Intertemporal Prospect Theory

Matthias Weber
University of St. Gallen

(joint work with Immanuel Lampe)

Seminar presentation at the National Bank of Ukraine

August 13, 2021

Outline

Motivation

- Prospect Theory (PT) in general describes decisions under risk better than Expected Utility Theory (EUT)
  - in atemporal settings (= outcomes materialize at one point in time)
  - PT in atemporal settings is well understood
- PT can explain several phenomena that EUT cannot explain
- Daniel Kahneman won the Nobel Prize for his work on PT

Motivation

- However, many (most?) important decisions in economics and finance involve a risk and a time dimension:
  - Saving and consumption (retirement savings)
  - Asset allocation
  - Buying a house vs. renting
  - Insurance
  - Etc.
- Still unclear how to apply prospect theory when outcomes materialize at multiple points in time (intertemporal contexts)
  - In particular, two potential application methods mentioned in the literature

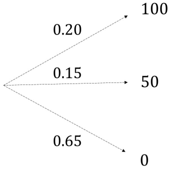
Motivation

What we do, in a nutshell:

- Conduct an experiment on a representative sample
- Subjects evaluate intertemporal lotteries
- Find out which application method describes risky choices best (out-of-sample prediction performance)
- Deliver a calibration for intertemporal PT

Prospect Theory

- A prospect/lottery consists of outcomes arising with given probabilities, $(x_1 : p_1; \ldots; x_n : p_n)$, e.g., $(100 : 0.2;\ 50 : 0.15;\ 0 : 0.65)$



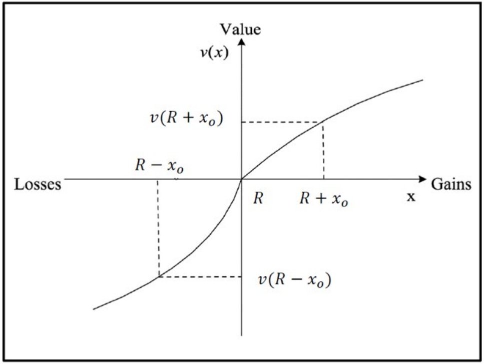
- Value of the prospect under EUT, in utility terms:

$$(x_1 : p_1; \ldots; x_n : p_n) \overset{EUT}{\to} \sum_{i=1}^{n} p_i u(x_i)$$

- Two differences in PT
  - Different utility/value function (incl. reference dependence)
  - Probability weighting

PT Value Function

- Gains and losses with respect to a reference point $R = 0$
- Kink around 0 (loss aversion)
- Often slightly concave for gains, slightly convex for losses

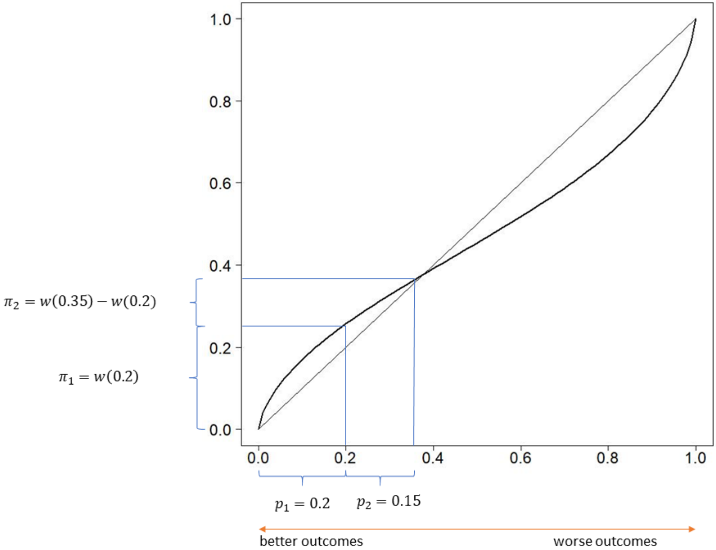


Probability Weighting

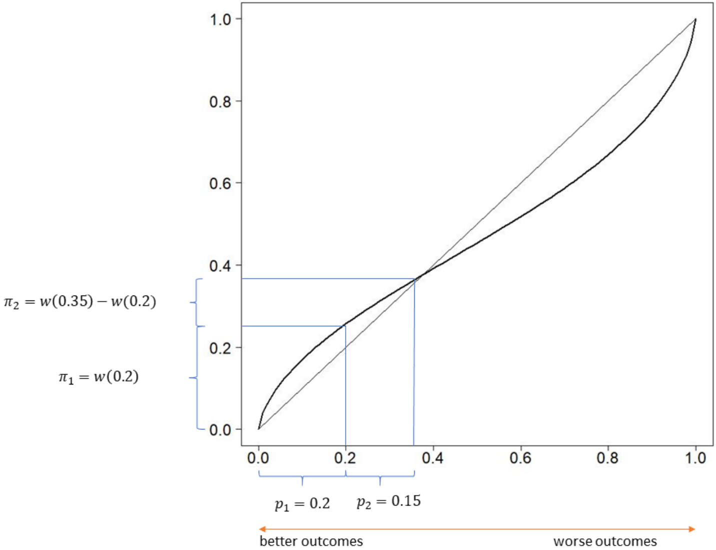
- Value of the prospect under PT, in utility/value terms:

$$(x_1 : p_1; \ldots; x_n : p_n) \overset{PT}{\rightarrow} \sum_{i=1}^{n} \pi_i v(x_i)$$

- How does this probability weighting work?
  - There is a weighting function w
  - The weighting is not $\pi_i = w(p_i)$



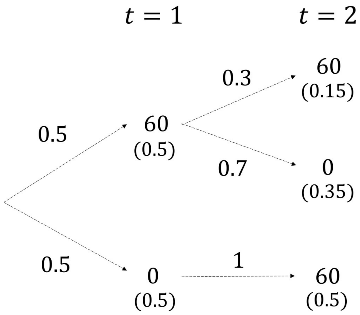
- Done separately for gains and losses



- Done separately for gains and losses

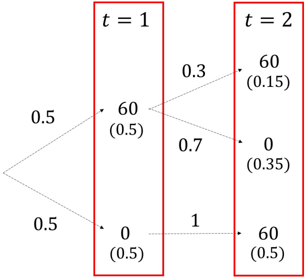
Intertemporal Prospects

- An intertemporal prospect yields (uncertain) payouts at different points in time.



▸ Additional example

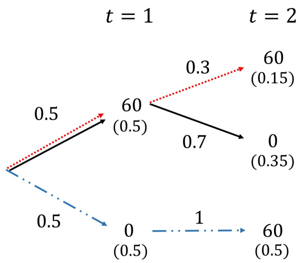
Time-separation Method



$$PT = \delta(1)PT_1 + \delta(2)PT_2$$

$$PT_1 = w(0.5)v(60) + (1 - w(0.5))v(0)$$
$$PT_2 = w(0.65)v(60) + (1 - w(0.65))v(0)$$

Present-value Method



$$PV_1 = \delta(1)v(60) + \delta(2)v(60)$$
$$PV_2 = \delta(1)v(60) + \delta(2)v(0)$$
$$PV_3 = \delta(1)v(0) + \delta(2)v(60)$$

$$PT = \pi_1 PV_1 + \pi_2 PV_2 + \pi_3 PV_3$$

Application Methods

- These two methods have been proposed in the literature
  - Time-separation method: e.g., Andreoni and Sprenger (2012), Krause et al. (2020)
  - Present-value method: e.g., Halevy (2008), Epper and Fehr-Duda (2015)
- Note: without (= with linear) probability weighting, both methods give the same results

Research Question and Literature

1. Which application method describes risky choices best?
   - Only two papers try to do this thus far, with approaches very different to ours
   - Both are laboratory experiments trying to find violations only in line with one method or the other
   - Andreoni et al. (2017) find support for the time-separation method
   - Rohde and Yu (2020) find support for the present-value method
2. What are good calibrations to apply prospect theory to intertemporal contexts?
   - Intertemporal applications usually use parametric specifications from atemporal contexts
   - Good reasons to assume that calibrations should be different in intertemporal contexts (e.g., Abdellaoui et al., 2013)

Preview of the Results

- Present-value method performs much better than time-separation method
- Calibration:
  - almost linear value functions (in loss and gain domains)
  - loss aversion parameter close to one
  - inverse-s shaped probability weighting functions (as in atemporal PT)
  - moderate discounting of all quarters (exponential) or distinction between now and future (quasi-hyperbolic; both versions predict equally well)

Experiment

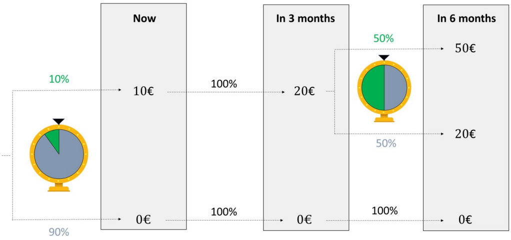
- Experiment on a sample representative for the Dutch population
- Carried out by CentERdata in September and October 2020
- Study was pre-registered (data analysis follows pre-analysis plan)

Details

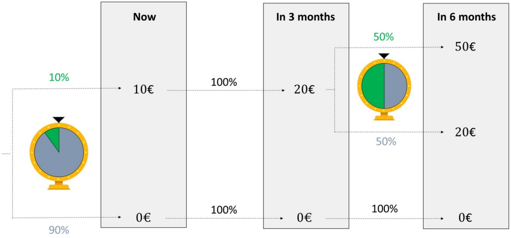
Experiment

- A total of 48 decision tasks.
- In each task subjects see a lottery ("risky option") with three uncertain payouts (today, in three months, in six months).
- We elicit the switching point from the lottery to a safe option that yields three certain and identical payouts (multi-period certainty equivalent; CE).
- For 75% of subjects hypothetical choices (T1), for the rest (T2) part of the choices incentivized
  - T1: 15 EUR for participation
  - T2: 15 EUR for participation, on average 84 EUR in addition
  - Incentivized and hypothetical choices do not differ [no strategic interaction, social image, self image]

Decision Screen



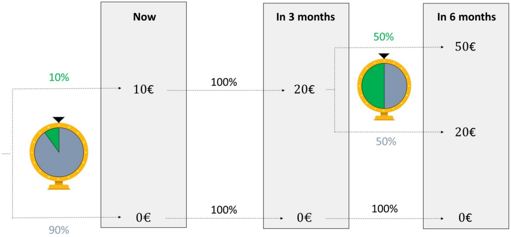
Decision Screen



I choose the **risky option** if the payout amount of the safe option that I receive three times (now, in 3 months and in 6 months) lies between:	I choose the **safe option** if the payout amount of the safe option that I receive three times (now, in 3 months and in 6 months) lies between:
0€ and **7€**	**8€** and 27€

0€ — 27€

Decision Screen



I choose the **risky option** if the payout amount of the safe option that I receive three times (now, in 3 months and in 6 months) lies between:	I choose the **safe option** if the payout amount of the safe option that I receive three times (now, in 3 months and in 6 months) lies between:
0€ and **7€**	**8€** and 27€

0€ 27€

↑ CE: 7-8

Lotteries

- 6 sets of 8 lotteries each

	Gains	Losses	Mixed
Small Stakes	Set 1	Set 2	Set 3
Large Stakes	Set 4	Set 5	Set 6

48 Lotteries

36 Calibration Lotteries
For any combination of the value, weighting and discount function TS = PV
$\Rightarrow$ Estimation leads to same parameters for both methods

12 Test Lotteries (2 from each set)
1-6: TS $>$ Linear Probability Weighting $>$ PV
7-12: TS $<$ Linear Probability Weighting $<$ PV

Parametric Specification

- For each application method, we use 12 combinations of value, probability weighting, and time-discount functions

		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
Value	Power	x	x	x	x	x	x						
	Exponential							x	x	x	x	x	x
Weighting	T+K (1992)	x	x					x	x				
	Prelec (1998)			x	x					x	x		
	G+E (1987)					x	x					x	x
Discount	Exponential	x		x		x		x		x		x	
	Quasi hyp.		x		x		x		x		x		x

Details

Maximum Likelihood Estimation (Calibration Set)

Assumptions (standard) for the log-likelihood function:

- Stated certainty equivalents are affected by noise $\epsilon_{i,j}$, with $\epsilon_{i,j} \sim \mathcal{N}(0, \sigma^2_{i,j})$.
- Standard deviation $\sigma_{i,j} = \epsilon_i w_j$
  - subject-specific ϵ_i
  - proportional to lotteries payout range w_j

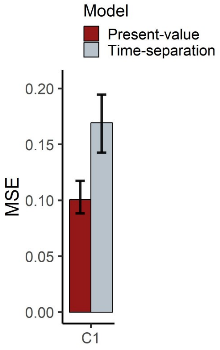
▸ Details

Measurement of Prediction Performance (Test Set)

(Weighted) MSE of participant i:

- $MSE_i = \frac{1}{12}\sum_{j=1}^{12}\left(\frac{1}{w_j}\left(CE_{i,j} - \widehat{ce_j}\right)\right)^2$
  - $\widehat{ce_j}$: predicted certainty equivalent for lottery j by given model
  - $CE_{i,j}$: certainty equivalent reported by player i for lottery j
  - w_j: payout range of lottery j
- **Main outcome variable:** $MSE = \frac{1}{n}\sum_{i=1}^{n} MSE_i$.
- Standard errors are bootstrapped (at the participant level)
- Tests are paired bootstrap tests

Main Result

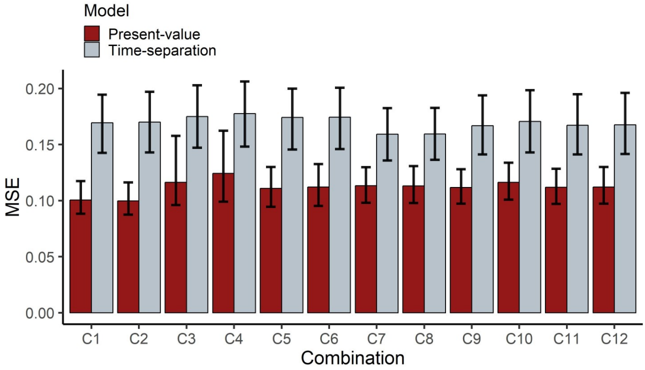


- For any combination (C1,C2,...,C12), the present-value method predicts decisions better than the time-separation method [even holds for all lotteries!]

Illustration I

Illustration II

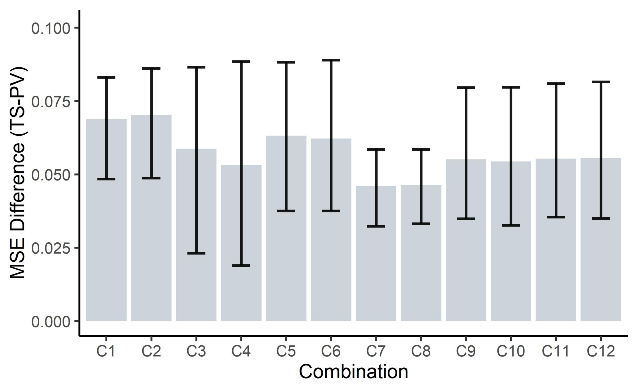
Main Result



- For any combination (C1,C2,...,C12), the present-value method predicts decisions better than the time-separation method [even holds for all lotteries!]

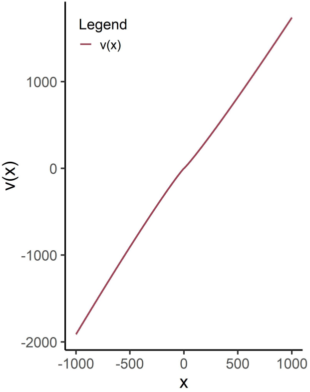
▸ Illustration I ▸ Illustration II

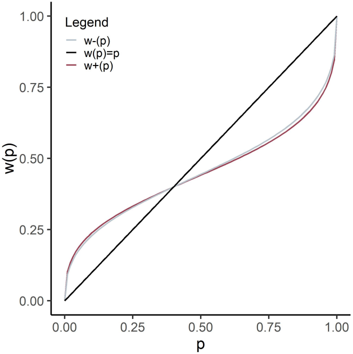
Statistical Tests



Example Calibration

- C6: i) power value function, ii) G+E (1987) probability weighting functions, iii) quasi-hyperbolic discounting)

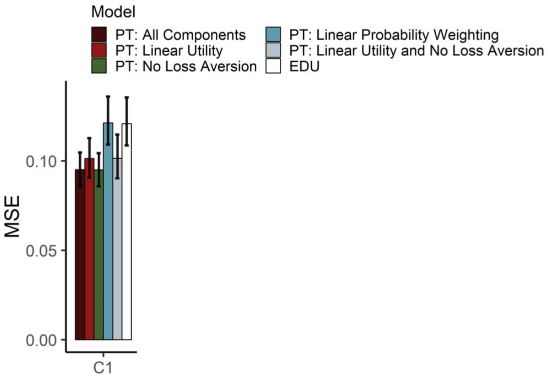




- quasi-hyperbolic disc.: $\delta(t) = k \exp(-rt)$, $k = 0.884$, $r = 0.001$

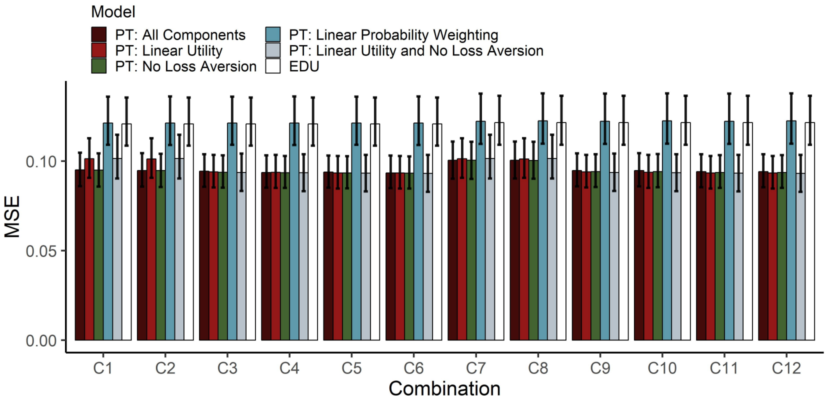
▸ MPV

PT Components and Comparison to EDU



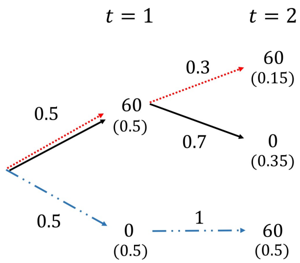
- It's all about probability weighting

PT Components and Comparison to EDU



- It's all about probability weighting

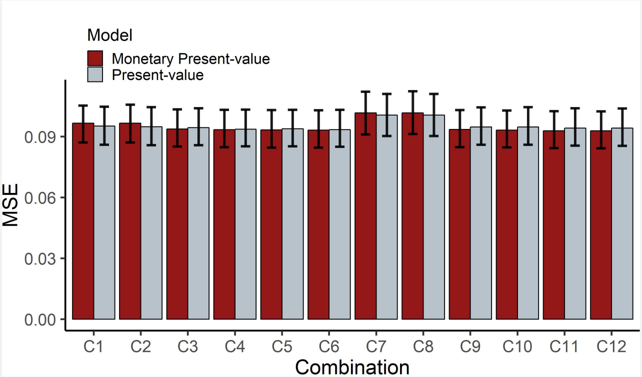
Monetary Present-value Method



$$MPV_1 = \delta(1)60 + \delta(2)60$$
$$MPV_2 = \delta(1)60 + \delta(2)0$$
$$MPV_3 = \delta(1)0 + \delta(2)60$$

$$PT = \pi_1 v(MPV_1) + \pi_2 v(MPV_2) + \pi_3 v(MPV_3)$$

Monetary Present-value Method

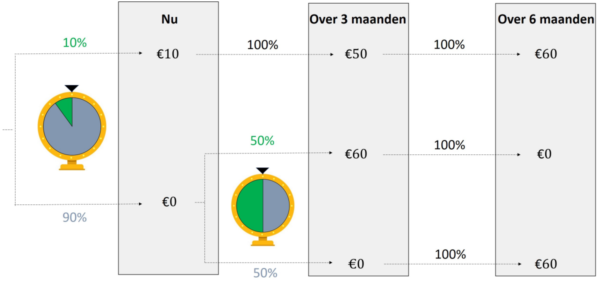


Conclusion

- Present-value method performs much better than time-separation method
- Monetary present-value method as good as present-value method
- Calibration:
  - almost linear value functions (in loss and gain domains)
  - loss aversion parameter close to one
  - inverse-s shaped probability weighting functions (as in atemporal PT)
  - moderate discounting of all quarters (exponential) or distinction between now and future (quasi-hyperbolic; both versions predict equally well)

Thank you for your attention!

Example Lottery (Same Evaluation)



Return

Number of Subjects

- 378 subjects completed the experiment
- Data exclusion is strict (ensures that results are not driven by carelessness or misunderstandings) and follows pre-registration:
  - subjects stated comprehension difficulties or low attention in at least one post-experimental question
  - short median decision times
- Left with 100 subjects
- Main result (comparison of the methods) identical when conducted with all subjects
- Demographic variables between excluded and general subject pool very similar

Return

Function Specifications

Specification	Parameters
Value functions	
Power utility: $v(x) = \mathbb{1}_{x \geq 0} x^{\alpha_1} - \mathbb{1}_{x<0} \lambda_1 (-x)^{\alpha_1}$	α_1, λ_1
Exponential utility: $v(x) = \mathbb{1}_{x \geq 0} \frac{1-\exp(-\alpha_2 x)}{\alpha_2} - \mathbb{1}_{x<0} \lambda_2 \frac{1-\exp(\beta x)}{\beta}$	α_2 , β, λ_2
Probability weighting functions (gains and losses)	
Tversy and Kahnemann (1992): $w(p) = \frac{p^{\gamma_1}}{(p^{\gamma_1} + (1-p)^{\gamma_1})^{1/\gamma_1}}$	γ_1^+, γ_1^-
Prelec (1998): $w(p) = \exp(-\eta(-ln(p))^{\gamma_2})$	$\eta_1^+, \eta_1^-, \gamma_2^+, \gamma_2^-$
Goldstein Einhorn (1987): $w(p) = \frac{\eta p^{\gamma_3}}{\eta p^{\gamma_3} + (1-p)^{\gamma_3}}$	$\eta_2^+, \eta_2^-, \gamma_3^+, \gamma_3^-$
Time-discount functions	
Exponential discounting: $\delta(t) = \exp(-r_1 t)$	r_1
Quasi-Hyperbolic discounting: $\delta(t) = k \exp(-r_2 t)$	k, r_2

Return

Maximum Likelihood Procedure (Details)

- ce_j = the certainty equivalent of lottery j resulting from an evaluation under one model specification.
- $CE_{i,j}$ = certainty equivalent player j reports for lottery i.
- Assumptions (as Hey et al., 2009; Bruhin et al., 2010)
  - Noise: $CE_{i,j} = ce_j + \epsilon_{i,j}$, with $\sim \mathcal{N}(0, \sigma_{i,j}^2)$.
  - SD subject specific and payout range dependent $\sigma_{i,j} = \epsilon_i w_j$.
- contribution of participant i:

$$f(\theta, \epsilon_i | CE_i) = \prod_{j=1}^{36} \frac{1}{\sigma_{i,j}} \phi \left(\frac{CE_{i,j} - ce_j(\theta)}{\sigma_{i,j}} \right)$$

- All n participants

$$log\ L(\theta, \epsilon | CE) = \sum_{i=1}^{n} log\ f(\theta, \epsilon_i | CE_i) = \sum_{i=1}^{n} \sum_{j=1}^{36} log\ \left[\frac{1}{\sigma_{i,j}} \phi \left(\frac{CE_{i,j} - ce_j(\theta)}{\sigma_{i,j}} \right) \right]$$

Return

Mean Absolute Prediction Error (By Lottery)

	Low-stake Lotteries						High-stake Lotteries					
	L7	L8	L15	L16	L23	L24	L31	L32	L39	L40	L47	L48
Payout Range	23	20	23	20	30	50	467	400	467	400	600	1000
Mean Error TS	6	8.1	7.3	6.5	7.9	16.7	144.1	148.6	146	124.9	180	319.9
Mean Error PV	5.2	5.6	6.1	5.6	7.2	14.6	108.5	118	114.4	115	150.7	276.8

Notes: The mean absolute prediction error of Lottery j is calculated as mean($|CE_{i,j} - \widehat{ce}_j|$), with $CE_{i,j}$ denoting the certainty equivalent subject j reported for Lottery j and $\widehat{ce}_j$ denoting the predicted certainty equivalent resulting from the parameters estimated on the calibration set.

Return

Participant Types

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
Time-separation Types	10	5	1	1	1	1	5	5	3	1	2	1
Present-value Types	81	84	84	81	89	88	87	87	86	89	88	88
Unclassified	9	11	15	18	10	11	8	8	11	10	10	11

Notes: Participant i is classified as time-separation type if $MSE_i^{PV} - MSE_i^{TS} > SE(\triangle MSE)$ or as present-value type if $MSE_i^{TS} - MSE_i^{PV} > SE(\triangle MSE)$.

Return

References I

Abdellaoui, M., Bleichrodt, H., l'Haridon, O., and Paraschiv, C. (2013). Is there one unifying concept of utility? an experimental comparison of utility under risk and utility over time. *Management Science*, 59(9):2153–2169.

Andreoni, J., Feldman, P., and Sprenger, C. (2017). A stream of prospects or a prospect of streams: On the evaluation of intertemporal risks. Technical report, National Bureau of Economic Research.

Andreoni, J. and Sprenger, C. (2012). Risk preferences are not time preferences. *American Economic Review*, 102(7):3357–76.

Bruhin, A., Fehr-Duda, H., and Epper, T. (2010). Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica*, 78(4):1375–1412.

References II

Epper, T. and Fehr-Duda, H. (2015). Risk preferences are not time preferences: balancing on a budget line: comment. *American Economic Review*, 105(7):2261–71.

Halevy, Y. (2008). Strotz meets allais: Diminishing impatience and the certainty effect. *American Economic Review*, 98(3):1145–62.

Hey, J. D., Morone, A., and Schmidt, U. (2009). Noise and bias in eliciting preferences. *Journal of Risk and Uncertainty*, 39(3):213–235.

Krause, J., Ring, P., and Schmidt, U. (2020). Hyperopic loss aversion. SSRN Working Paper, No 3582050.

Rohde, K. I. M. and Yu, X. (2020). Aggregating over risk and time in risky intertemporal choice: Which order? Working Paper, Erasmus University Rotterdam.