noise} is applied to the noisiness as well as two dynamic weights: $(1 - T(t_j))$ and h_{choice}. The first weight $(1 - T(t_j))$ means that efficiency only comes into consideration once pressure has begun to build. The second, h_{choice}, means that as choice entropy reaches 0 so does the weight given to efficiency.

Last, there is a term for how much weight to give to the subjective value of the choice options:

$$a_{value} = e^{-T(t_j)w_{info}h_j}\mu_j(t) \tag{17}$$

Thus, the decider is more likely to attend to options she currently believes to be greater value, $\mu_j(t)$. This is weighted by $e^{-T(t_j)w_{info}h_j}$ which increases to 1 as each of the three things occur: time pressure $T(t_j)$ builds, uncertainty h_j about the values of each of the choice options approaches 0, or the weight given to information search w_{info} reduces to 0.

Therefore, the total value of attending to option j is given by:

$$a_j = a_{info} + a_{noise} + a_{value} \tag{18}$$

Though Eq. 18 represents the complete model, simulations using various parameter settings suggest that giving weight to the uncertainty about option values does not produce data that resembles the behavioral data. Therefore, we have set $w_{info} = 0$, which simplifies the attention value equation to:

$$a_j = -h_{choice}(1 - T(t_j))w_{noise}\sigma^2_{\epsilon,j} + \mu_j(t) \tag{19}$$

To calculate the probability of attending to the gamble, the attention values are submitted to a logit function:

$$PrFix(gamble) = (1 + e^{-\tau(t)(a_{gamble} - a_{sure})})^{-1} \tag{20}$$

where $\tau(t)$ is a sensitivity parameter that increases as time pressure builds (i.e., attention becomes more exploitative and less exploratory as time pressure increases):

$$\tau(t) = sens_{attend}(1 + sens_{pressure}(1 - T(t + AttendPressure))) \tag{21}$$

where $sens_{attend}$ is the sensitivity to difference in attention value where there is no time pressure and $sens_{pressure}$ is how much sensitivity will increase as a function of time pressure.

For our simulations, we set the parameter values to:

- $w_{info} = 0.0$
- $w_{noise} = 0.5$
- $AttendPressure = 5.0$
- $SwitchCost = 1.0$

- $sens_{attend} = 0.1$
- $sens_{pressure} = 0.5$

2.8 Terminating information search

Last, the decider must arbitrate between either drawing another sample or terminating information search and making a choice. The value of drawing another sample is a weighted function of the uncertainty surrounding the subjective values of the options and the choice uncertainty:

$$SampValue = T(t)h_{choice} + (1 - e^{-h_{choice}})w_{info} \sum_{j=1}^{2} T(t_j)h_j \tag{22}$$

According to this equation, as choice entropy reaches 0, so does the term for option value uncertainty. However, because we've set $w_{info} = 0$, this equation reduces to $SampValue = T(t)h_{choice}$. Thus, if no priority is given to reducing uncertainty about the subjective values of the options, whether to continue sampling or not is entirely driven by choice uncertainty weighted by time pressure. The value of sampling is then submitted to a modified logit equation:

$$PrSamp = \frac{2}{1 + e^{-sens_{sample} \cdot SampValue}} - 1 \tag{23}$$

where $sens_{sample}$ controls the sensitivity to $SampValue$.

For our simulations, we set the following values:

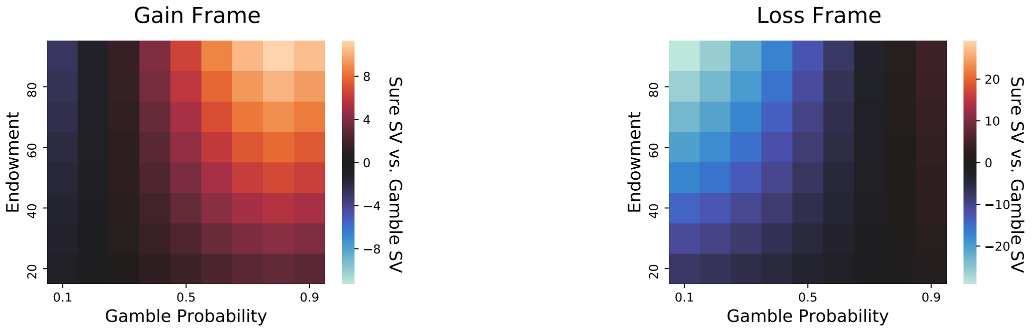
- $sens_{sample} = 3.0$

3 Simulations

This section presents results from simulations using the current parameter settings.

3.1 Subjective Values

With the current prospect theory parameters, Fig. 5 shows how the subjective values of the sure options and their corresponding gambles compare.



(a) Gain Frame

(b) Loss Frame

Figure 5: Sure vs. Gamble Subjective Values

3.2 Time Pressure on First Fixations

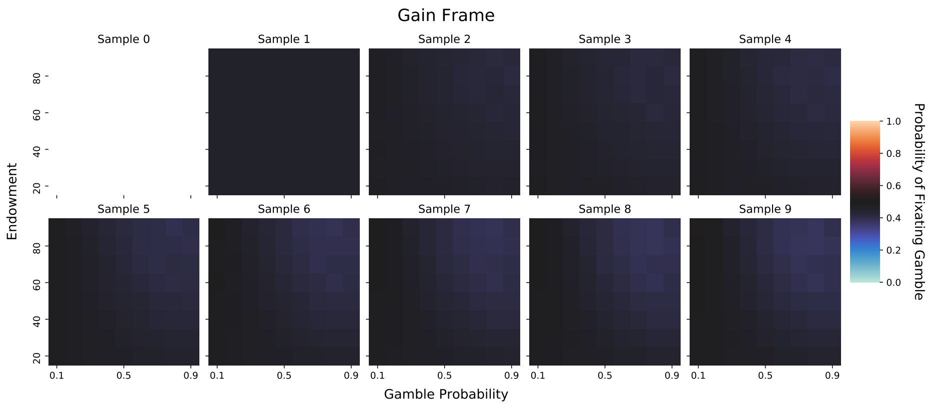
Reducing the amount of time to decide, increases the tendency to fixate the sure option first for gain frame trials (Fig. 6) but has less of an effect on loss frame trials (Fig. 7).

3.3 Time Pressure on Choice Probabilities

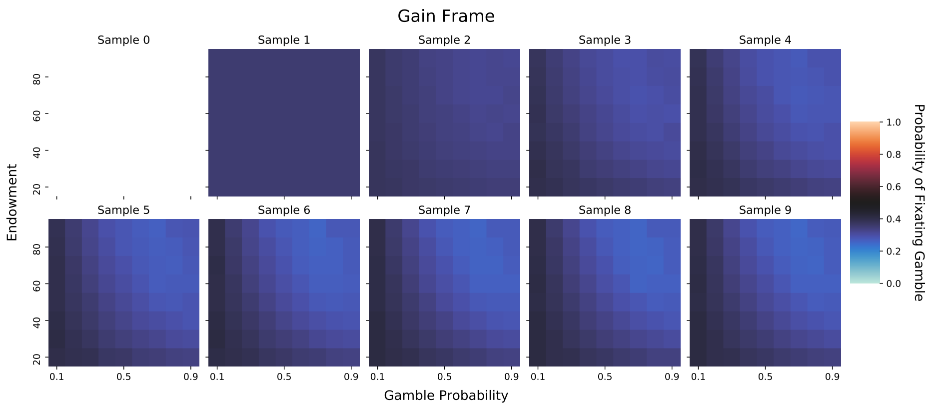
Reducing the amount of time to decide, increases the tendency to choose the sure option gain frame trials (Fig. 8) but has less of an effect on loss frame trials (Fig. 9).

3.4 Fixations on Choice Entropy

Manipulating the fixation pattern so that participants only fixate one option, fixating the sure option reduces choice entropy more than fixating the gamble for gain frame trials (Fig. 10) but the opposite is true for loss frame trials (Fig. 11).

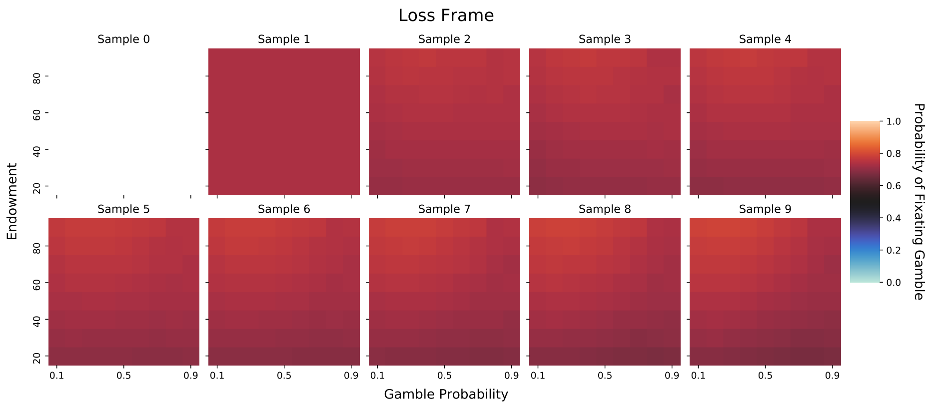


(a) No Time Pressure

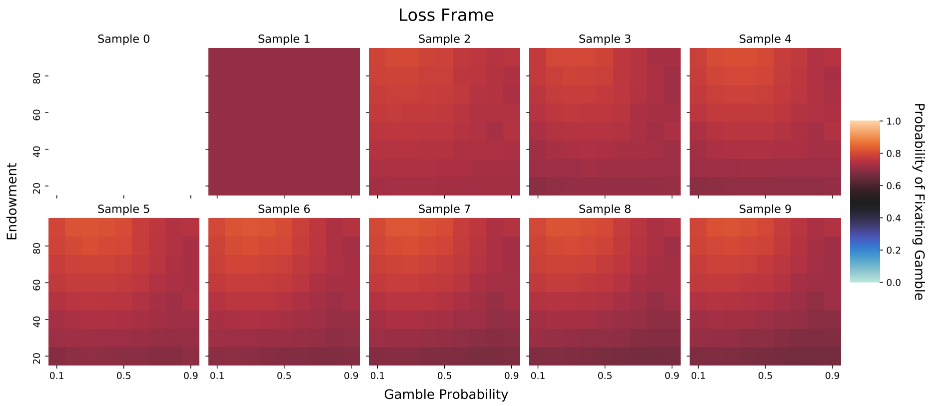


(b) Time Pressure

Figure 6: Gain Frame: Probability of Fixating Gamble

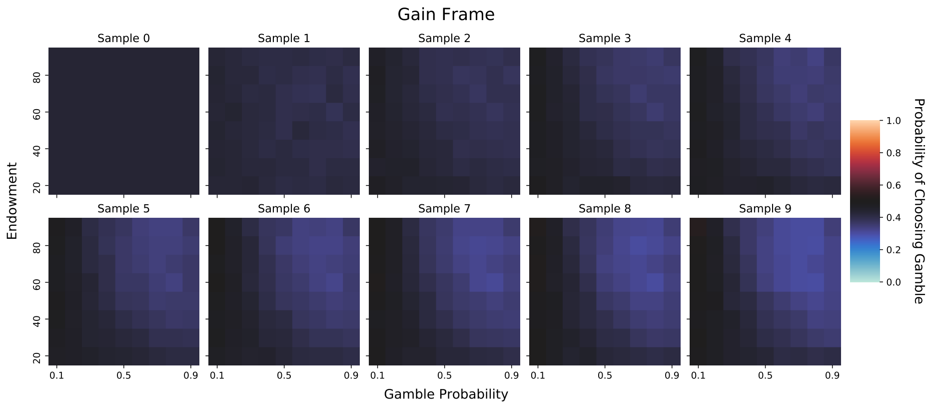


(a) No Time Pressure

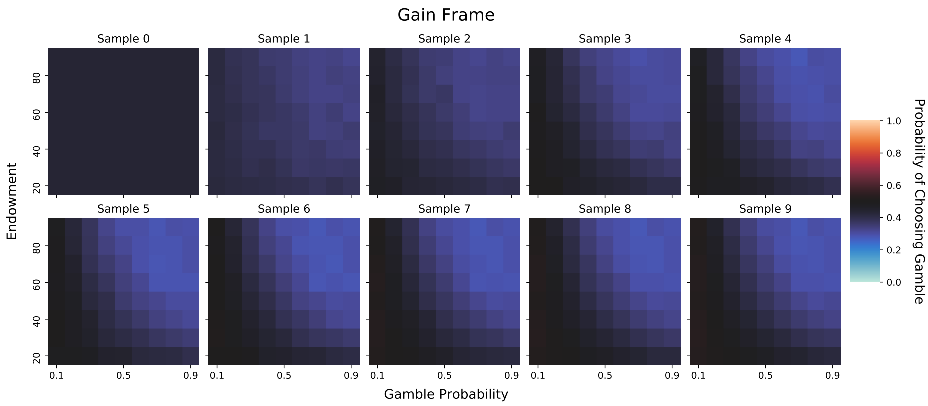


(b) Time Pressure

Figure 7: Loss Frame: Probability of Fixating Gamble

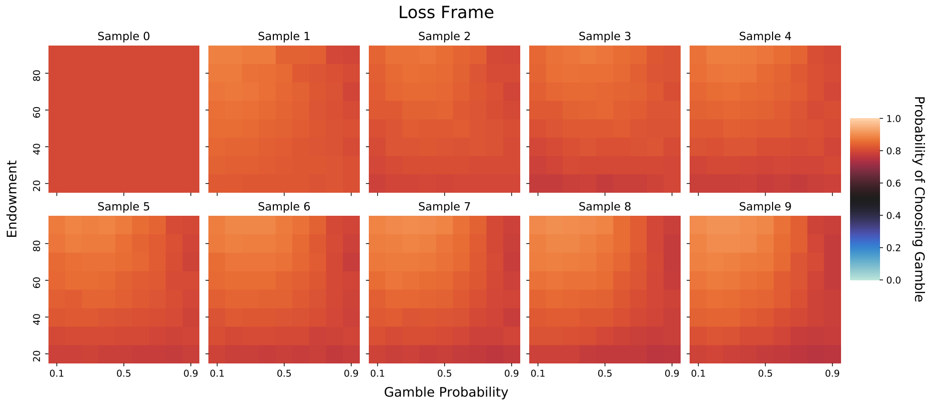


(a) No Time Pressure

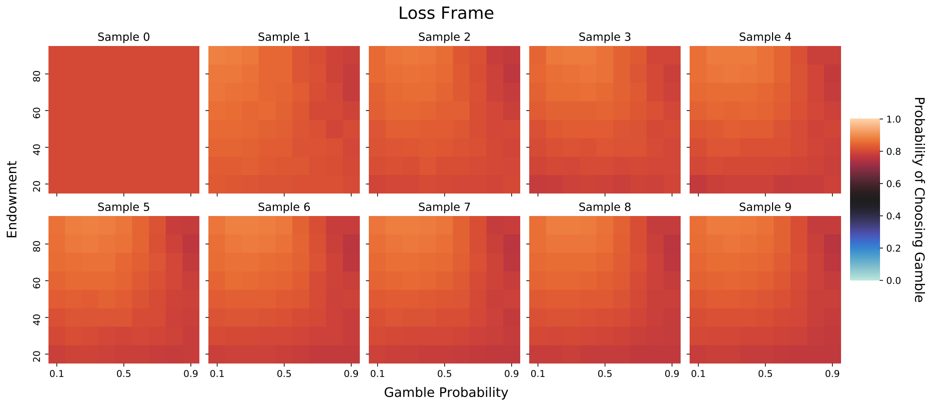


(b) Time Pressure

Figure 8: Gain Frame: Probability of Choosing Gamble

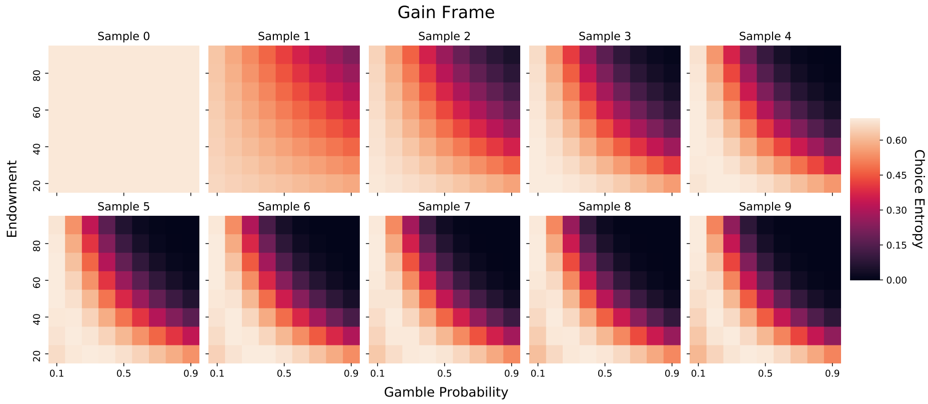


(a) No Time Pressure

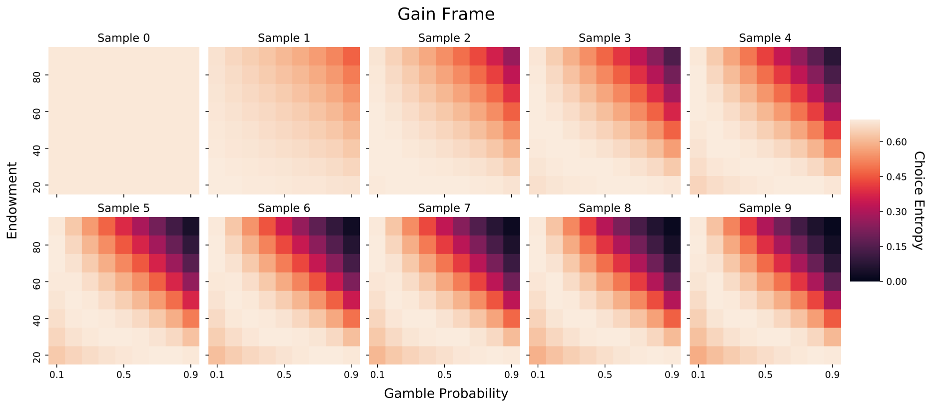


(b) Time Pressure

Figure 9: Loss Frame: Probability of Choosing Gamble

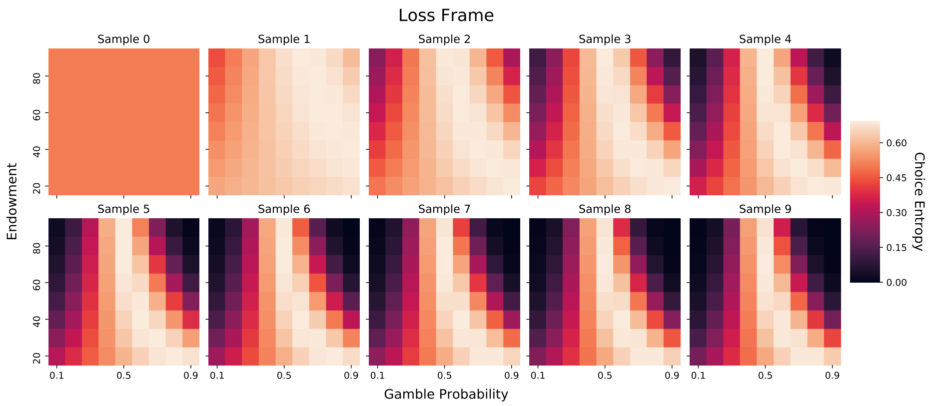


(a) Only Fixate Sure Option

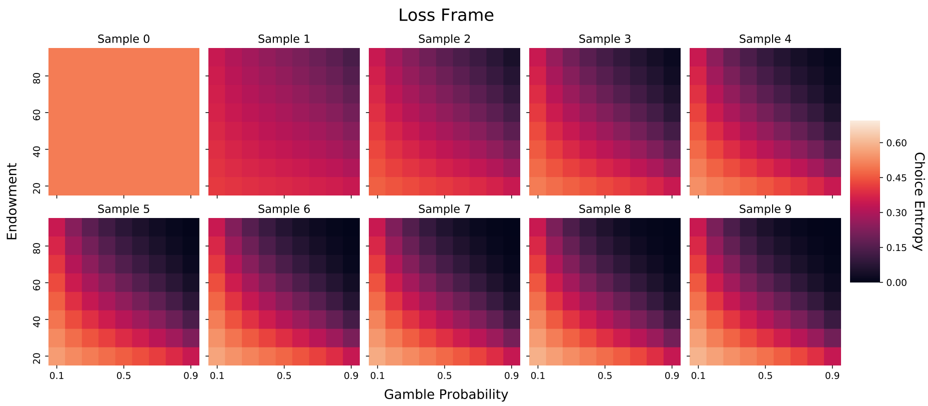


(b) Only Fixate Gamble

Figure 10: Gain Frame: Choice Entropy



(a) Only Fixate Sure Option



(b) Only Fixate Gamble

Figure 11: Gain Frame: Choice Entropy