A Primer in Behavioral Economics

2025

Main themes

- 1. The neoclassical approach vs. the behavioral approach
- The empirical facts from the lab and the field
- Competing theories: different predictions, sometimes same predictions (strategic complementary and substitute)
- Making better decisions: default, cooling-off, experience, competition, let others decide for you, team decision, commitment, …
- 2. Doing theory: axiomatization, adding more parameters, adding psychological motivations, **new utility functions, new equilibrium**
- 3. Doing empirics
- Lab experiment: between-subject design, within-subject design
- Field experiment: artefactual, framed, and natural field experiment
- Natural experiment: instrumental variable, DID, regression discontinuity

$\succsim$ is a binary relation on the set of alternatives X, allowing the comparison of pairs of alternatives $x, y \in X$.

COM: $\forall x, y \in X$, we have that $x \succsim y$ or $y \succsim x$.

TRA: $\forall x, y, z \in X$, we have that if $x \succsim y$ and $y \succsim z$, then $x \succsim z$.

} Rational

CON:

WARP: If x is ever chosen when y is available, then there can be no budget set containing x, y for which y is chosen but x is not.

IIA: If $x \in B \subseteq A$ and $x \in C(A)$, then $x \in C(B)$ x B A

Choice Theory under Certainty

- 1. Preference-based approach

Axioms: completeness (COM), transitivity (TRA), continuity (CON) ?

Utility representation theorems: $\forall x, y \in X$, $x \succsim y \iff u(x) \geq u(y)$.

- X is finite, COM + TRA
- X is a general set, COM + TRA + CON

- 2. Choice-based approach: WARP, IIA (Property α), GARP
- 3. Behavioral facts

框架效应 - Framing effect: people change their choices when the changes in the options are inconsequential, different wordings, settings, etc.

禀赋效应 - Endowment effect (lost aversion): people want give additional valuation for an object they own.

维持现状 - Status quo bias (default): people choose an object which they are initially given.

- Asymmetric dominance effect / Compromise effect

  two options are tradeoff, add the third which is strictly dominated by one.

- Mental Accounting.
- Sunk Cost Fallacy.

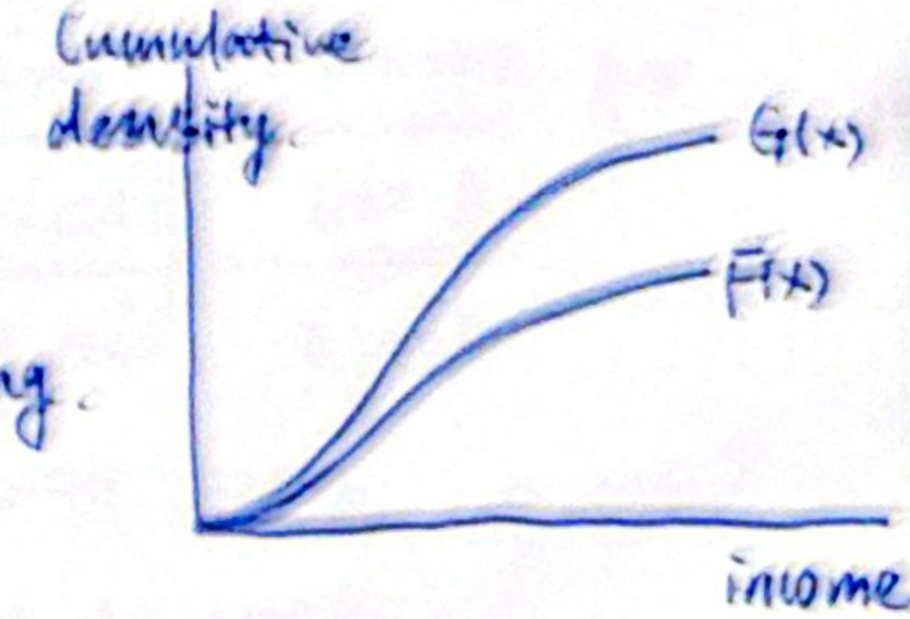
Probabilistic Judgment

$$Pr(B|D) = \frac{Pr(D|B)\cdot Pr(B)}{Pr(D)} = \frac{Pr(D|B)\cdot Pr(B)}{Pr(D|B)\cdot Pr(B) + Pr(D|\bar{B})\cdot Pr(\bar{B})}.$$

- 1. Probability theory: **Bayesian theorem**, law of large number, …
- 2. Heuristics and biases (**Tversky and Kahneman, 1974**) **
  - System 1 and 2 — exaggerate the small probability. — underestimate.
  - **Representativeness**: law of small numbers, conjunction fallacy, disjunction fallacy, base-rate neglect — $P(A\wedge B) \le P(A)$, overestimate
  - Availability
  - Anchoring 锚定点
- 3. Overconfidence
  - overestimate one's ability (over-optimism) the real performance is lower than the estimation.
  - overestimate the precision of one's estimates (over-precision) confidence intervals too narrow.
- 4. Prediction market : the wisdom of the crowded.

Risk averse = prefers the mean of the gamble over the gamble. $u(\alpha x+(1-\alpha)y) \geq \alpha u(x)+(1-\alpha)u(y)$, $\forall x, y, \alpha \in [0,1]$, or $u''(x)<0$.

First-Order Stochastic Dominance: The distribution $F(\cdot)$ yields unambiguously higher returns than the distribution $G(\cdot)$.
For every amount of money x, the probability of getting at least x is higher under $F(\cdot)$ than $G(\cdot)$



COM, TRA: same as before

CON: $\succsim$ on space $\mathcal{L}$ is continuous if for any $L, L', L'' \in \mathcal{L}$, the sets $\{\alpha \in [0,1] : \alpha L + (1-\alpha)L' \succsim L''\} \subset [0,1]$ and $\{\alpha \in [0,1] : L'' \succsim \alpha L + (1-\alpha) L'\} \subset [0,1]$ are closed. remark: small change in Pr don't change the ordering between two lotteries.

IND: $\forall P, Q, R \in \mathcal{L}$, and every $\alpha \in (0,1)$, $P \succsim Q$ iff $\alpha P + (1-\alpha) R \succsim \alpha Q + (1-\alpha) R$. remark: mixture doesn't change the ordering.

Choice under Risk and Uncertainty

- 1. **Expected utility**
- Axioms: COM, TRA, CON, Independence (IND), on the lottery space L
- Expected Utility Theorem: $\forall L = (p_1, \cdots, p_N)$ and $L' (p_1', \cdots, p_N')$, $L \succsim L'$ iff $\sum_{n=1}^{N} u_n p_n \geq \sum_{n=1}^{N} u_n p_n'$ $(n = 1, 2, \cdots, N)$.
- △ 2. **Prospect theory (Kahneman and Tversky, 1979)** **
- Value function: reference dependence, diminishing sensitivity, loss aversion

$$v(x) = \begin{cases} (x-r)^{\alpha}, & x \geq r \\ -\lambda (r-x)^{\beta}, & x < r. \end{cases}$$

$\lambda = 2.25$, $\alpha = \beta = 0.88$ from K&T.

- Probability weighting: nonlinear, 1979 vs. 1992 version
- 3. Paradoxes and facts
  - St. Petersburg Paradox (边际效用递减) $E = \sum_{n=1}^{\infty} \frac{1}{2^n} \cdot 2^n = \infty$.
  - △ Allais Paradox, fourfold risk pattern (确定性偏好) ⇒

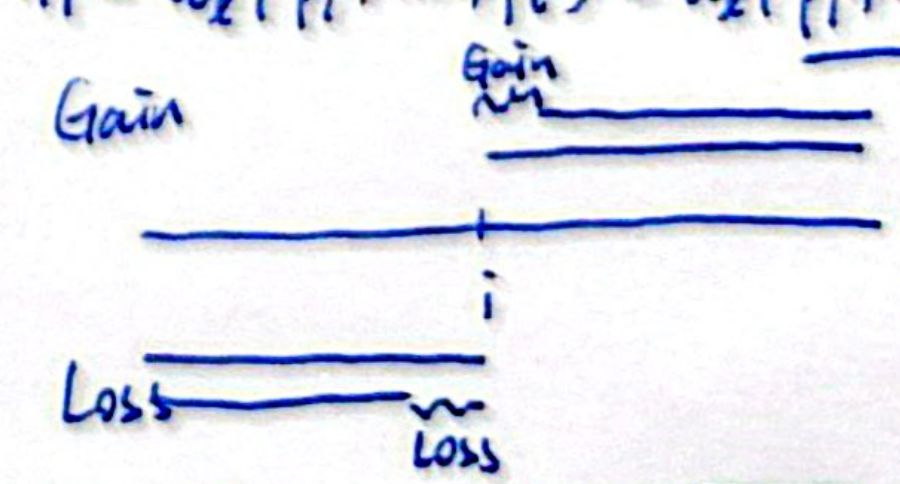
	High Pr	Low Pr
Gain	Risk averse (salary)	Risk seeking (Buying lottery)
Loss	Risk seeking (facing a fine)	Risk averse (Buying insurance)

- 4. Reference dependence utility (status quo or expectation based)
  - Shopping
  - Labor supply for taxi drivers and laboratory tasks

Rank the lottery outcomes from the worst x_1 to the best x_n. $x_1 \sim x_L$ Loss, $x_{L+1} \sim x_n$ gain.
w_g is a function on the decumulative events $x > x_i$:

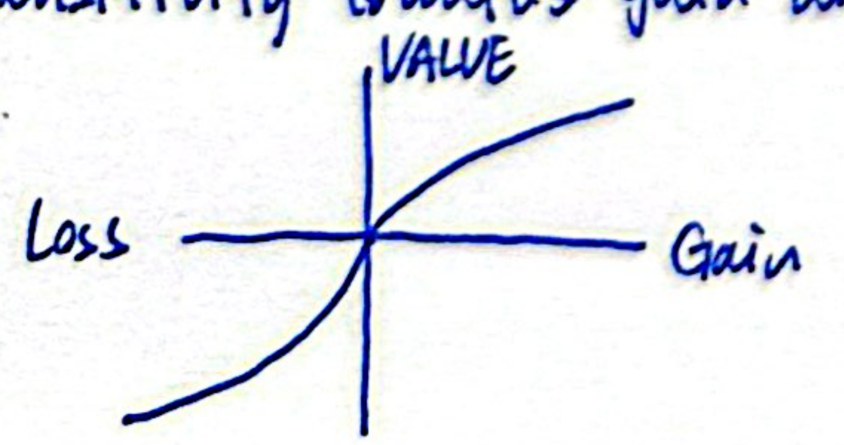
Gain: $\pi_i^+ = w_g(p_i + \cdots + p_n) - w_g(\underline{p_{i+1} + \cdots + p_n})$

Loss: $\pi_i^- = w_l(p_1 + \cdots + p_i) - w_l(\underline{p_1 + \cdots + p_{i-1}})$



$V(L) = \sum_{gains} \pi_i^+ v(x_i) + \sum_{losses} \pi_i^- v(x_i)$.

Valuation Function:
Gain/Loss relative to reference dependence
Diminishing sensitivity towards gain and loss.
Loss aversion.



Reference Point:

1. Linear Function: $U(c,r) = \begin{cases} c+\eta(c-r) & \text{if } c \geq r \\ c+\lambda\eta(c-r) & \text{if } c<r \end{cases}$, where η is the weight for the gain-loss utility, and $\lambda>1$ captures the loss aversion intuition.

e.g.

outcomes	$p_1=0.5$	$p_2=0.5$
$q_1=0.5$	(300, 300)	(200, 300)
$q_2=0.5$	(300, 250)	(200, 250)

utility	$p_1=0.5$	$p_2=0.5$
$q_1=0.5$	$300+\eta(300-300)$	$200+\eta\lambda(200-300)$
$q_2=0.5$	$300+\eta(300-250)$	$200+\eta\lambda(200-250)$

2. Shopping: a pair of shoes price p, reservation value I.

There are two dimensions of the choice, $m(c)=c_1+c_2$. $c_1 \in \{0,1\}$, buy or not, c_2 is wealth.

		Shoes $m(c_1)=c_1$	Payment $m(c_2)=c_2$	Shoes $\mu[m(c_1)-m(r_1)]$	Payment $\mu[m(c_2)-m(r_2)]$	
Classical Theory	Buy	1	$-p$			
	Not Buy	0	0			
Expect to Buy	Buy	1	$-p$	0	0	$1-p > \eta\lambda(0-1)+\eta(0-(-p)) \Rightarrow p \leq (1+\eta\lambda)/(1+\eta)$
	Not Buy	0	0	$\eta\lambda(0-1)$	$\eta(0-(-p))$	Easier to buy.
Not Expect to Buy	Buy	1	$-p$	$\eta(1-0)$	$\eta\lambda((-p)-0)$	$1-p+\eta(1-0)+\eta\lambda((-p)-0) \geq 0 \Rightarrow p \leq (1+\eta)/(1+\eta\lambda)$
	Not Buy	0	0	0	0	More difficult to buy.

Factor 1: Labor force participation
Factor 2: Education
Factor 3: Labor-force experience and work hours
Factor 4: Motherhood
Factor 5: Occupations, industries and firms
Factor 6: Discrimination

Gender

Factor 7: Norms, psychological attributes, and noncognitive skills

Men enter tournament more than women.
like to compete
more overconfident
less risk averse
less averse to feedback

- Gender gap in wage and leadership
- Discrimination, ability differences, preference for competition
- Niederle and Vesterlund (2007): main design

Task 1: Price Rate: each correct answer worth 50 cents.
Task 2: Tournament: who solves most gets $ for each correct answer while others 0.
Task 3: Choice of Compensation Scheme for Future Performance: tournament or not.
Task 4: Choice of Compensation Scheme for Past Piece-Rate Performance: individual choices.

Stage 1: Price Rate Stage 2: Tournament Stage 3: Choice
Stage 4: Choice + Belief Elicitation Stage 5: Choice + Belief Elicitation + Additional Feedback
Stage 6: Non-Competitive Choice List

Types of the decision maker:

TC: Time consistent preference ($\beta = 1$)

Time inconsistent preference: ($\beta < 1$)

Sophisticates: correctly foresee they will have self-control problems.

Naifs: people do not foresee these self-control problems.

A person must perform an activity exactly once. There are T periods.

$v := (v_1, v_2, \cdots, v_T)$ be the reward schedule. $c = (c_1, c_2, \cdots, c_T)$ be the cost schedule

$v_t \geq 0$ and $c_t \geq 0$ for each period.

Immediate costs: $U^t(\tau) = \begin{cases} \beta v_\tau - \underline{c_\tau} & \text{if } \tau = t \\ \beta v_\tau - \beta c_\tau & \text{if } \tau > t \end{cases}$

Immediate rewards: $U^t(\tau) = \begin{cases} \underline{v_\tau} - \beta c_\tau & \text{if } \tau = t \\ \beta v_\tau - \beta c_\tau & \text{if } \tau > t \end{cases}$

Strategy $S := (S_1, S_2, \cdots, S_T)$, $S_t \in \{Y, N\}$.

TCs: for all $t < T$, $S_t = Y$ iff $U^t(t) \geq U^t(\tau)$ for all $\tau > t$.

Naifs: for all $t < T$, $S_t = Y$ iff $U^t(t) \geq U^t(\tau)$ for all $\tau > t$.

Sophisticates: for all $t < T$, $S_t = Y$ iff $U^t(t) \geq U^t(\tau')$, where $\tau' = \min_{\tau > t} \{S_\tau = Y\}$.

e.g. $T = 4$, $\beta = 1/2$, $v = (v, v, v, v)$, $c = (3, 5, 8, 13)$

TCs: (Y, Y, Y, Y) Naifs: (N, N, N, Y) $U^3(3) = 0.5v - 8$ $U^3(4) = 0.5(v - 13)$.

Sophisticates: (N, Y, N, Y) $U^3(3) = 0.5v - 8$ $U^3(4) = 0.5(v - 13)$ $U^2(2) = 0.5v - 5 \geq U^2($

Time Preferences

- Exponential discounting
  - $U_0 = u_0 + \delta u_1 + \delta^2 u_2 + \delta^3 u_3 + \cdots$
- Hyperbolic discounting
  - $U_0 = u_0 + \frac{1}{1+r} u_1 + \frac{1}{1+2r} u_2 + \frac{1}{1+3r} u_3 + \cdots$
- Quasi-hyperbolic discounting:
  - $U_0 = u_0 + \beta\delta u_1 + \beta\delta^2 u_2 + \beta\delta^3 u_3 + \cdots = \underline{u_0 + \beta[\delta u_1 + \delta^2 u_2 + \delta^3 u_3 + \cdots]}$
- Application: **O'Donoghue and Rabin (1999)** **
  - TC, Naifs, Sophisticates: real behavior vs. planned behavior
  - Immediate costs, Immediate rewards sophisticates act before naifs.
- Gul & Pesendorfer (2001) model: set betweenness axiom $U(A) = \max_{x \in A} [u(x) + v(x)] - \max_{y \in A} v(y)$.

  u. long run utility v. short run utility
- Commitment devices: Wertenbroch and Ariely (2002) homework assignments
- Projection bias: Chang, Huang, and Wang (2018) insurance purchase

$\forall A, B$. s.t. $A \succ B \Rightarrow A \succsim A \cup B \succsim B$.

Projection bias: beliefs systematically based toward current state. 健康保险、饿时点餐.

A: every 7 days B. 21 days

C. self-control.

Predictions: neo: $B = C > A$

naive. $A > B = C$

Sophiscated: $C > A > B$

partially: $\underline{A > C > B}$

Neoclassical theory: people are selfish - care only their own consumption

WARP: If A is <u>directly</u> revealed preferred to B, then B is not <u>directly</u> revealed preferred to A.

SARP: If A is <u>indirectly</u> revealed preferred to B, then B is not <u>directly</u> revealed preferred to A.

GARP: If A is <u>indirectly</u> revealed preferred to B, then B is <u>not strictly directly</u> revealed preferred to A.

Violation of WARP $\Rightarrow$ Violation of SARP $\Leftarrow$ Violation of GARP.

Social preferences

- Games: dictator, ultimatum, trust, public goods, gift exchange
- **Fehr and Schmidt (1999): inequality aversion ****
  - $U_{\underline{1}}(x_1, x_2) = \begin{cases} x_1 - \alpha_1(x_2 - x_1)\ if\, x_2 \geq x_1 \\ x_1 - \beta_1(x_1 - x_2)\ if\, x_1 \geq x_2 \end{cases}$

    $\alpha_1 \geq 0$, $\beta_1 \geq 0$: inequity aversion

    $\beta_1 \leq \alpha_1$: suffers more from inequity to his disadvantage

    $\beta_1 < 1$: throw money away suffers more.
- Charness and Rabin (2002): efficiency motivation
  - $u_s(\pi_s, \pi_o) = (\rho r + \sigma s + \theta q)\pi_o + (1 - \rho r - \sigma s - \theta q)\pi_s$

    competitive prefer: $\sigma \leq \rho < 0$.

    selfish prefer: $\sigma = \rho = 0$

    difference aversion prefer: $\sigma < 0 < \rho < 1$

    social welfare prefer: $0 < \sigma \leq \rho \leq 1$.
- Rabin (1992): fairness equilibrium, intention matters

Two players i and j, action $a_i \in A_i$, $a_j \in A_j$, material payoff $\pi_i(a_i, a_j)$, $\pi_j(a_i, a_j)$, belief: First-order (b_i, b_j) Second-order (c_i, c_j).

By choosing a_i, i is choosing $\pi(b_j) = \{\pi_i(a_i, b_j), \pi_j(b_j, a_i) \mid a_i \in A_i\}$.

<u>Def.</u> i's kindness to j. $f_i(a_i, b_j) = \dfrac{\pi_j(b_j, a_i) - \pi_j^e(b_j)}{\pi_j^h(b_j) - \pi_j^{min}(b_j)}$. if $\pi_j^h(b_j) - \pi_j^{min}(b_j) = 0$, $f_i(a_i, b_j) = 0$. $\pi_j^e(b_j) = \dfrac{\pi_j^h(b_j) + \pi_j^l(b_j)}{2}$.

e.g.

Player 1	Player b_2
a_1^1	3, 9
a_1^2	4, 5
a_1^3	7, 1
a_1^4	-1, -1

$f_1(a_1^1, b_2) = \dfrac{9 - \frac{9+1}{2}}{9 - (-1)} = 0.4 > 0.$

<u>Def.</u> i's belief about j's kindness $\tilde{f}_j(b_j, c_i) = \dfrac{\pi_i(c_i, b_j) - \pi_i^e(c_i)}{\pi_i^h(c_i) - \pi_i^{min}(c_i)}$. Utility $u_i(a_i, b_j, c_i) \equiv \pi_i(a_i, b_j) + \tilde{f}_j(b_j, c_i)[1 + f_i(a_i, b_j)]$.

<u>Def.</u> The pair of strategies $(a_1, a_2) \in (S_1, S_2)$ is a fairness equilibrium if, for $i = 1, 2$, $j \neq i$.

(1) $a_i \in \underset{a \in S_i}{\arg\max}\, u_i(a, b_j, c_i)$ (2) $c_i = b_i = a_i$. Actions are consistent with belief.

Level-0 players: play at random with $P_r = 1/N$ in each of their strategies
Level-1 players: Think that all the other players play at Level-0.
Level-2 players: Think that all the other players play at Level-1.
...

Limited Strategic Thinking

- Mixed strategy equilibrium: proposition
- △ **Level-k and cognitive hierarchy model**: best response, non-equilibrium ** : no mutual consistency between belief and action.
- (QRE) Quantal response equilibrium: better response with noise, equilibrium
- Applications   A QRE, σ^* is a fixed point of the mapping $\pi(\sigma^*) = \sigma^*$. That is for all players and strategies, $\sigma_{ij} = \pi_{ij}(\bar{u}_{ij}(\sigma))$.
  - △ Beauty contest game
  - Coordination game
  - Market entry game
  - Matching penny game, hide and seek game
  - Penalty kick in football

实验设计三原则: 对照, 随机, 重复.

Completely randomization: minimizes the risk that the treatment is correlated with subjects characteristics, but the variance of outcome may large.

Factorial design: examine the interaction between the two treatments. 2×2 design is popular.

Block design: randomization is performed within, but not between blocks.

Within-subject design: eliminate the subject-specific effect and greatly improve the precision of the estimates requires a smaller sample size.

Crossover design: uses AB and BA sequences to cancels out the order effect.

Experimental Design

- Potential outcome framework and ATE $ATE = E[Y_i(1)] - E[Y_i(0)] = E[Y_i(1)|T=1] - E[Y_i(0)|T=0]$.
- Randomization
  - completely randomization, factorial design, block design, within-subject design, crossover design
  - level of randomization
- Power analysis
  - Optimal sample size
  - Replication crisis

P-hacking

The significance level (Type-I error).

The power of the test (Type-II error).

The minimal detectable effect size.

$$t = \frac{\bar{y}_2 - \bar{y}_1}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \overset{n_1=n_2}{=} \frac{\bar{y}_2 - \bar{y}_1}{s_p\sqrt{\frac{2}{n}}} \sim t_{n-2}.$$

$$s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$$ pooled sample s.d.