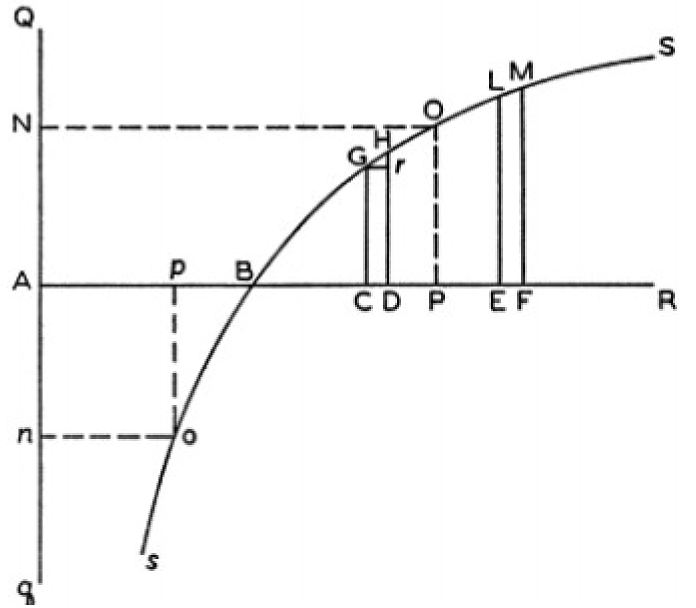
Decision Theory: Explaining Choice Under Risk

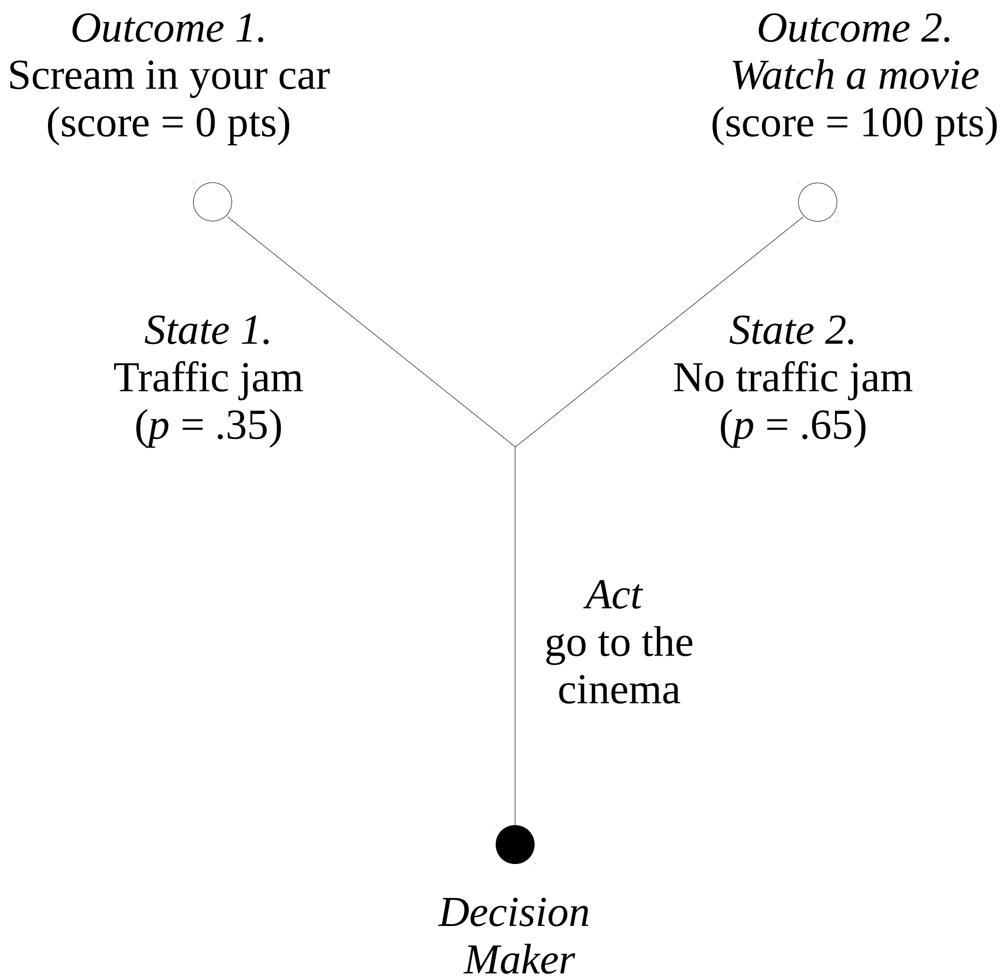
Dr Goran S. Milovanović

Data Scientist

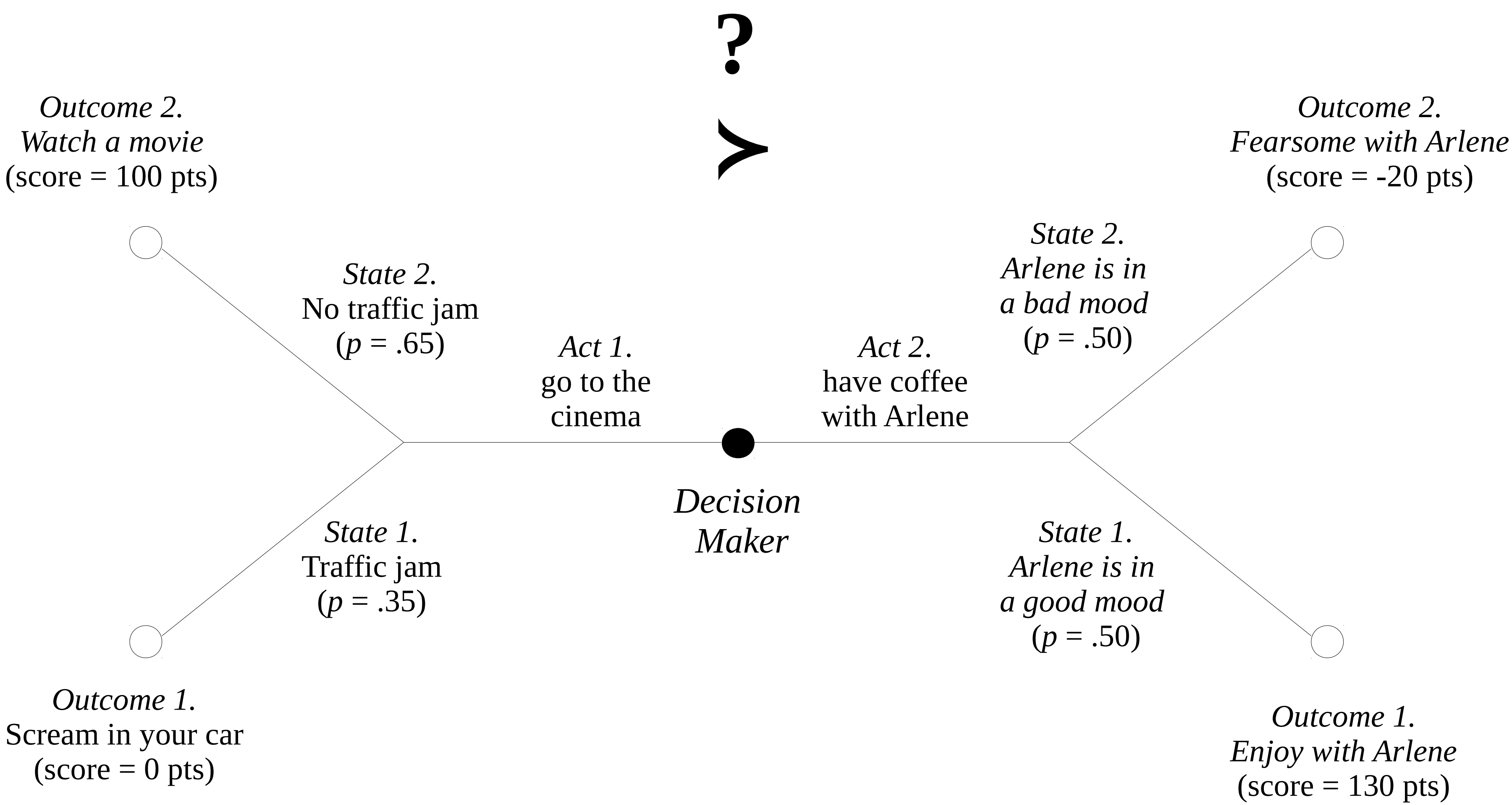
Wikimedia Deutschland, Berlin
DataKolektiv, Belgrade



The Problem of Choice Under Risk



The Problem of Choice Under Risk



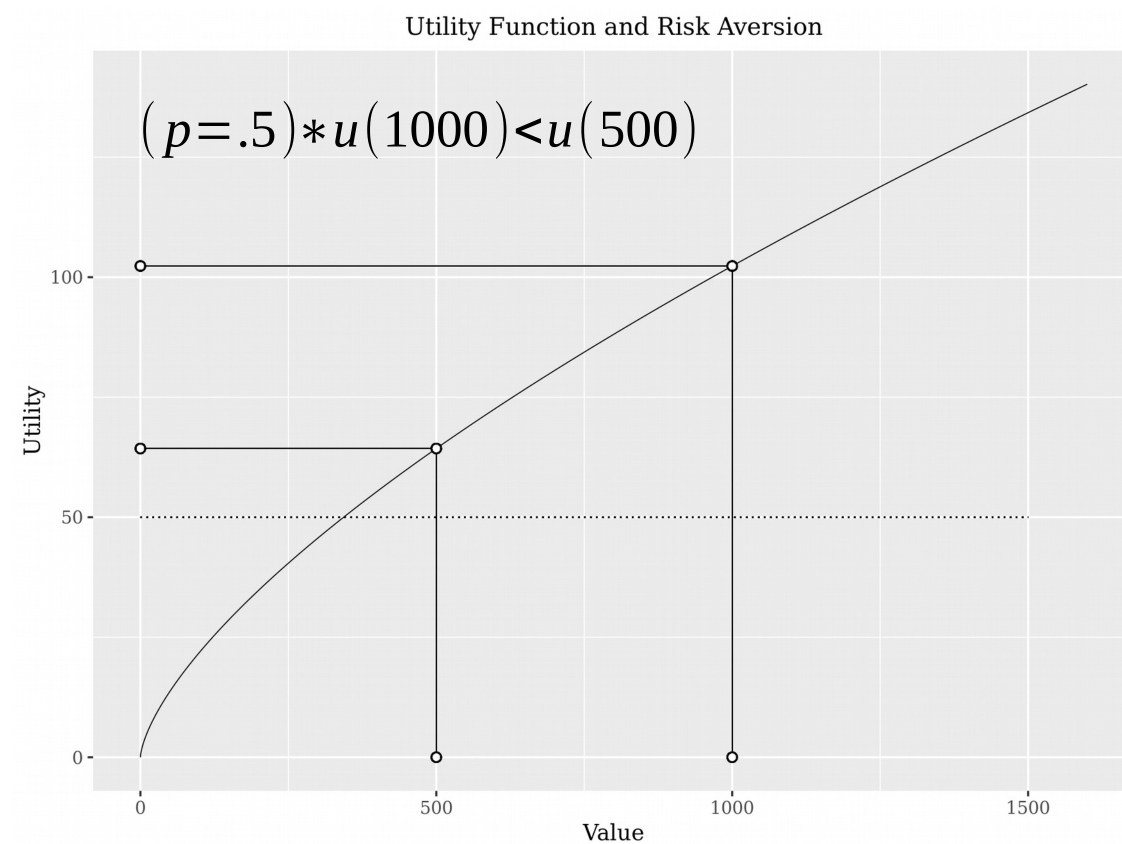
Risk Attitudes and Utility

50% to win EUR 1000 / 50% to win nothing	?	Sure thing EUR 500 **Risk Aversion**

$$EV\left(\left(p=.5\right)*\left(v=1000\right)+\left(p=.5\right)*\left(v=0\right)\right)=500$$

1738, *Specimen theoriae novae de mensura sortis* (*Exposition of a New Theory on the Measurement of Risk*)

Use not value (of money), but utility (of money), more like *log*(money)...



$$\left(p=.5\right)*u\left(1000\right)<u\left(500\right)$$

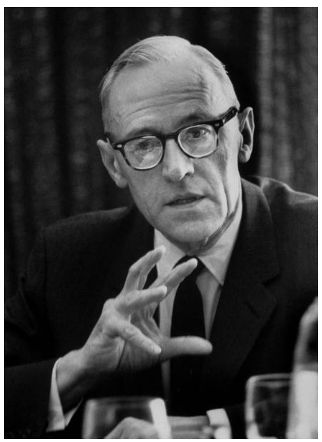
Power utility

$$u\left(v\right)=v^{\rho}$$

Exponential utility

$$u(c)=\begin{cases}(1-e^{-ac})/a & a\neq 0\\ c & a=0\end{cases}$$

Axiomatic Foundations of Utility



Theory of Games and Economic Behavior (1944, 2nd edition 1947): If the following set of axioms + some of more technical importance hold:

A1 *Completeness.* For all *p*, *q*: either $p \geq q$ *or* $q \geq p$.

A2 *Transitivity*. For all *p, q, r*: if $p \geq q$, i $q \geq r$, then $p \geq r$.

A3 *Independence*. For all *p*, *q*, *r* and any real number α between 0 and 1: $p \geq q$ iff $\alpha \cdot p + (1- \alpha) \cdot r \geq \alpha \cdot q + (1- \alpha) \cdot r$.

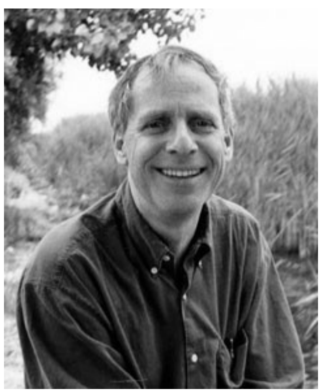
then a decision maker will make choices under risk *as if* he maintains a Bernoulli utility function *u*(*v*) and follows the Principle of Maximum Expected Utility:

$$EU(p;V) = \sum_j p(v_j) u(v_j)$$

Paradoxes of Expected Utility

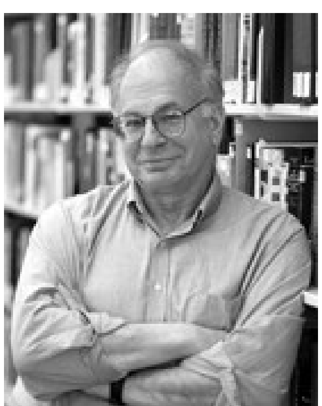
5% to win EUR 1000 / 95% to win nothing	*Certainty Equivalent* = EUR 60	95% to win EUR 1000 / 5% to win nothing	*Certainty Equivalent* = EUR 890
	Risk Seeking		***Risk Aversion***

This is not possible if a decision maker is characterized by a single utility function.



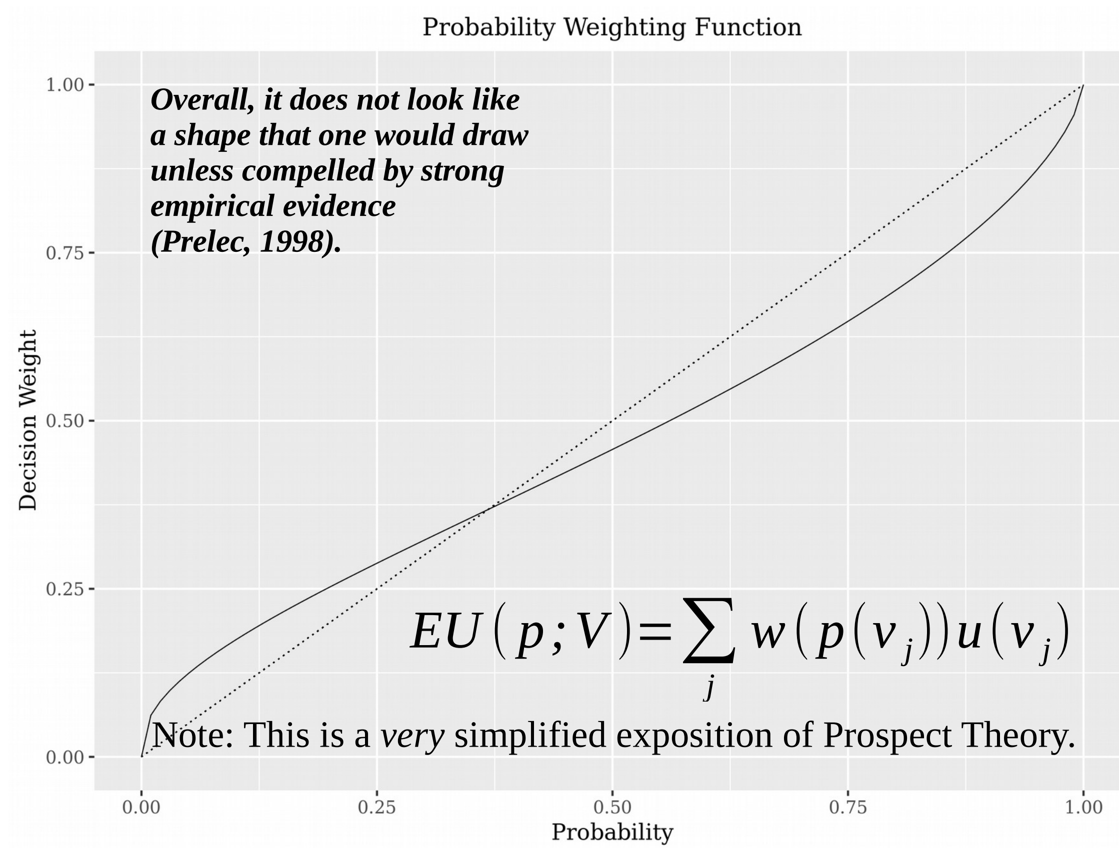
Kahneman & Tversky (1979), Tversky & Kahneman (1992): **Prospect Theory**.

Many different probability weighting functions are used in practice.



The *Prelec one-parameter form*:

$$w(p) = e^{-((-\log(p))^{\gamma})}$$



Axioms for Behavioral Choice Under Risk

Let α, β, γ, δ be some outcomes.

Let P, Q be any two lotteries characterized by the *same* probability distribution:

P: $(p_1, x_2; p_2, x_2;.. ; p_n, x_n)$, Q: $(p_1, \mathrm{y}_1; p_2, y_2;.. ;p_n, y_n)$

A lottery α_jP is obtained by replacing the *j*-th value in P by α; in an analogous way we obtain β_jQ, γ_jP, *and* δjQ.

Let's introduce a *tradeoff relation*, $\succeq^*$ by

$\alpha, \beta \succeq^* \gamma, \delta$ *iff* $\alpha_j \mathrm{P} \succeq \beta_j \mathrm{Q}$ *and* $\gamma_j \mathrm{P} \prec \delta_j \mathrm{Q}$.

Tradeoff Consistency: there are no outcomes α, β, γ, δ such that $\alpha, \beta \succ^* \gamma, \delta$ *and* $\alpha, \beta \precsim^* \gamma, \delta$ hold.

In Expected Utility, this implies that there are no outcomes α, β, γ, δ such that $\mathrm{u}(\alpha) - \mathrm{u}(\beta) \geq \mathrm{u}(\gamma) - \mathrm{u}(\delta)$ and $\mathrm{u}(\alpha) - \mathrm{u}(\beta) < \mathrm{u}(\gamma) - \mathrm{u}(\delta)$.

Add completeness, transitivity, and monotonicity and EU is axiomatized.

(Wakker & Tversky, 1993).

Axioms for Behavioral Choice Under Risk

In Rank-Dependent Utility (RDU) theories (upon which Prospect Theory is developed):

Define *comonotonic lotteries*:

P		**Q**
15% 35% 50%	$\succ$	15% 35% 50%
$\$50 < \$25 < \$5$		$\$10 < \$5 < \$1$

P’		**Q’**
10% 30% 60%	$\succ$	10% 30% 40%
$\$50 < \$25 < \$5$		$\$10 < \$5 < \$1$

Tradeoff Consistency: there are no outcomes $\alpha, \beta, \gamma, \delta$ such that $\alpha, \beta \succ^* \gamma, \delta$ *and* $\alpha, \beta \precsim^* \gamma, \delta$ hold.

In Expected Utility, tradeoff consistency *holds for all lotteries*.

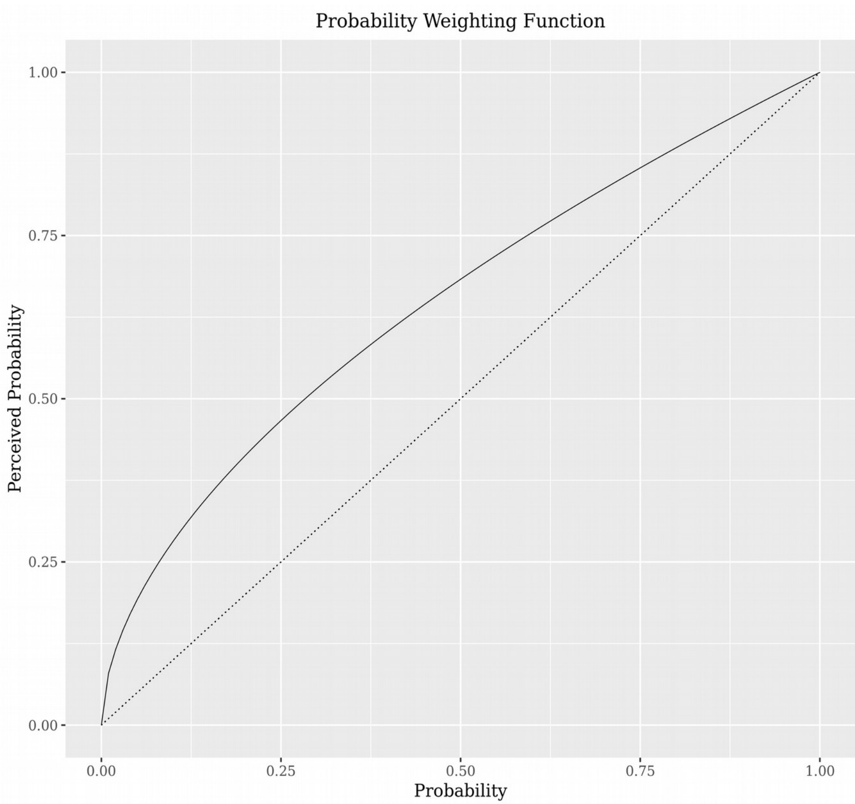
In RDU theories, tradeoff consistency *holds for comonotonic lotteries only*.

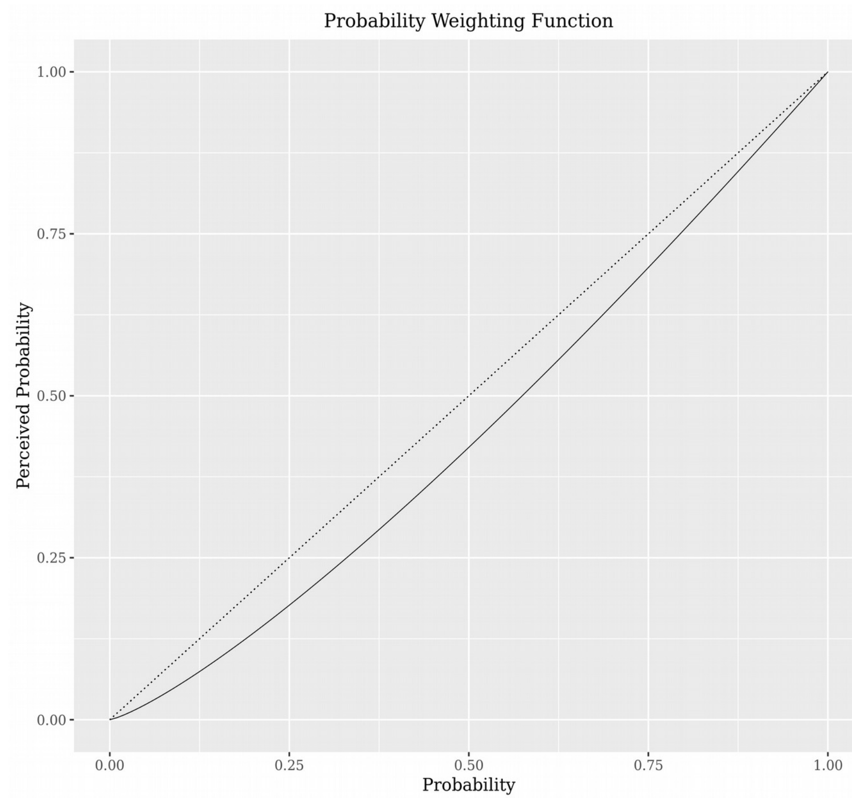
In Prospect Theory, tradeoff consistency *holds for comonotonic, co-signed lotteries only*.

Alternative History of Probability Weighting

E.g. a monotonic function of probability (c.f. Birnbaum, 2008)

$$t(p) = p^{\tau}$$





Gain domain: **optimism**
Loss domain: **pessimism**

Gain domain: **pessimism**
Loss domain: **optimism**

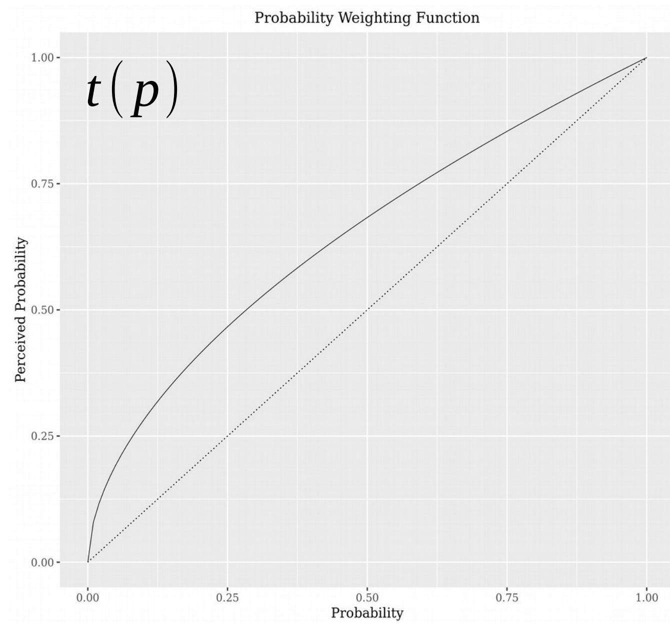
Let's assume that *all* non-cognitive, non-perceptual factors that could influence choice exhibit their effects *solely* on the shape of this function, reflecting in optimism/pessimism attitudes towards probabilities.

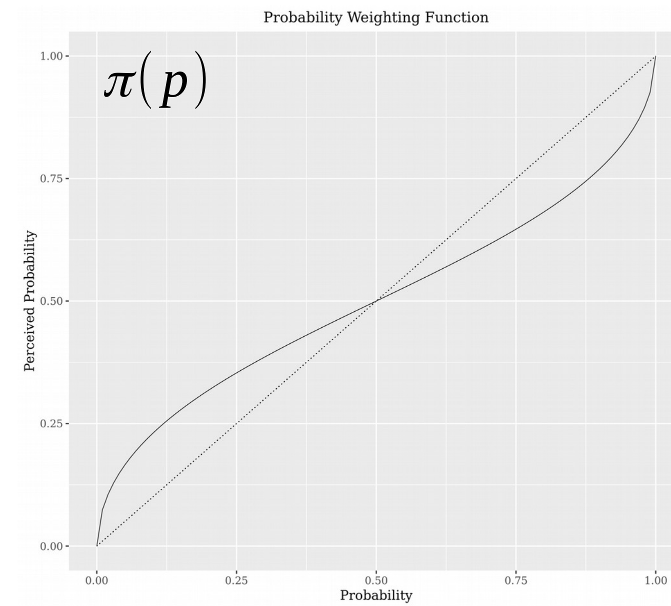
Alternative History of Probability Weighting

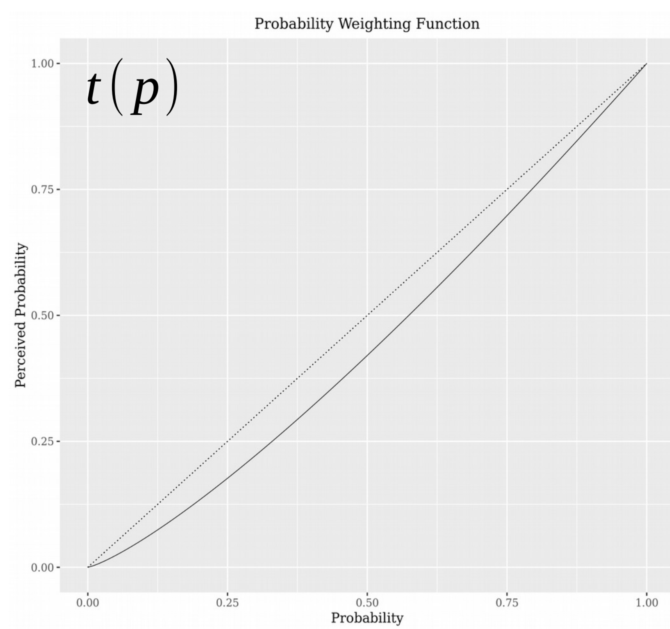
Spence (1990)

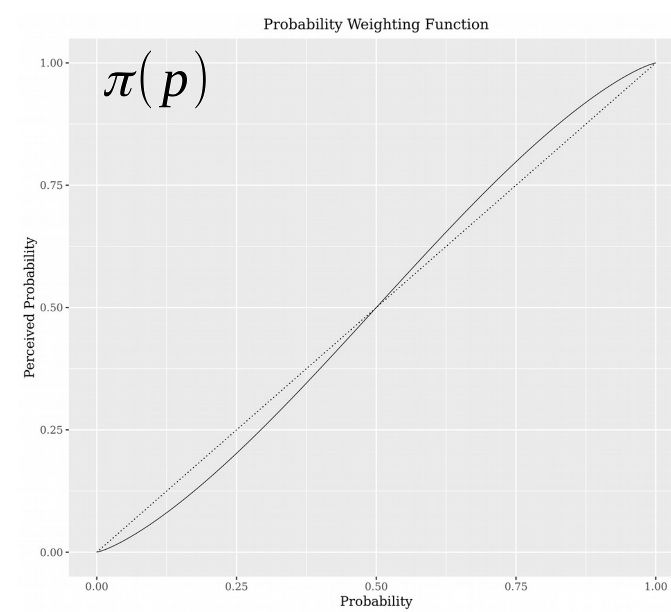
$$t(p) = p^{\tau}$$

$$\pi(p) = \frac{t(p)}{\sum_{j=1}^{n} t_j}$$









Alternative History of Probability Weighting

Viscusi (1989) Perspective Reference Theory

Lottery L:(x, p_x; y, p_y) is a binomial distribution:

$$p(x;p_x)=\binom{n}{x}p_x{}^{x}(1-p_x)^{n-x}$$

It's conjugate prior in Bayesian inference is Beta(α, β):

$$Beta(\alpha,\beta)=\frac{1}{B(\alpha,\beta)}p_x{}^{\alpha-1}(1-p_x)^{\beta-1}$$

The posterior follows a Beta distribution with:

$$\alpha''=\alpha+x$$

$$\beta''=\beta+n-x$$

Note. Extension to a multinomial case with Dirichlet conjugate priors is straightforward.

Alternative History of Probability Weighting

Viscusi (1989) Perspective Reference Theory

Let's introduce an alternative parametrization of the Beta distribution:

$$\gamma = \alpha + \beta$$

γ is the "*informational content of the individual's prior beliefs*" - a virtual sample size upon which the prior belief was developed, the strength of prior.

$$\Theta_p = \alpha / (\alpha + \beta)$$

Θ is the mean of the Beta distribution, and we will use it as our estimate of $\boldsymbol{p_x}$.

Under this choice of parameters, the estimate of the mean of the Beta posterior, $\boldsymbol{p''_x}$ takes the following form:

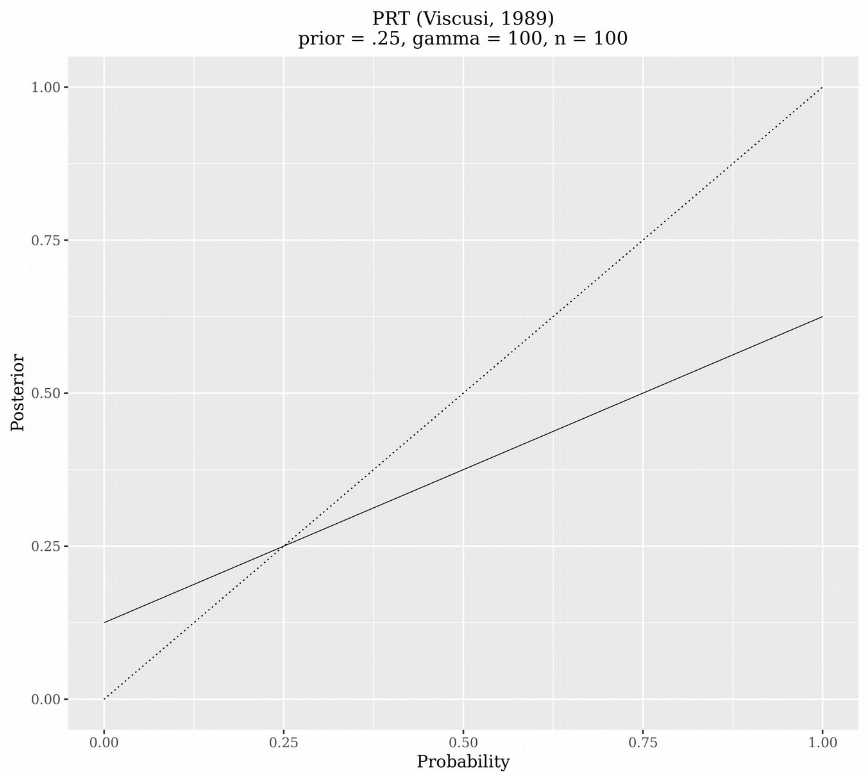
$$p''_x = \Theta_{p_x} = \frac{\gamma p'_x + n p_x}{\gamma + n}$$

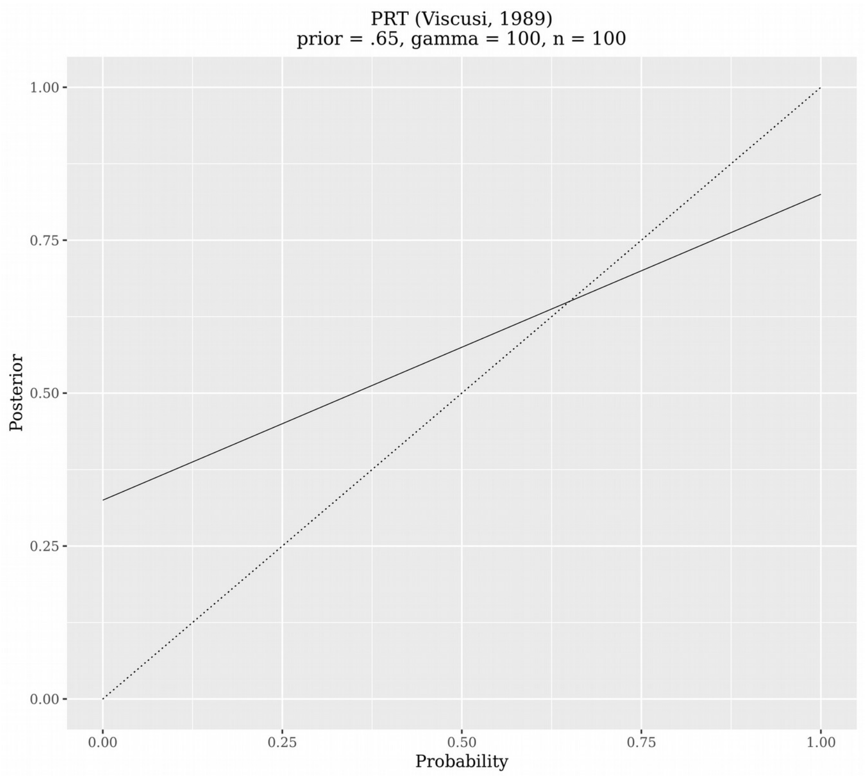
which is a *linear function* of the stated probability $\boldsymbol{p_x}$, with a slope of $n/(\gamma+n)$ and an intercept (controled by the prior and γ) of $\gamma p'_x/(\gamma+n)$.

Alternative History of Probability Weighting

Viscusi (1989) Perspective Reference Theory

Probability weighting in Viscusi’s PRT



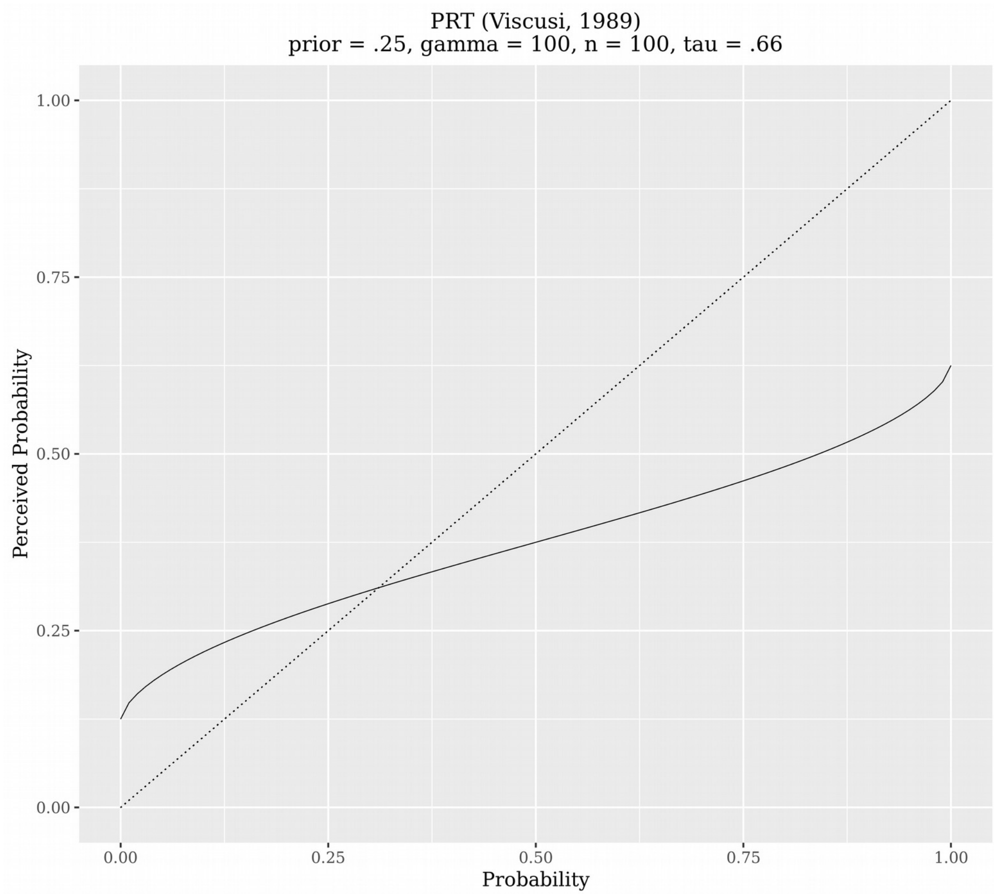


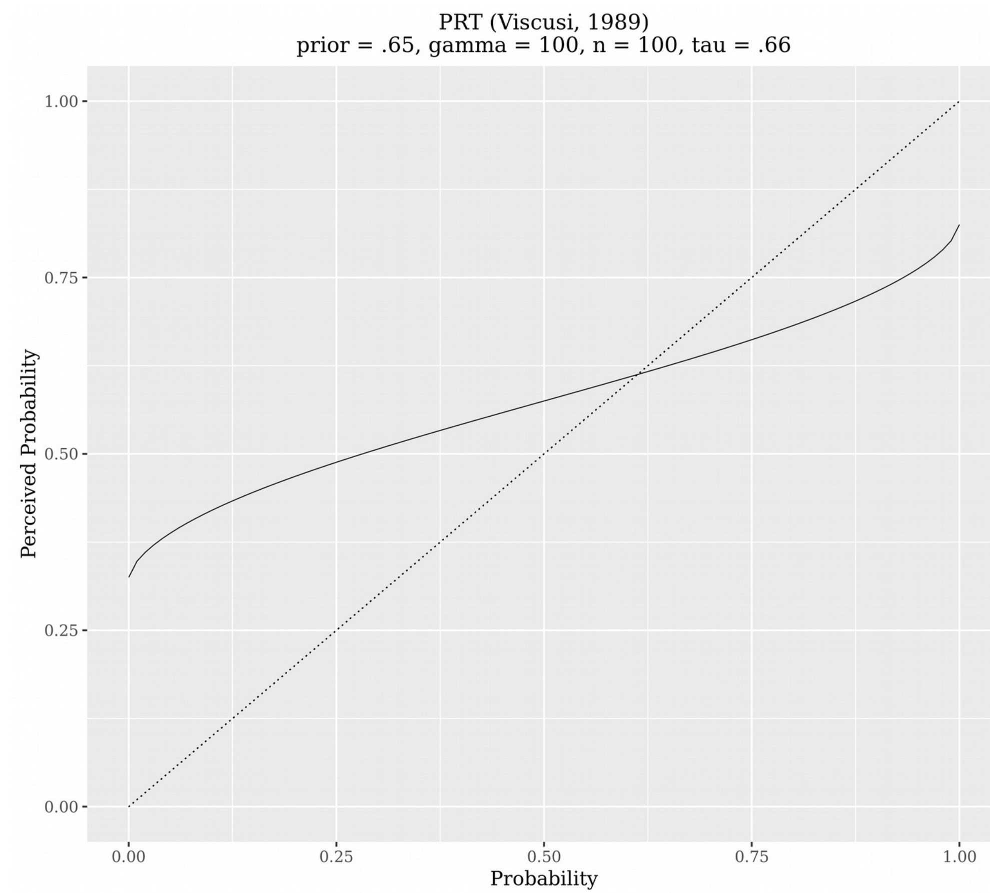
Note. This form of probability weighting alone can explain the most robust behavioral deviations from the vNM Expected Utility.

Alternative History of Probability Weighting

Introduce a monotonic transform of probability to the PRT weighting mechanism:

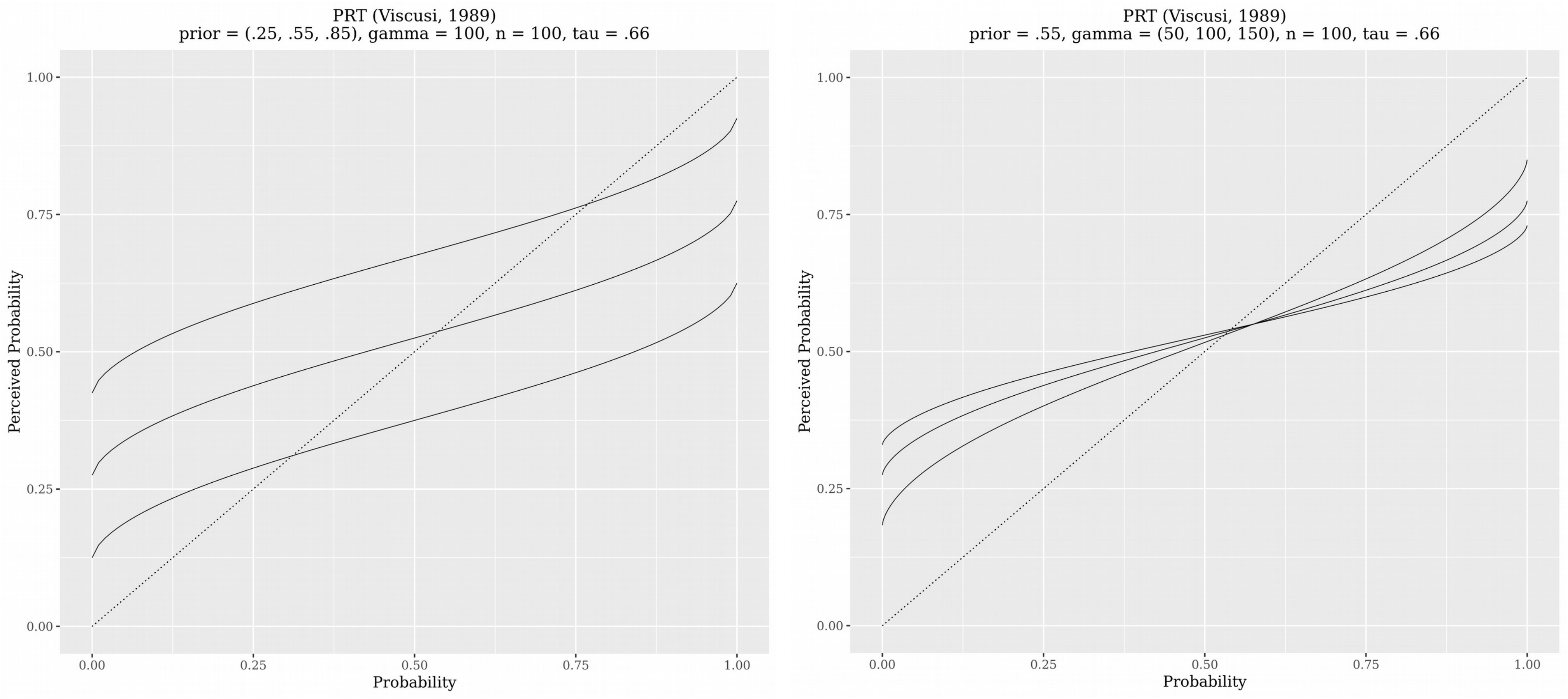
$$t(p) = p^{\tau} \qquad \pi(p) = \frac{t(p)}{\sum_{j=1}^{n} t_j}$$





Explaining Probability Weighting

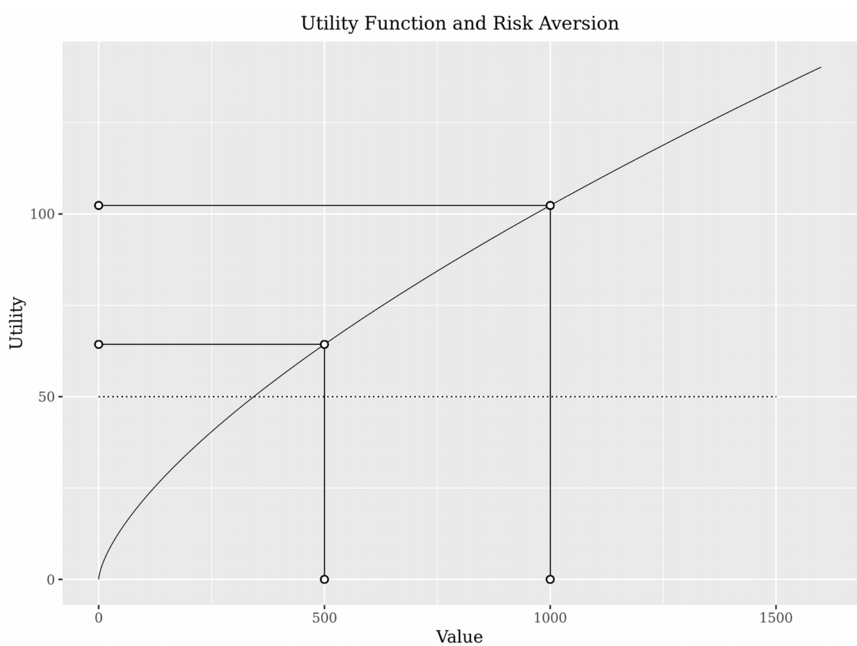
… by simple Bayesian inference + normalized monotonic probability transform



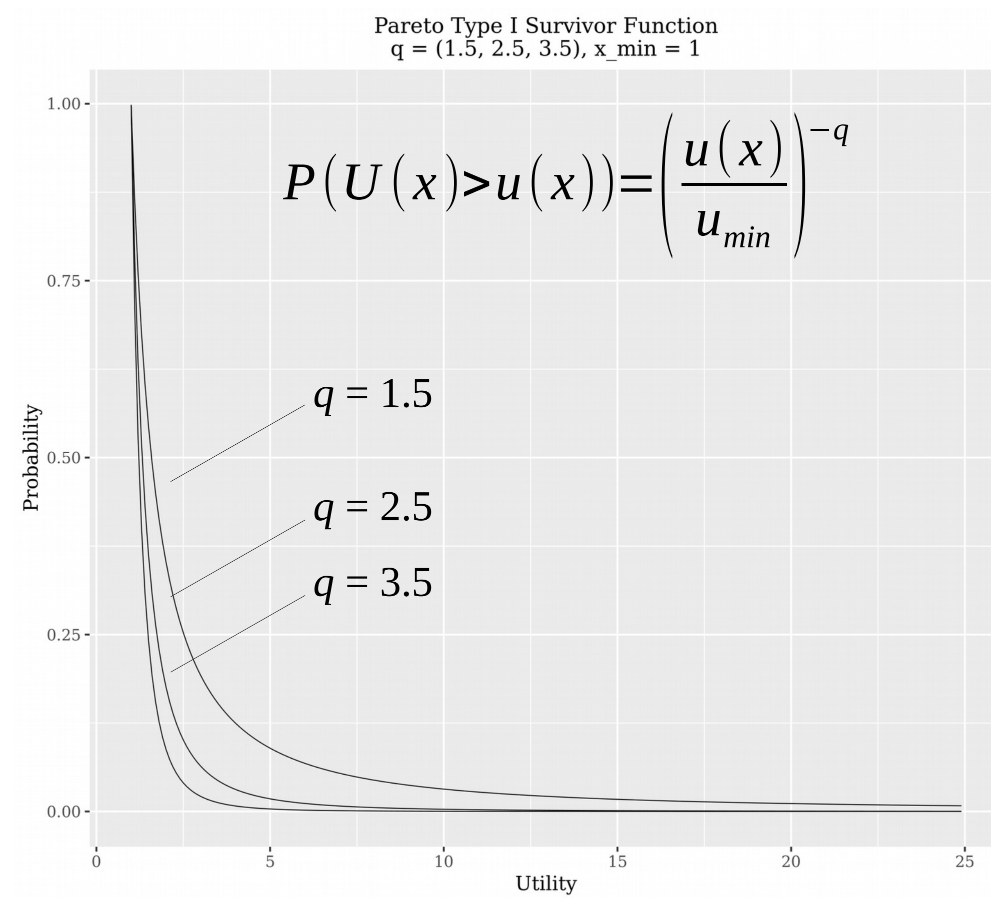
On the Origin of Priors

Viscusi (1989) noted, but did not elaborate on a *multiple reference point case* in the Perspective Reference Theory, where the decision maker does not hold that all a priori probabilites are equal.

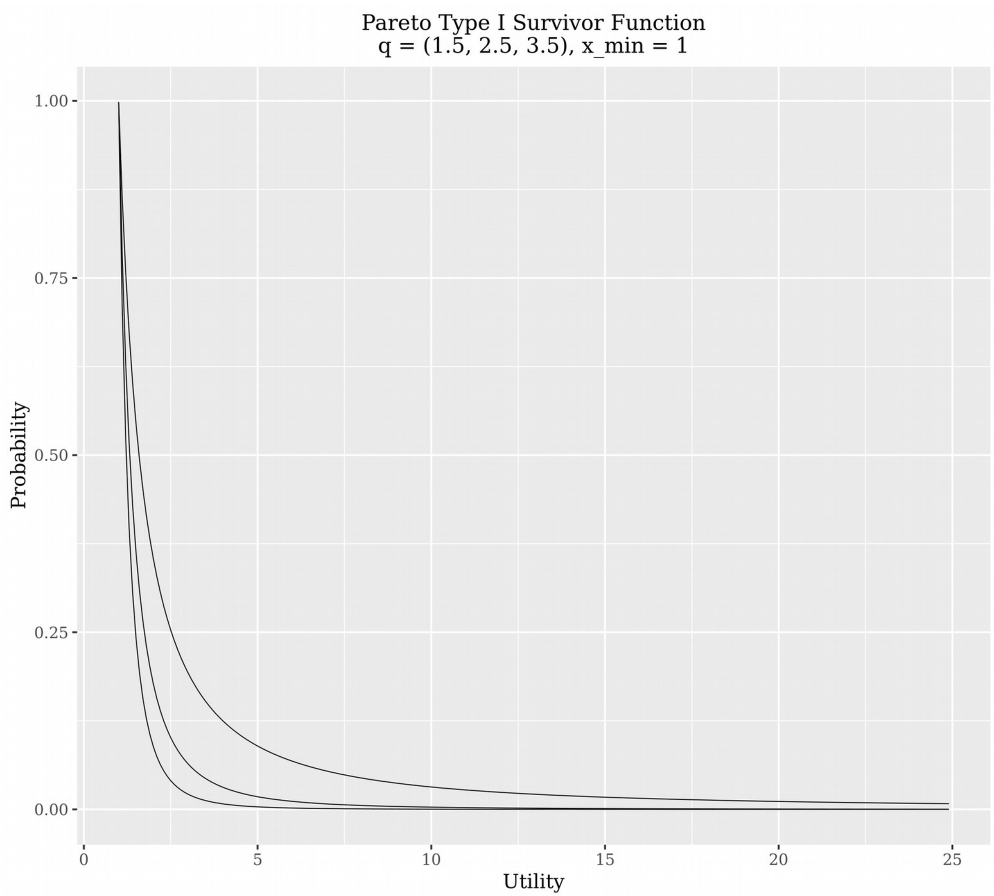
We are dealing with two parameters: the *prior distribution over lottery outcomes*, and γ, the strength of the prior. How do we determine the values of these parameters?



$S(u(x))$: the distribution of the prior belief on $u(x)$



Confidence Theory: On the Origin of Priors



1. Lottery L:(x, p_x; y, p_y) is a binomial distribution.
2. What is the prior belief on $S(u(x))$?
3. What is the prior belief on $S(u(y))$?
4. Prior beliefs p'_x and p'_y are:

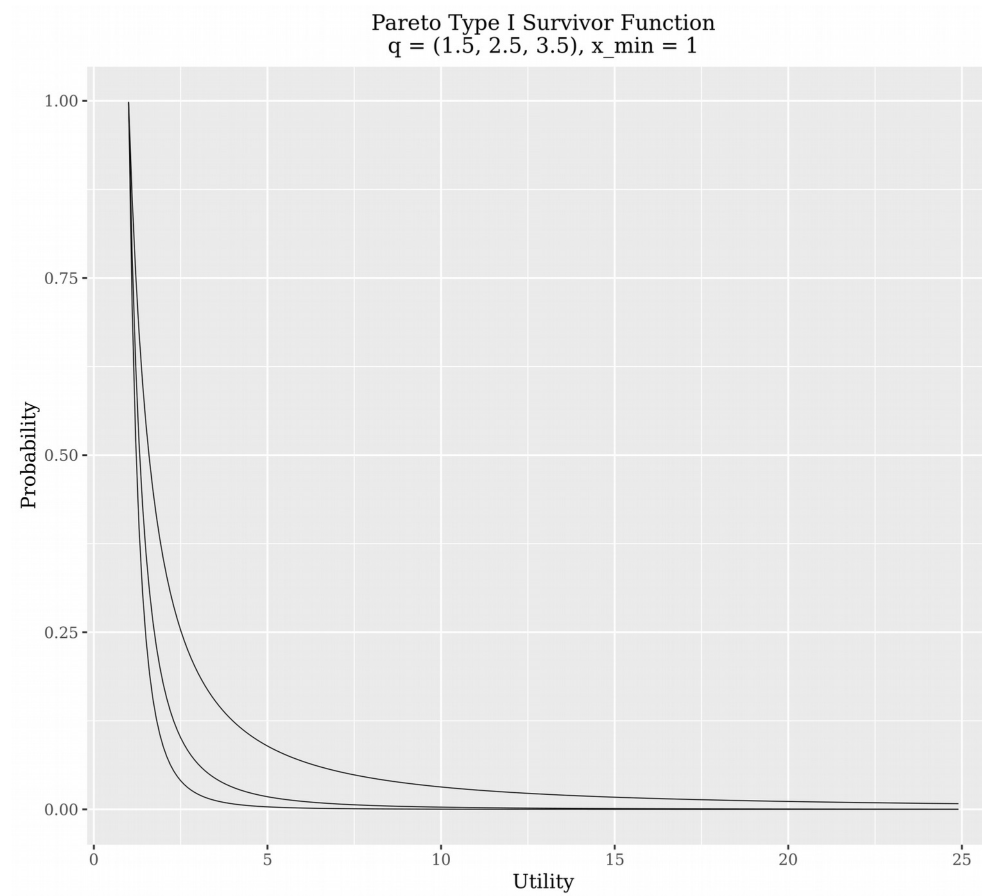
$$p'_x = \frac{s(u(x))}{s(u(x)) + s(u(y))}$$

$$p'_y = \frac{s(u(y))}{s(u(x)) + s(u(y))}$$

N.B. What happens with the prior distribution when *x* and *y* are *similar* (i.e. close in value and hence close in utility)?

Confidence Theory: On the Strength of Priors

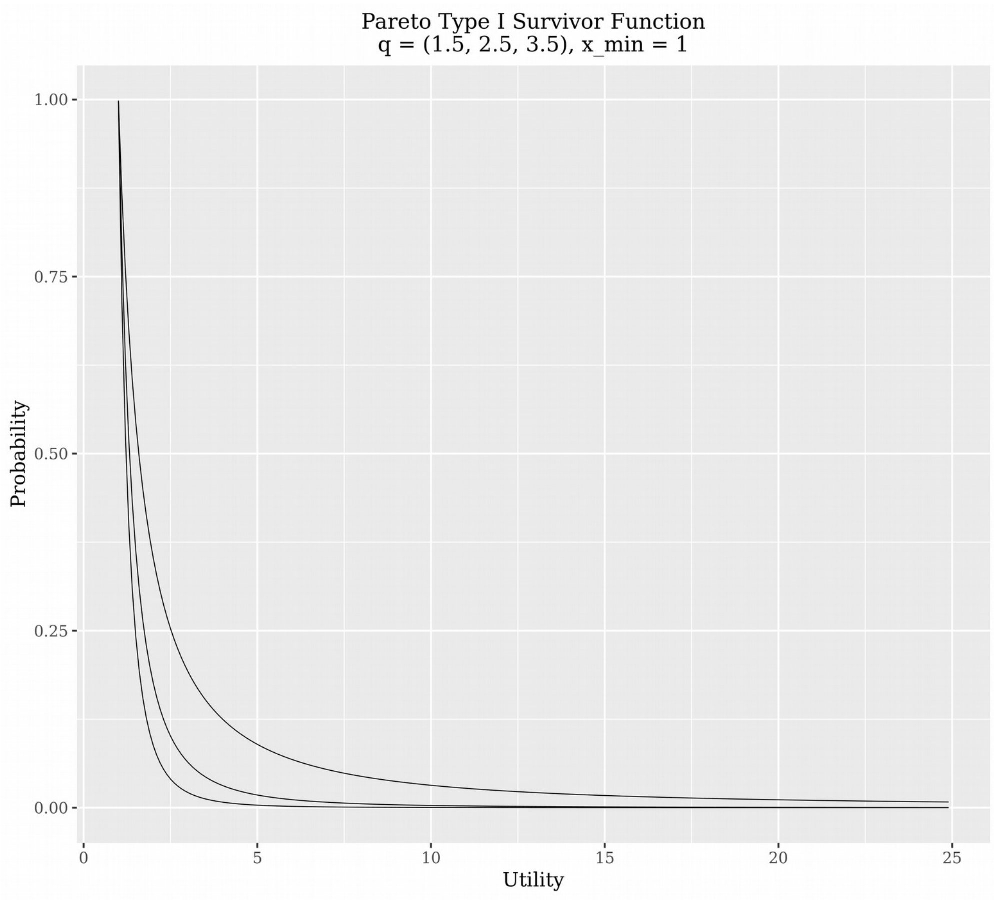
5% to win EUR 10 / 95% to win EUR 2000	95% to win EUR 10 / 5% to win EUR 2000
5% to win EUR 100 / 95% to win EUR 101	51% to win EUR 100 / 49% to win EUR 101



Observation. In our natural environments, where real decisions are made, *similar outcomes* occur with *similar probabilities*.

Confidence Theory: On the Strength of Priors

What happens with the prior distribution when *x* and *y* are *similar* (i.e. close in value and hence close in utility)? *Its entropy increases.*



$$p'_x = \frac{s(u(x))}{s(u(x)) + s(u(y))}$$

Shanon’s Diversity Index

$$\Omega = \frac{H(p')}{H_{max}} n$$

We will use Ω as our measure of γ
- the strength of prior beliefs.

More entropic prior distributions of utility will gain more power in Bayesian inference.

Confidence Theory: The Decision Model
[ρ, τ, q, x_{min}]

Probability

$$t(p) = p^{\tau}$$

$$\pi(p) = \frac{t(p)}{\sum_{j=1}^{n} t_j}$$

Prior

$$p'_x = \frac{s(u(x))}{s(u(x)) + s(u(y))}$$

$$\Omega = \frac{H(p')}{H_{max}}$$

$$\gamma = \Omega n$$

Posterior

$$p''_x = \frac{\gamma p'_x + n \pi_x}{\gamma + n}$$

Posterior Expected Utility

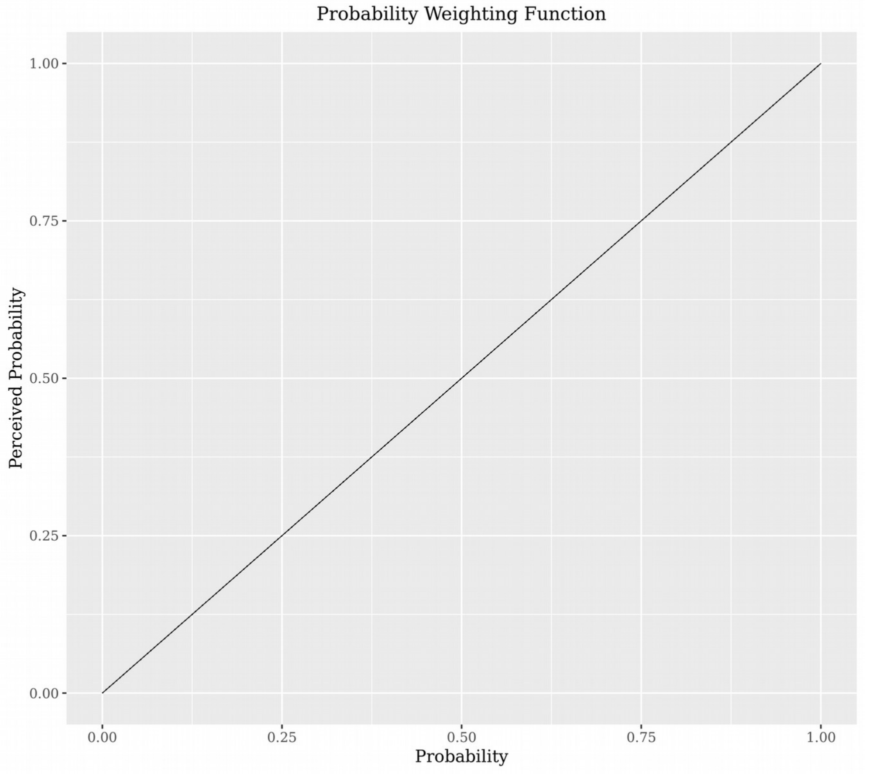
$$EU''(p;V) = \sum_j p''(v_j) u(v_j)$$

$$EU''(p;V) = \frac{\gamma}{\gamma + n} p'_x u(V) + \frac{n}{\gamma + n} \pi_x u(V)$$

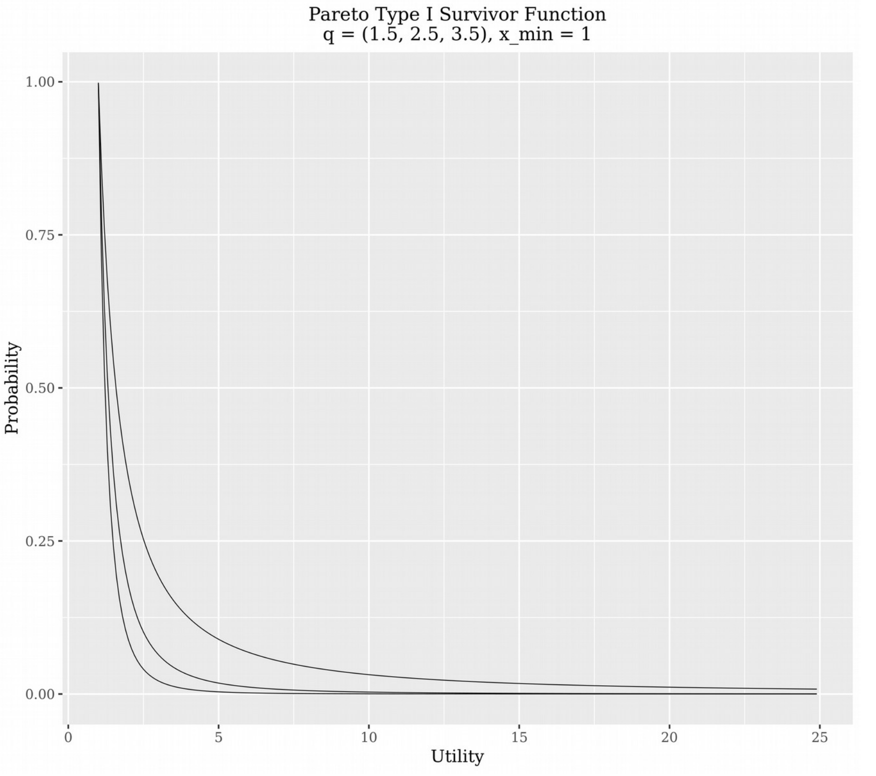
Rationality in the sense of vNM under Confidence Theory

When does a decision maker make rational choices in sense of von Neumann & Morgenstern's axioms?

Condition 1. Objective perception of probability (no optimism, no pessimism).



Condition 2. The internal model of the environmentaly relevant probability distribution of value under consideration is true.



Note on Condition 2. When prior and stated probabilites are equal, Bayesian inference has no effect and the decision model is vNM Expected Utility (given that the Condition 1 holds; otherwise, it reduces to some form of Subjective Expected Utility (SEU).

Empirical Tests

A. Measurment of Certainty Equivalents Data Sets

- **Milovanović (2013), Experiments 2a, 2b**: Confidence Theory *without* the monotone probability transform *t*(*p*) achieves lower RMSE values than Prospect Theory (CPT model).

- **Milovanović (2014), re-analysis of the Gonzales & Wu (1999) data, Bayesian model selection procedure**: Confidence Theory outperforms both Prospect Theory (CPT model) and Birnbaum's (2008) TAX, irrespective of whether a power or an exponential utility function is assumed; it explains the violations of preference homogeneity while retataining the power utility function, which is not possible in Prospect Theory.

B. Choice Experiment Data Sets

- **Milovanović (2014), re-analysis of Birnbaum's (2008) new paradoxes of risky choice**:

- all experimental findings in these data sets *falsify* Prospect Theory (*i.e.* no C(PT) parametric model and no combination of parameters allows for the respective violations);

- Confidence Theory numerical simulations and model fits show that it can reproduce *every observed behavioral pattern* that violates Prospect Theory in this series of experiments.

“(*He must so to speak throw away the ladder, after he has climbed up on it.*) He must transcend these propositions, and then he will see the world aright.”

(Ludwig Wittgenstein,
***Tractatus Logico-Philosophicus,* 1921.)**

A Strategy to Develop a Decision Theory

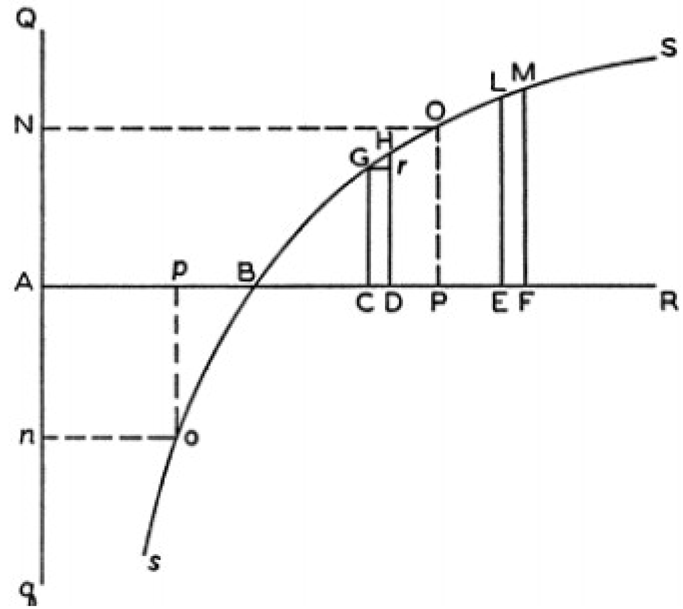
Formal analysis

The complexity of formal, axiomatic analysis of choice under risk (and uncertainty) might well be a consequence of ***the approach that attempts to axiomatize choice functions in isolation***, i.e. without previously considering what other – possibly well-known – cognitive functions influence decision making.

Known empirical principles

For example, the normalizations used in Confidence Theory – applied to ensure that the decision maker always operates on a probability scale – are *ubiquitous* in cognition and perception. If *we know something about the human cognitive system, we know that it is* **sensitive to relative and not absolute magnitudes** of the environmental stimuli.

Furthermore, the idea that organisms adapt by planning their actions in respect to their prior experiences – in other words, that they adapt by ***learning*** - is a prominent idea in any behavioral science. In Confidence Theory *we have assumed that the decision maker has some prior beliefs about the probability to improve or worsen upon its present condition in units of utility*. That assumption is well-aligned with the classic form of explanation in behavioral sciences.



Q
S
L
M
N
O
H
G
r
A
p
B
C
D
P
E
F
R
n
o
s
q