Decisions From Experience

Motivation

- Oftentimes individual agents have limited understanding of the environment.
- The agent may not understand the exact incentive structure and compute the implied equilibria.
- Rather, the agent might need to rely on experience with similar situations.
- Consider the distinction between "decisions from description" and "decisions from experience."

Decisions From Experience

- Studies on decisions from experience do not receive a prior description of the incentive program.
- Such studies focus on situations for which rational choice does not have clear predictions.
- As a result, almost any behavior can be justified as rational given certain prior beliefs.
- Decisions from experience are intended to expand on the set of situations that can be addressed with economic models that provide clear and useful predictions.

Clicking Paradigm

The current experiment includes many trials. Your task, in each trial, is to click on one of the two keys presented on the screen. Each click will be followed by the presentation of the keys' payoffs. Your payoff for the trial is the payoff of the selected key.

The Clicking Paradigm focuses on the effect of experiencing monetary payoffs without knowledge of the incentive structure. Subjects base their decisions on the feedback (observed monetary outcomes) of previous decisions.

Review

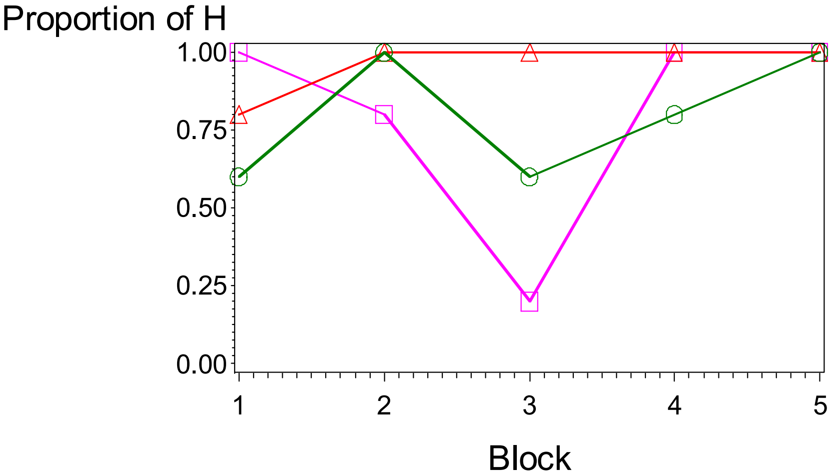
- The 2-alternative Clicking Paradigm with complete feedback and a static payoff rule is used.
- After each choice in the experiments, the agents receive feedback concerning their obtained payoff (the payoff from the key they selected) and the foregone payoff (the payoff that could be obtained had they selected the alternative key).
- The payoff of each key is drawn from a payoff distribution associated with that key.
- The fact that the payoff rule is static implies that the distributions do not change during the experiment.
- Based on this setup, we observe 6 robust behavioral regularities.

I. The Law of Effect

- Thorndike (1898) studied how cats learn to escape from puzzle boxes.
- The experiment included several trials: Each trial started with a placement of a cat in a puzzle box and ended when the cat exited the box.
- Evaluation of the learning curves (time to escape as a function of trial number) led Thorndike to conclude that the learning was gradual and stochastic.
- *Choices that have led to good outcomes in the past are more likely to be repeated in the future.*

I. The Law of Effect

- Studies that use the clicking paradigm reveal a similar pattern.
- The experiment involved a trivial choice task: option 'H' (high payoff) that always provides a payoff of 1 Shekel and option 'L' (low option) that always provides a payoff of 0.
- Participants did not receive prior information concerning the payoff rule and could rely only on feedback concerning the obtained and foregone payoffs.
- The results are presented in five blocks of five trials each.



The learning process is noisy. The proportion of optimal choices of the circle subject go up to 100% by the second block, then go down to 60% in the third block, and then go up to 100% in the fifth block.

II. The Payoff Variability Effect

- Myers and Sadler (1960) studied decisions from experience using a card flipping paradigm.
- In each trial, the participant saw one side of a card and had to decide whether to accept the payoff written on the side (the safe alternative) or the payoff written on the unobserved side of the card (the riskier option).
- Participants received feedback concerning their payoffs after each choice (the card was flipped only if the participant chose the riskier option).
- The results revealed that an increase in the payoff variability of the risky option reduced the proportion of choices that maximized the expected payoff.
- Busemeyer and Townsend (1993) termed this pattern the "payoff variability effect."

II. The Payoff Variability Effect

- The pattern is replicated using the clicking paradigm.
- Problems 1, 2 and 3 were run in the same experiment in a within-participant design.
- Each of 20 participants faced each problem for 200 rounds with complete feedback.
- The order of the three problems was random.
- The participants did not receive a description of the problems, but were informed that the experiment includes three independent parts and when a new part starts.

Problem 1: (r=200, n=20, FB=complete, payoff in shekels in a randomly selected trial)

H	1 with certainty	[H-rate: 96%]
L	0 with certainty	

Problem 2 (same procedure as in Problem 1)

H	+11 with probability 0.5 -9 otherwise (EV = 1)	[H-rate: 58%]
L	0 points with certainty	

Problem 3 (same procedure as in Problem 1)

H	0 with certainty	[H-rate: 53%]
L	9 with probability 0.5 -11 otherwise (EV = -1)	

The Payoff Variability Effect

- The higher EV maximization rate (H-rate) in Problem 1 (96%) compared to Problem 2 (58%) can be described as an indication of risk or loss aversion. H was less attractive in Problem 2 when the variance increased and was associated with losses.

- However, this "risk and/or loss aversion" explanation is inconsistent with a comparison of Problem 2 and Problem 3.

- In Problem 3, risk and loss aversion implies maximization of 'H' choices. The results show an H-rate of only 53%.

- Additional studies demonstrate the robustness of the payoff variability effect.

III. Underweighting of Rare Events

- Kahneman and Tversky (1979) demonstrate that two of the best known violations of mainstream economic theory, the tendency to buy both insurance and lotteries can be explained as indications of overweighting rare events.
- That standardized paradigm focuses on decisions from description.
- Problems 7 and 8 demonstrate the evidence for underweighting of rare events in decisions from experience.
- These problems were studied by Nevo and Erev (2012) using the clicking paradigm with complete feedback.

III. Underweighting of Rare Events

Problem 7 (r=100, n=24, FB=complete, payoff in shekels in a randomly selected trial)

S	0 with certainty	[S-rate = 43%]
R	+1 with probability 0.9; -10 otherwise (EV = -0.1)	

Problem 8 (same procedure as in Problem 7)

S	0 with certainty	[S-rate = 72%]
R	+10 with probability 0.1; -1 otherwise (EV = +0.1)	

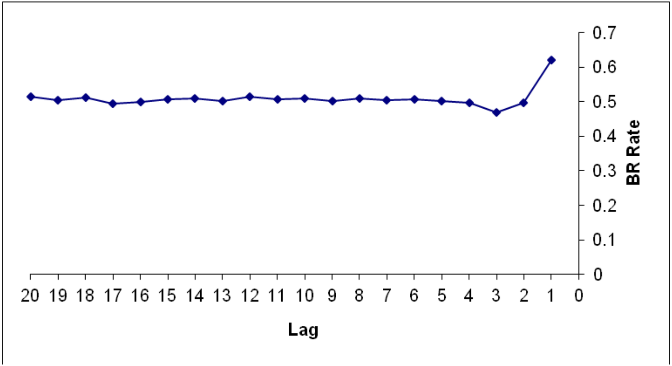
III. Underweighting of Rare Events

- In Problem 7, the safer option has a higher expected value, but the participants tend to select the gamble.
- Problem 8 reflects the opposite risk preference: the gamble has a higher expected value, but the participants tend to select the safer option.
- Baron and Erev note that this pattern can be a reflection of insufficient sensitivity to the rare and extreme outcomes (the extreme outcomes that occur in 10% of the trials).
- Thus, the participants behave as if they believe "it won't happen to me."

IV. The Very Recent Effect

- Analysis of the effect of recent outcomes on choice behavior in probability learning tasks led Estes to conclude that the most common pattern is positive recency: decision makers are more likely to select the alternative that led to the best outcome in recent trials.
- Let Best-Reply-L be the choice rate of the alternative that led to the best outcomes exactly L trials before the current trial.
- Figure 4 presents the values of Best-Reply-1 to Best-Reply-20 (based on the data from trial 21 until 200 in Problems 2 and 3).
- The results reveal a large qualitative difference between Best-Reply-1 and the other values.
- Nevo and Erev (2012) call it the "very recent effect."

Figure 4: The very recent effect: The proportion of choices (at trial t) of the alternative that led to the best outcome in trial t-Lag. Thus, Lag=1 (on the right) present the best reply rate to the most recent trial, and Lag=2 present the best reply rate to the outcome occur in the trial before the most recent. The analysis is based on trial 21 to 200 in Problems 2 and 3.



What is the proportion of choices, had I chosen in trial t the alternative that led to the best outcome in trial t-Lag?

V. Inertia and Surprise-Triggers-Change

- Analysis of the relationship between recent and current choice reveals strong positive correlation i.e. inertia.
- Decision makers tend to repeat their last choice.
- In Problems 2 and 3, for example, participants repeated their last choice in 68% of the trials.
- Moreover, inertia is a better predictor of behavior than positive recency.
- When inertia and positive recency lead to contradicting predictions, the participants are more likely to exhibit inertia.

VI. Individual Differences

- There exist individual differences that are captured in the speed of learning.
- Some participants are fast learners, other participants are slow learners and others never learn.
- Consider the Iowa Gambling Task.

Iowa Gambling Task

- Bechara, Damasio, Damasio and Anderson (1994) studied patients with lesions in the orbitofrontal cortex.
- This syndrome involves intact IQ and reasoning skills but poor decision-making capacities.
- The task is presented as a choice between four decks of cards.
- Each alternative results in a sure gain or with some probability to a loss.
- The information is limited to the obtained payoff after each trial.

Iowa Gambling Task

Dis R: Win \$100 with probability 0.9; lose \$1150 otherwise (EV = -25)
Dis S: Win \$100 with probability 0.5; lose \$150 otherwise (EV = -25)
Adv R: Win \$50 with probability 0.9; lose \$200 otherwise (EV= +25)
Adv S: Win \$50 with probability 0.5; 0 otherwise (EV = +25)

Bechara, Damasio, Damasio and Anderson (1994) found that the patients with lesions in the orbitofrontal cortex did not learn to avoid the disadvantageous alternatives, while the participants in the control group (without lesions in the orbitofrontal cortex) did.

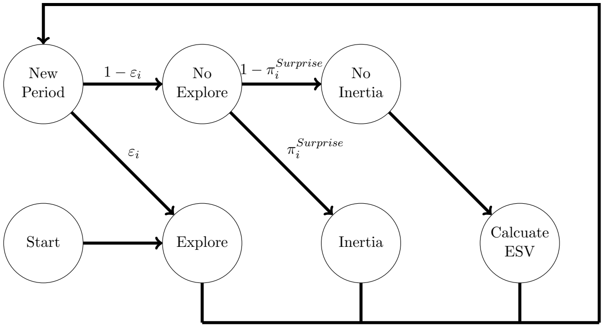
Inertia, Sampling and Weighting (I-SAW)

- I-SAW is an instance-based model, which allows for three response modes: *exploration, inertia and exploitation.*
- In each period, a player enters one of the modes with different probabilities.
- There are n players in the game.
- Each player has a set of parameters $(p_A, \varepsilon_i, \pi_i, \mu_i, \rho_i, \omega_i)$.
- The parameter $p_A \in [0,1]$ is the same for all agents.
- The other parameters are idiosyncratic with $\varepsilon_i \sim U[0,\varepsilon]$, $\pi_i \sim U[0,\pi]$, $\mu_i \sim U[0,\mu]$, $\rho_i \sim U[0,\rho]$ and $\omega_i \sim U[0,\omega]$.

Inertia, Sampling and Weighting (I-SAW)

- For simplicity, assume two actions.
- Player i's action set is $\mathcal{A}_i = \{A, B\}$.
- Let a_i^t be the action of player i that was played in period t, where $h_i(t_1, t_2) = \left\{a_i^{t_1}, a_i^{t_1+1}, \ldots, a_i^{t_2}\right\}$ for $t_1 \leq t_2$.
- Similarly, let a_{-i}^t be the actions of players other than i in period t, where $h_{-i}(t_1, t_2) = \left\{a_{-i}^{t_1}, a_{-i}^{t_1+1}, \ldots, a_{-i}^{t_2}\right\}$ for $t_1 \leq t_2$.
- We explain next the three response modes.

Schematic of I-SAW



Notes: I-SAW allows for three response modes: exploration, inertia and exploitation. In exploration trials, a player chooses amongst actions with some probability. Exploration occurs with probability ε_i. Inertia occurs with probability $(1-\varepsilon_i) \times \pi_i^{Surprise(t)}$. In this mode, a player repeats the last action. Exploitation occurs with probability $(1-\varepsilon_i) \times (1-\pi_i^{Surprise(t)})$. In exploitation trials, a player selects the action with the highest Estimated Subjective Value (ESV).

Exploration

- In exploration, each player chooses action A with probability p_A and action B with probability $1 - p_A$. The probabilities are the same for all players.

Inertia

- The decision to enter the inertia mode depends on an endogenous parameter $Surprise(t) \in [0, 1]$.
- A player might enter the inertia mode after period 2 with probability $\pi_i^{Surprise(t)}$, where $\pi_i \in [0, 1]$.
- The probability of inertia is low when surprise is high and vice versa.

Exploitation

- In exploitation trials, an individual selects the action with the highest Estimated Subjective Value (ESV).
- To determine the ESV, player i randomly selects μ_i elements from $h_{-i}(0, t-1)$ with replacement; let us call this set $M_{-i}(0, t-1)$.
- This set is chosen according to the following: with probability ρ_i player chooses a_{-i}^{t-1} and with probability $1 - \rho_{-i}$ player chooses uniformly over $h_{-i}(0, t-1)$.
- The same set M_{-i} is used for each $a_i \in \mathcal{A}_i$.
- The sample mean for action a'_i is then defined as

$$SampleM(a'_i, t) = \frac{1}{|M_{-i}(0, t-1)|} \sum_{a_{-i} \in M_{-i}(0, t-1)} g_i(a'_i, a_{-i}).$$

Exploitation

- The $GrandM(a_i', t)$ is defined as

$$GrandM(a_i', t) = \frac{1}{|h_{-i}(0, t-1)|} \sum_{a_{-i} \in h_{-i}(0,t-1)} g_i(a_i', a_{-i}).$$

- Then, player i's ESV of action a_i' is

$$ESV(a_i') = (1 - \omega_i) \cdot SampleM(a_i', t) + \omega_i \cdot GrandM(a_i', t),$$

where ω is the weight assigned on the payoff based on the entire history ($GrandM$) and $1 - \omega$ is the weight assigned on the payoff based on the sample from the history ($SampleM$).

- Then, the player simply chooses the a_i' that maximizes ESV (and chooses randomly in ties).

Inertia, Sampling and Weighting (I-SAW)

1. The Law of Effect – The observed deviations from 'best reply to past experiences' can be an indication of an exploration effect of selecting the 'L' key. I-SAW captures this observation with the assertion that in certain trials the subjects choose an *exploration mode*.
2. The Payoff Variability Effect – I-SAW captures this effect (and correlation effects) with the assertion that the subjects tend to rely on a small sample of past experiences.
3. Underweighting of Rare Events – I-SAW captures the tendency to underweight rare events with the assertion that people rely on small samples of past experiences.

Inertia, Sampling and Weighting (I-SAW)

4. The Very Recent Effect – The effect is captured with a minimal addition onto I-SAW. The addition assumes that the most recent outcome is particularly likely to affect the next choice, while all older experiences are sampled with equal probability.

5. Inertia and Surprise-Triggers-Change – The observed inertia and the complex recency pattern is captured in I-SAW with the hypothesis that in certain trial periods, participants choose an inertia mode and simply repeat their last choice. The probability of terminating the inertia mode increases with surprise.

Inertia, Sampling and Weighting (I-SAW)

6. Individual Differences – I-SAW assumes that the observed individual differences reflect quantitative rather than qualitative differences. That is, the different tendencies can be captured as indications of individual-specific learning parameters. I-SAW distinguishes between two classes of parameters. One class involves the parameters that describe individual agents ("traits"). A second class involves the parameters that capture the distribution of traits in the population.