What are we weighting for?

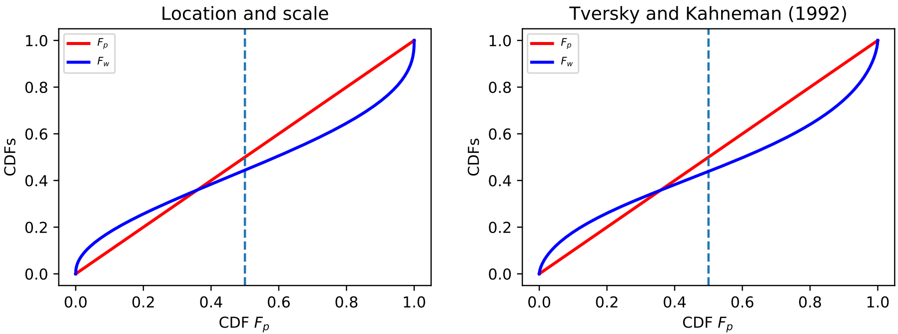
A mechanistic model for probability weighting

Ole Peters Alexander Adamou Yonatan Berman Mark Kirstein

D-TEA 2020, 16 June 2020

Max Planck Institute for Mathematics in the Sciences
UNIVERSITÄT LEIPZIG
Mathematisches Institut
LONDON MATHEMATICAL LABORATORY

Main results

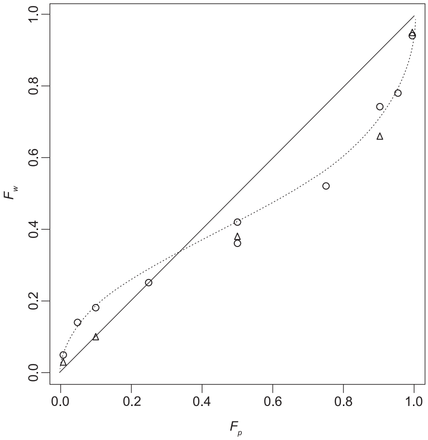


XXX label on left figure: change to "Ergodicity economics" to mirror K&T?

1. inverse-S shape can be explained by difference in uncertainty
2. cautious estimation of probabilities generates such differences

▸ PW K&T 1979

Definition of Probability Weighting (PW)



(Tversky and Kahneman 1992, p. 310, Fig. 1,relabelled axes)

- low probabilities treated as higher; high probabilities treated as lower
- stable empirical pattern: inverse-S shape

Received wisdom:

- PW = maladaptive irrational cognitive bias

In search of a mechanism

↪ How does this pattern emerge?

↪ Can we derive a functional form (rather than fit a function)?

Setup

Task: model payout, x, of a gamble as a random variable.

Disinterested Observer (DO)

DO assigns
probabilities $p(x)$
CDF $F_p(x)$

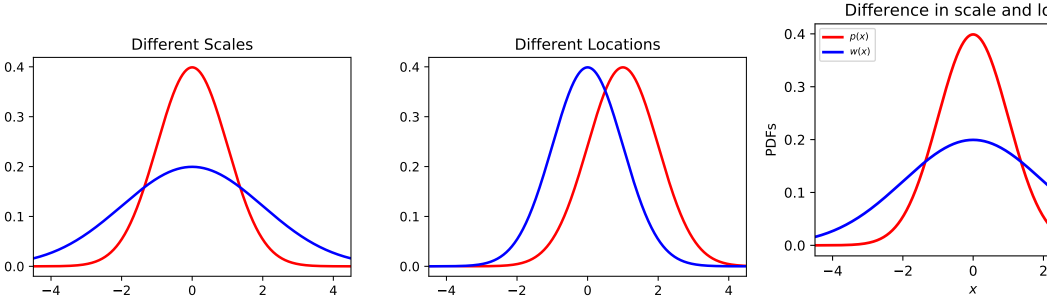
Decision Maker (DM)

DM assigns different
probabilities $w(x)$ (decision weights)
CDF $F_w(x)$

Mark Kirstein

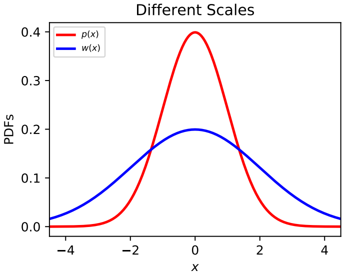
Possible model differences

Locations, Scales, Shapes



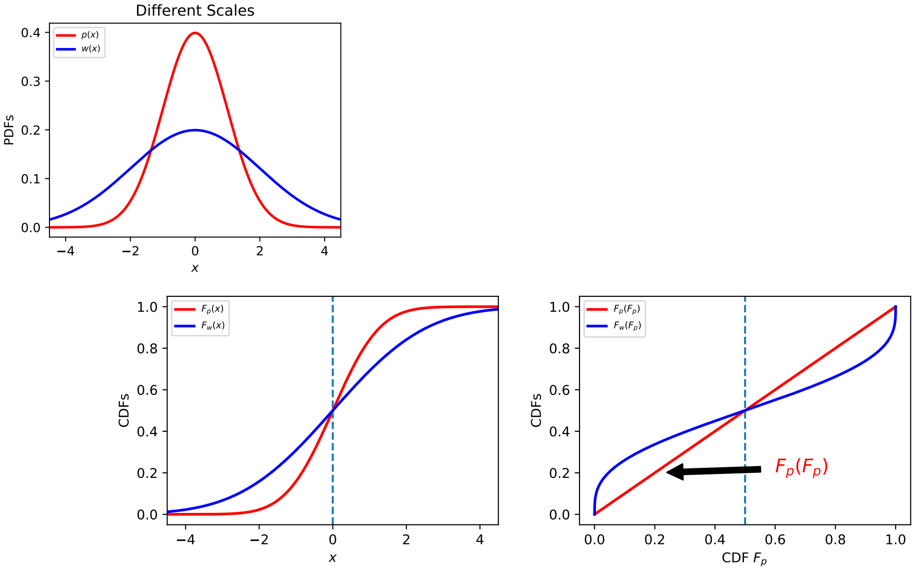
XXX produce figure with Gaussian and Student-t with heavy tails to illustrate shape difference

Thought experiment: DM assumes greater scale



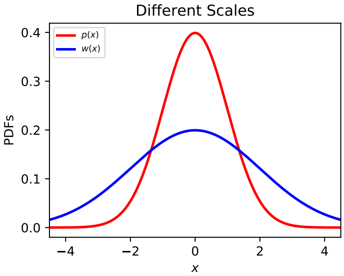
Mark Kirstein

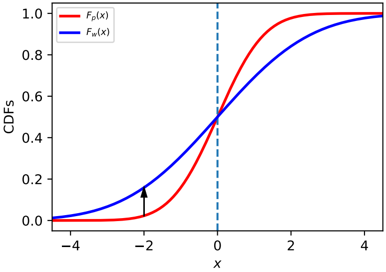
Thought experiment: DM assumes greater scale

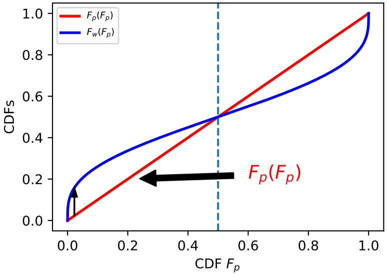


Mark Kirstein

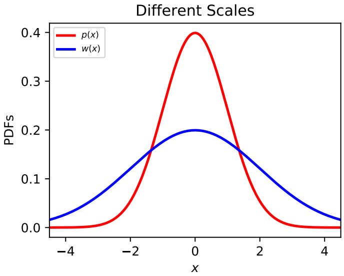
Thought experiment: DM assumes greater scale

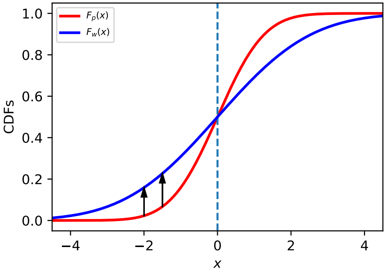


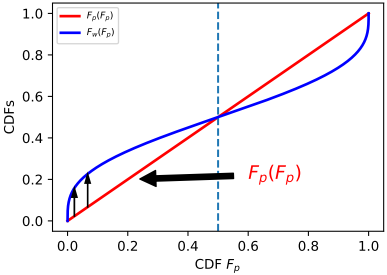




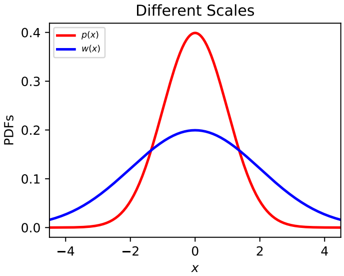
Thought experiment: DM assumes greater scale

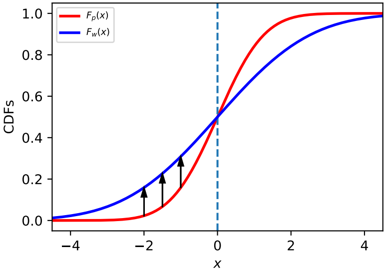


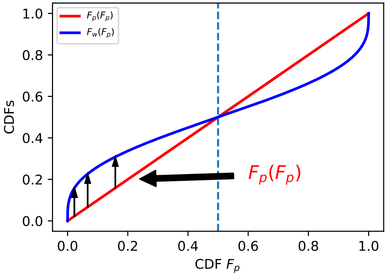




Thought experiment: DM assumes greater scale

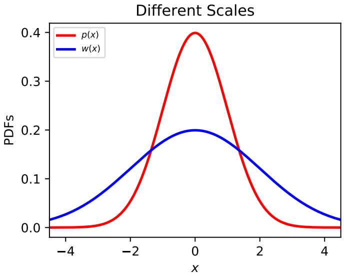


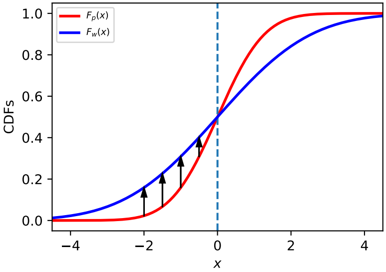


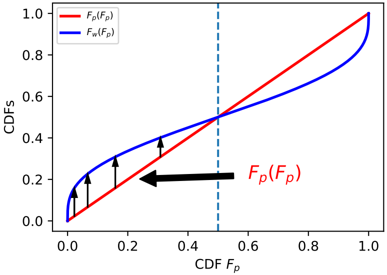


Mark Kirstein

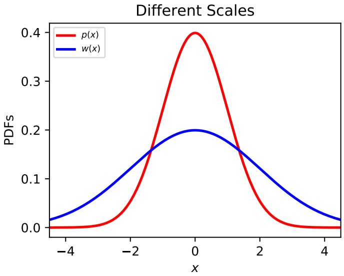
Thought experiment: DM assumes greater scale

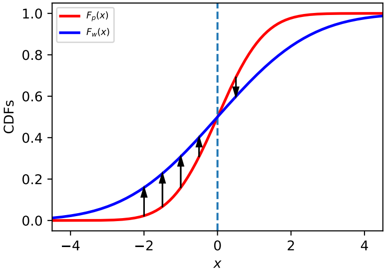


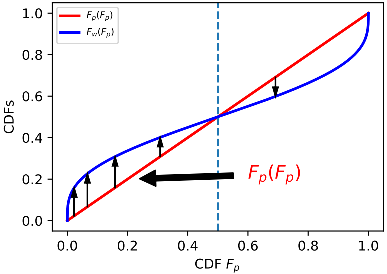




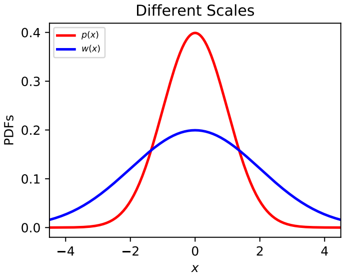
Thought experiment: DM assumes greater scale

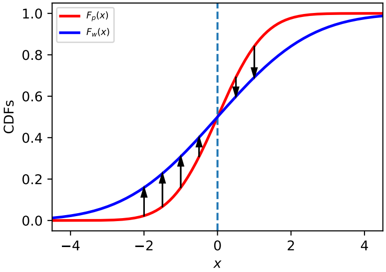


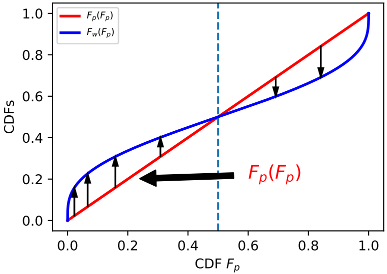




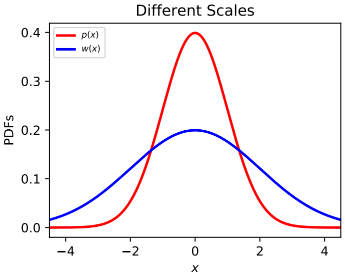
Thought experiment: DM assumes greater scale

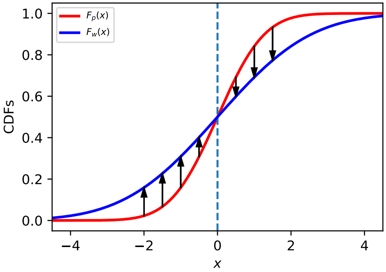


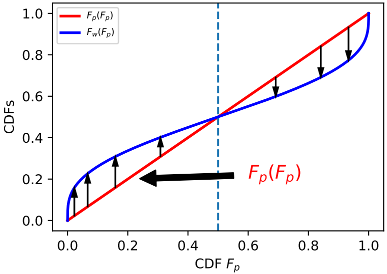




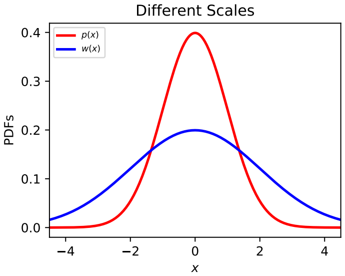
Thought experiment: DM assumes greater scale

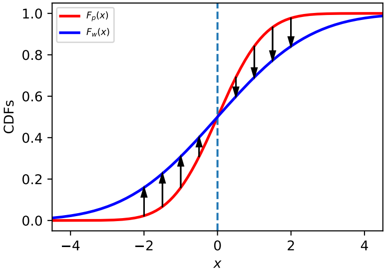


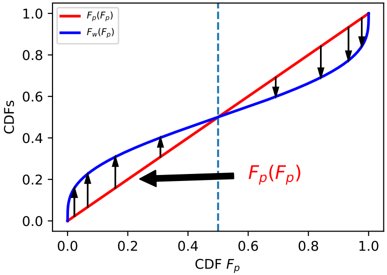




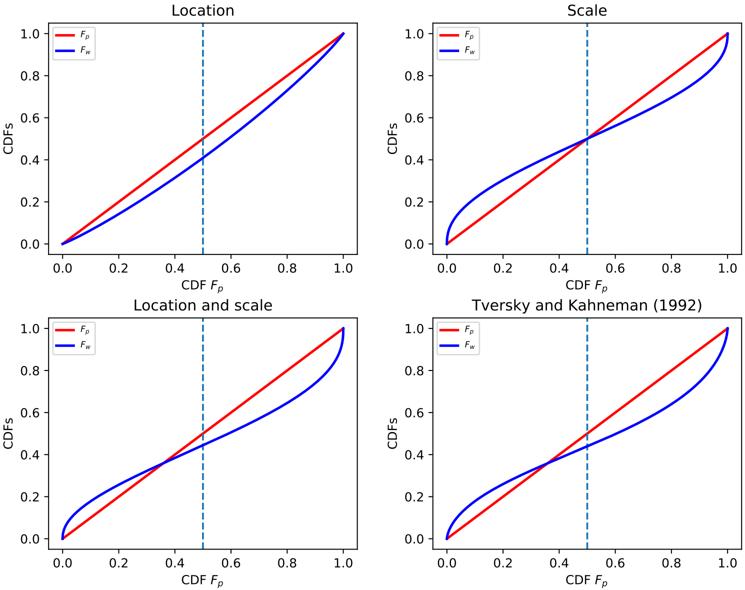
Thought experiment: DM assumes greater scale







Asymmetry from different locations



Numerically easy for any pair of distributions (models):

1. list values of DO's CDF, $F_p(x)$, at set x_i
2. list values of DM's CDF, $F_w(x)$, at same x_i
3. plot $F_w(x)$ *vs.* $F_p(x)$

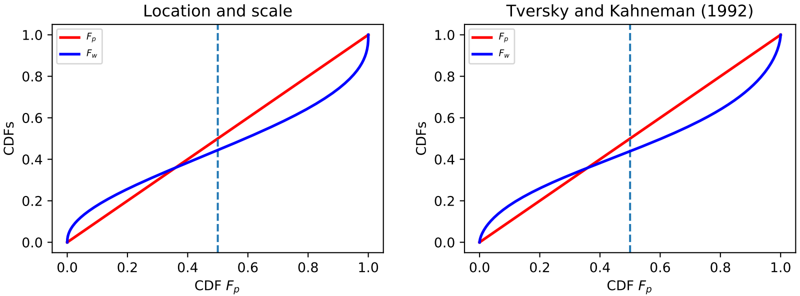
XXX illustrate with corresponding lists and figure

Interim conclusion

- DM's greater scale gives inverse-S shape (unimodal distributions)
- difference in locations gives asymmetry
- reproduces observations of probability weighting

Job done. Thank you for your attention ;)

Functional Forms



The Ergodicity Question

Typical DO concern

What happens on average to the ensemble of subjects?

$\neq$

Typical DM concern

What happens to me on average over time?

Why DM's greater scale?

- DM has no control over experiment
- experiment may be unclear to DM
- DM may not trust DO
- . . .

Experiencing probabilities

- probabilities are not observable
- probabilities encountered as
  - known frequency in ensemble of experiments (DO)
  - frequencies estimated over time (DM)

$\hookrightarrow$ **estimates have uncertainties – cautious DM accounts for these**

Estimating probabilities

Rare Event

- $p(x) = 0.0001$
- 10 000 observations
- $\sim 99.5\%$ of such time series will contain 0 or 1 events
- Naïve estimation: $\hat{p}(x) = 0$ or $\hat{p}(x) = 0.0001$

$\hookrightarrow$ either impossible or 1000% (over)estimation

Common Event

- $p(x) = 0.1$
- 10 000 observations
- $\sim 99.5\%$ of time series would contain between 50 and 150 events,
- Naïve estimation: $0.05 < \hat{p}(x) < 0.15$

$\hookrightarrow$ only $\approx 50\%$ error in $\hat{p}(x)$

$\hookrightarrow$ small $p(x)$, small count
$\hookrightarrow$ small count, big uncertainty

Relative estimation error is large for rare events

Asymptotic probability	Most likely count	Standard error in count	Standard error in probability	Relative error in probability
0.1	1000	32	0.003	3%
0.01	100	10	0.001	10%
0.001	10	3	0.0003	30%
0.0001	1	1	0.0001	100%

Table: $T = 10\,000$, assuming Poisson statistics, relative estimation errors $\sim {}^1/\sqrt{\text{count}}$

DMs don't like surprises

To avoid surprises, let's say DMs add estimation uncertainty $\varepsilon\left[p(x)\right]$ to every estimated probability, then normalize, s.t.

$w(x) = \frac{p(x)+\varepsilon\left[p(x)\right]}{\int\left(p(s)+\varepsilon\left[p(s)\right]\right)ds}$

This allows us to derive a functional form, *e.g.* for the Gaussian case ... (find in manuscript)

...visually similar to function chosen by Kahneman and Tversky.

XXXNot sure we need much more. I'd just have one figure that gives a nice inverse S, for a Gaussian, say, based on estimation error.

Conclusion

Ergodicity Economics and probability weighting

- inverse-S shape appears as neutral indicator of a difference in opinion
- reported observations consistent with DM's extra uncertainty
- may arise from DM estimating probabilities over time

$\hookrightarrow$ Probability weighting is rational cautious behaviour under uncertainty over time

- Manuscript at https://www.researchers.one/article/2020-04-14
- Interactive code at https://bit.ly/lml-pw-code
  XXX researchers.one link is better: an opportunity to advertise R1 and explain it's open for public review.
  XXX code link goes to wrong place – best to link to binder for interactive notebook.

Mark Kirstein

Thank you for your attention!

I'm looking forward to the discussion
Comments & questions are very welcome, here or to

✉ m.kirstein@lml.org.uk
@nonergodicMark

Submit an open peer review to this paper on
bit.ly/lml-pw-r1

RESEARCHERS.ONE

BACK UP

Probability Weighting as an Estimation Issue

"It is important to distinguish overweighting, which refers to a property of decision weights, from the overestimation that is commonly found in the assessment of the probability of rare events. [...] In many real-life situations, overestimation and overweighting may both operate to increase the impact of rare events." (Kahneman and Tversky 1979, p. 281)

$\hookrightarrow$ distinguish between

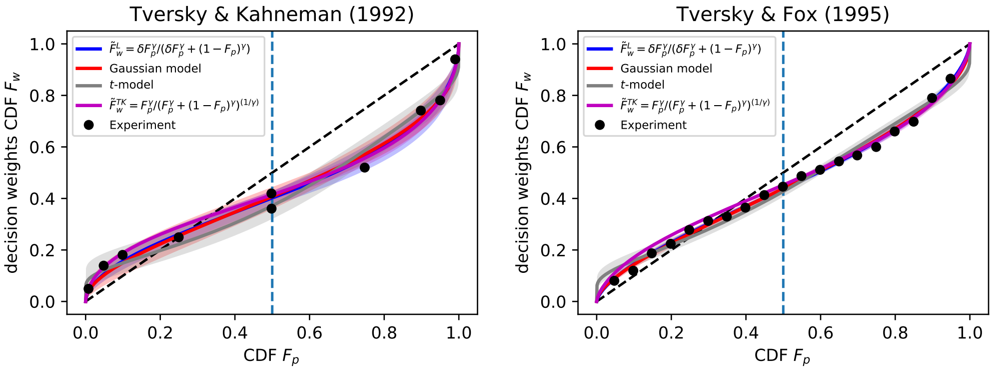
- uncertainty estimation and
- "weighting"

we analyse the former and find very good agreement with the empirical inverse-S pattern

$\hookrightarrow$ How big is the residual "probability weighting" after accounting for uncertainty estimation?

◂ Main Results

Estimation Error Explains 99% of Probability Weighting



- similar fits of Gaussian & t-distributed model

$\hookrightarrow$ How big is the residual “probability weighting” after accounting for estimation errors?

Main Results

Functional Forms

Gaussian

Tversky and Kahneman (1992, $\gamma = 0.68$)

$$\tilde{F}_w^{TK}\left(F_p;\gamma\right) = \left(F_p\right)^{\gamma} \frac{1}{\left[\left(F_p\right)^{\gamma} + \left(1 - F_p\right)^{\gamma}\right]^{1/\gamma}} \tag{1}$$

Lattimore, Baker, and Witte (1992)

$$\tilde{F}_w^{L}\left(F_p;\delta,\gamma\right) = \frac{\delta F_p^{\gamma}}{\delta F_p^{\gamma} + \left(1 - F_p\right)^{\gamma}} \tag{2}$$

We derive decision weight as a function of probability with $(\alpha\sigma)^2$ as the DM's scale

$$w(p) = p^{\frac{1}{\alpha^2}} \frac{\left(2\pi\sigma^2\right)^{\frac{1-\alpha^2}{2\alpha^2}}}{\alpha}\,, \tag{3}$$

which is a power law in p with a pre-factor to ensure normalisation

Linking Probability Weighting to Relative Uncertainties

Decision weight w is the normalised sum of the probability $p(x)$ and its uncertainty $\varepsilon\left[p(x)\right]$

$$w(x)=\frac{p(x)+\varepsilon\left[p(x)\right]}{\int_{-\infty}^{\infty}\Big(p(s)+\varepsilon\left[p(s)\right]\Big)\,ds}\;. \tag{4}$$

This can be expressed as

$$w(x)=p(x)\left(\frac{1+\frac{\varepsilon[p(x)]}{p(x)}}{\int_{-\infty}^{\infty}p(s)\left\{1+\frac{\varepsilon[p(s)]}{p(s)}\right\}ds}\right)\;, \tag{5}$$

where $\frac{\varepsilon[p(x)]}{p(x)}$ is the relative error, which is large (small) for small (large) probabilities
In the long-time limit $w(x)\rightarrow p(x)$

References I

Kahneman, Daniel and Amos Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk". *Econometrica* 47 (2), pp. 263–291. DOI:10.2307/1914185 (cit. on p. 28).

Lattimore, Pamela K., Joanna R. Baker, and A. Dryden Witte (1992). "Influence of Probability on Risky Choice: A Parametric Examination". *Journal of Economic Behavior and Organization* 17 (3), pp. 377–400. DOI:10.1016/S0167-2681(95)90015-2 (cit. on p. 30).

Tversky, Amos and Daniel Kahneman (1992). "Advances in Prospect Theory: Cumulative Representation of Uncertainty". *Journal of Risk and Uncertainty* 5 (4), pp. 297–323. DOI:10.1007/BF00122574 (cit. on pp. 3, 30).