Intertemporal Prospect Theory*

Immanuel Lampe† Matthias Weber‡

February 23, 2022

Abstract

Prospect Theory is well understood in contexts without a time dimension. In intertemporal contexts, however, it is unclear how prospect theory should be applied. In particular, it is unclear whether probabilities should be weighted within time periods or whether probabilities of present values should be weighted. Furthermore, it is unclear what parametric specifications of probability-weighting and value functions should be used. We find in a pre-registered experiment on a representative sample that the version weighting probabilities of present values predicts decisions best. Estimated probability weighting functions are inverse-S shaped, and value functions are almost linear.

Keywords: Prospect theory; intertemporal choice; time-first vs risk-first evaluation; decisions under risk.
JEL-codes: D81; D90; C90; G40; D15.

*We thank Renée Adams, Francesco Audrino, Peter Bossaerts, Martin Brown, Patrick Chuard-Keller, Thomas Epper, Özgür Gürerk, Frank Heinemann, Lorenz Kueng, Boris van Leeuwen, Jona Linde, Paul Smeets, David Solomon, Peter Wakker, and Michael Weber for comments and suggestions, as well as conference participants at ASFEE Conference, Behavioral Macroeconomics Workshop, Workshop on Theoretical and Experimental Macroeconomics, Experimental Finance Conference, SSES Congress, ZEW Public Finance Conference, WEAI Meetings, D-TEA Conference, and seminar participants at University of Technology Sydney, National Bank of Ukraine, Heidelberg University, the Baltic Economic Association, and the University of St. Gallen. IRB approval was obtained from the University of St. Gallen. We gratefully acknowledge funding from University St. Gallen GFF project funding. We are also grateful to the Society for Experimental Finance, which awarded us the Vernon L. Smith Young Talent Award for this paper.

†School of Economics and Political Science, University of St. Gallen. Email: immanuel.lampe@unisg.ch.
‡School of Finance, University of St. Gallen. Email: matthias.weber@unisg.ch.

1 Introduction

Analyzing decisions under risk is an important topic in research in economics, finance, and business. The most prominent contender of expected utility theory is (cumulative) prospect theory (Tversky and Kahneman, 1992). Prospect theory entails two fundamental breakaways from the classical expected utility model. First, instead of defining preferences over total wealth, preferences are defined over changes with respect to a reference point, and it is assumed that negative changes (losses) are treated differently than positive changes (gains). Second, outcomes are not weighted by objective probabilities but rather by transformed probabilities (specifically, unlikely large gains and unlikely large losses are overweighted). Considering financial decisions, prospect theory can, for example, explain the equity premium puzzle (Benartzi and Thaler, 1995; Barberis and Huang, 2006), the stock market participation puzzle (Dimmock and Kouwenberg, 2010), and the disposition effect (Odean, 1998; Weber and Camerer, 1998; Barberis and Xiong, 2009; Meng and Weng, 2017). In other fields of economics, prospect theory can explain the downward-sloping labor supply curve (Goette et al., 2004), the avoidance of probabilistic insurance (Wakker et al., 1997) and index insurance (Lampe and Würtenberger, 2020), and the choice of the deductible in insurance contracts (Sydnor, 2010). DellaVigna (2009) and Barberis (2013) provide overviews of many more phenomena that are inconsistent with expected utility theory but can be explained with prospect theory.

Most applications of prospect theory assume an atemporal setting, that is, a setting where the gains and losses materialize at one point in time (e.g., Benartzi and Thaler, 1995; Berkelaar et al., 2004; Gomes, 2005; and Barberis and Huang, 2006). In these settings, prospect theory has been thoroughly analyzed and is well understood. Many natural settings, however, involve gains and losses at different points in time. Recent studies integrate elements of prospect theory in such intertemporal settings (see for example Barberis and Huang, 2009; Barberis and Xiong, 2009; De Giorgi and Legg, 2012; van Bilsen et al., 2020; Hlouskova et al., 2017; Meng and Weng, 2017; Deng and Pirvu, 2019). However, in intertemporal applications, usually only some aspects of prospect theory are incorporated into a setting that is otherwise one of discounted expected utility maximization. A majority of these studies include loss aversion with respect to a reference point in their models as in prospect theory but neglect probability weighting (e.g., Kyle et al., 2006; Barberis and Huang, 2009;

Barberis and Xiong, 2009; Henderson, 2012; Easley and Yang, 2015; Hlouskova et al., 2017; Meng and Weng, 2017).

To date, there is no generally accepted version of prospect theory in intertemporal settings. There are not only minor disagreements about the calibration of the parameters, but there is not even agreement on how probability weighting should be applied. Considering that prospect theory is widely viewed as the most accurate available description of how people make risky decisions, this is surprising. This research project addresses this gap in the literature and aims to provide guidance on how prospect theory should best be applied to intertemporal settings.

We refer to the first possible method that can be used as the time-separation method. With this method, outcomes are ranked within each period and decision weights are calculated (with probability weighting functions similar to an atemporal setting). Given these decision weights, each period's PT value can be calculated and the overall PT value is then the time-discounted sum of these values. Note that the time discounting need not be exponential, any other form of discounting such as hyperbolic or quasi-hyperbolic discounting may also be used. This first method of applying prospect theory is, for instance, assumed in the analyses by Andreoni and Sprenger (2012); De Giorgi and Legg (2012); Guo and He (2017); Krause et al. (2020) and van Bilsen and Laeven (2020).

The second method, which we refer to as the present-value method, first calculates the decision maker's present value of each possible stream of outcomes, in utility terms (with a given form of time discounting). These present values are then ranked and their decision weights are calculated (decision weights are again calculated analogously to atemporal prospect theory). The overall PT value is then calculated as the atemporal PT value of the present values. This method (or a version of it with rank-dependent utility, not distinguishing between gains and losses but including probability weighting) is used in the studies by Chew and Epstein (1990), Bleichrodt and Eeckhoudt (2006), Halevy (2008), Drouhin (2015), Epper and Fehr-Duda (2015), Blavatskyy (2016), Andreoni et al. (2017), and Andersen et al. (2018).

In general, the two methods yield different results. For studies that neglect probability weighting, both methods predict the same evaluation of prospects, however. The first step towards providing a unifying standard of how to apply prospect theory in intertemporal settings is to choose between these two methods.

A second open issue when moving from atemporal to intertemporal prospect theory is the specification of the value and probability weighting functions (for a given method of applying prospect theory). Many studies that integrate elements of prospect theory into intertemporal models use functional forms and parameter calibrations obtained in atemporal settings.[1] It is a priori not clear that the specifications of the utility and probability weighting functions obtained in an atemporal setting are well suited to analyze intertemporal choices. The studies by Abdellaoui, l'Haridon and Paraschiv (2013) and Abdellaoui, Bleichrodt, l'Haridon and Paraschiv (2013), for instance, suggest that the utility function in intertemporal contexts is close to linear for gains and losses and that the loss aversion parameter is smaller than usually found in atemporal contexts (elicited in intertemporal settings without risk).

In this study, we elicit certainty equivalents of intertemporal lotteries in a representative population. We estimate parameter calibrations for a variety of different value, probability-weighting and time-discount functions and can then determine which of the two methods describes risky choices best. We also provide a benchmark of functions and parameter values for intertemporal applications. For a full specification of intertemporal prospect theory, we estimate value and probability-weighting functions jointly with the method of time discounting.

Very closely related literature is scarce. Epper and Fehr-Duda (2015) nicely illustrate the co-existence of the two methods of intertemporal prospect theory and show that the present-value method can explain a decision pattern that cannot be explained with the time-separation method (documented in Andreoni and Sprenger, 2012). However, their study does not attempt to establish which method in general predicts choices better (neither to estimate utility, probability-weighting, or time-discount functions).

The two existing studies that attempt to analyze which of the two methods of prospect theory explains human choices better provide conflicting evidence. Andreoni et al. (2017) find support for the time-separation method, while Rohde and Yu (2020) find support for the present-value method. Both studies rely on experimental designs that make the observations

[1] Many studies use the specifications of Tversky and Kahneman (1992), who provide a natural benchmark, despite the fact that specifications obtained in larger representative samples are available, (e.g., Booij et al., 2010). Note, however, that the parameters obtained are generally similar to those provided by Tversky and Kahneman (1992).

of certain outcomes more or less likely depending on whether experimental subjects use the time-separation method or the present-value method. Andreoni et al. (2017) investigate whether adding an independent common future risk affects the choice between two lotteries. Assuming functional forms as usually found in atemporal settings, the choice can change if participants use the present-value method to evaluate the lotteries but not if they use the time-separation method (because the PT value of the newly added time period is identical for both lotteries). Andreoni et al. (2017) interpret their findings as support for the time-separation method. The support for this interpretation is not very strong, however, because choices do change in roughly 50% and the difference in explanatory power between the methods is small.[2] Rohde and Yu (2020) investigate intertemporal correlation aversion in a laboratory experiment (in a study conducted at the same time as ours and independently of it). Intertemporal correlation aversion can arise if participants use the present-value method to evaluate the lotteries but not if they use the time-separation method. Rohde and Yu (2020) find that the majority of participants is averse to intertemporal correlation, their results thus lend support to the present-value method.

Our paper differs decisively from Andreoni et al. (2017) and Rohde and Yu (2020). Rather than presenting participants combinations of simple lotteries, where certain choices by participants are in line with one method or the other, we attempt to establish which application method has higher predictive power when participants evaluate many lotteries (in a setting that does not favor any of the two methods a priori). We also believe that thoroughly estimated value, probability-weighting, and time-discount functions in an intertemporal setting are of high value. Scholars and practitioners can only use the different methods in applications if such a parametrization is available (estimating the parameters anew also seems necessary to make sure that the comparison between the methods is not driven by potentially unfitting atemporal calibrations). To be able to obtain parameters that extend beyond a typical student sample (as used, for instance, in Andreoni et al., 2017, and Rohde and Yu, 2020), we conduct the study on a representative sample.[3]

[2]Furthermore, Andreoni et al. (2017) present the relevant choice lists (that only differ by the same added outcome of a 90% or 10% chance of 19 or 9 USD in four weeks) one after another to the same participants. This nudges people to make a consistent choice, which favors the time-separation method in their setup.

[3]There are additional differences, in particular when compared to Andreoni et al. (2017). Participants' choices in Andreoni et al. (2017), for instance, only influence the first periods' payouts; that is, what happens in the future is independent of participants' choices. This runs counter to many natural settings in which

To preview our results, the present-value method predicts observed choices significantly better than the time-separation method for all considered parametric specifications. The prediction performance is also much better than the prediction performance of expected discounted utility models. The estimated value functions are almost linear for gains and losses and have a loss aversion coefficient close to one. The estimated probability-weighting functions are inverse-S shaped and very similar to typical atemporal calibrations.[4]

While our main focus lies on the comparison of the time-separation method and the present-value method, we also compare these methods to a third application method, which is similar to the present-value method. This method also calculates the present values of outcome streams first, but it calculates these present values in monetary terms before entering them in a value/utility function (in contrast to the regular present value method, which calculates present values directly in utility terms). The third application method based on present values in monetary terms performs as well as the regular present-value method.

As a minor additional result, there is no noticeable difference between incentivized and hypothetical choices in our experiment. This is not surprising, given that the valuation of lotteries does not contain strategic elements or social- or self-image considerations. This is in line with the literature (e.g., Abdellaoui et al., 2011; Brañas Garza et al., 2020; Hackethal et al., 2020).

This paper is organized as follows. Section 2 introduces the two methods of applying prospect theory to intertemporal settings. Section 3 contains the experimental design and briefly discusses the pre-registration and the experimental procedures. Section 4 discusses parametric specifications, the estimation method, and the measure of prediction performance. Section 5 shows the results. Section 6 presents parameter calibrations that we recommend for intertemporal applications and concludes.

current decisions also influence the distribution of future risks (e.g., investments, education, health), and, in particular, it means that there is no possibility for 'diversification across time', which Epper and Fehr-Duda (2015) consider an important reason why the present-value method may explain peoples' choices better. In our experiment, future risks are in general not independent.

[4]In our experiment, we implement salient reference points of zero in the different periods (in later applications, it will of course be possible to work with different reference points). We show that our results are robust to allowing for different reference points. The model feature that increases prediction performance considerably over expected utility theory is probability weighting, not utility curvature or loss aversion (with or without variations in the reference points).

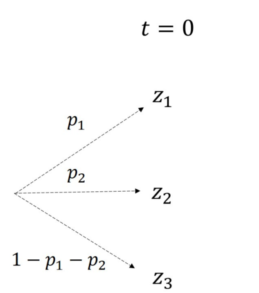
2 Modeling Options for Intertemporal Prospects

This section introduces intertemporal prospects and the two (main) possibilities of modeling intertemporal decisions with prospect theory.

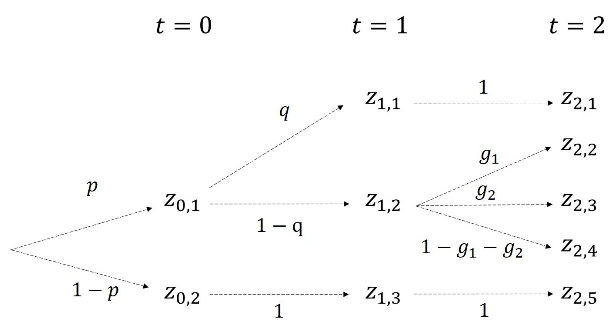
2.1 Intertemporal Prospects

An atemporal prospect yields one single payout at one time period. Figure 1 a) displays an explanatory atemporal prospect. The three potential outcomes (arriving at $t = 0$, without time discounting) are labeled z_1, z_2 and z_3. Outcome z_1, for instance, arrives with probability p_1.

An intertemporal prospect may yield payouts at different time periods. Figure 1 b) displays an explanatory intertemporal prospect. The potential outcomes of period t are labeled by $z_{t,1}, z_{t,2}...$. A period's outcomes are not necessarily distinct (that is, $z_{1,2} = z_{1,3}$ in Figure 1, for instance, is possible) and the labels may not reflect the outcomes' rank-order. In the first period ($t = 0$), either outcome $z_{0,1}$ arrives (with probability p) or outcome $z_{0,2}$ (with probability $1-p$). Outcomes at $t = 1$ and $t = 2$ depend in general on the outcomes of previous periods. The probability that one specific outcome arrives is, thus, equal to the product of its path probabilities. The probability that the outcome at $t = 1$ equals $z_{1,1}$ is, for instance, given by $p \cdot q$.



(a) Atemporal prospect



(b) Intertemporal prospect

Figure 1: Examples of an atemporal and an intertemporal prospect

2.2 Time-separation Method

To evaluate a prospect with outcomes in T periods, the time-separation method first calculates each period's PT value. These values are then time-discounted and added up, with a given time-discount function δ (generally with $\delta(0) = 1$). We denote the PT value of period t by V_t, $t = 0, ..., T-1$. Applying the time-separation method to the example prospect of Figure 1 implies that its PT value is the sum of its three time-discounted PT values, as illustrated in Figure 2.

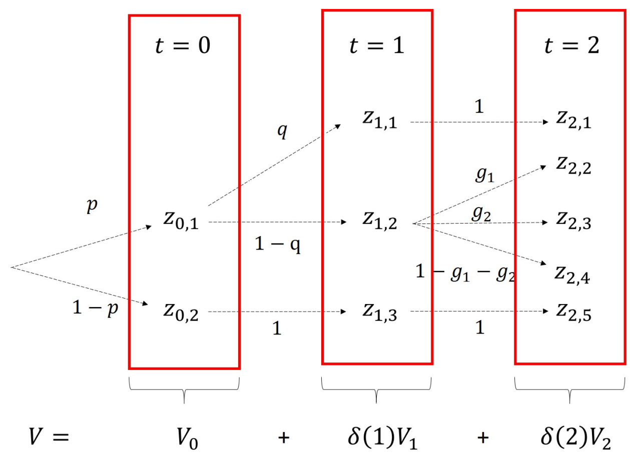


Figure 2: Illustration of the time-separation method

To calculate each period's PT value, prospect theory is applied as in an atemporal setting. Periods' outcomes are ranked and their objective probabilities are transformed into subjective decision weights. Denote by $x_{t,1}..., x_{t,n_t}$ the ordered and distinct outcomes of period t and by $p_{t,1}, ..., p_{t,n_t}$ their objective probabilities. Each period's outcome set may consist of gains and losses, so that $x_{t,1} > x_{t,2} > ... > x_{t,k_t} > 0 > x_{t,k_t+1} > ... > x_{t,n_t}$. Note that the objective probabilities are compound probabilities (for example, if $z_{1,1} > z_{1,2}$ and $z_{1,1} > z_{1,3}$ in Figures 1 b) and 2, then $x_{1,1} = z_{1,1}$ and $p_{1,1} = p \cdot q$). The PT value of period t can then simply be calculated with the formula

$$V_t = \sum_{i=1}^{k_t} \pi_{t,i}^{+} v(x_{t,i}) + \sum_{l=k_t+1}^{n_t} \pi_{t,l}^{-} v(x_{t,l}),$$

with v denoting the decision maker's value function (i.e., the utility function), which is strictly increasing. We assume that the reference point is at 0 and $v(0) = 0$. Further, the decision weights $\pi^+_{t,i}$ and $\pi^-_{t,i}$ are defined by

$$\begin{aligned}
&\pi^+_{t,1} = w^+(p_{t,1}) \ , \ \pi^-_{t,n_t} = w^-(p_{t,n_t}),\\
&\pi^+_{t,i} = w^+(p_{t,1} + \ldots + p_{t,i}) - w^+(p_{t,1} + \ldots + p_{t,i-1}), \ \text{for } 1 < i \le k_t,\\
&\pi^-_{t,l} = w^-(p_{t,l} + \ldots + p_{t,n_t}) - w^-(p_{t,l+1} + \ldots p_{t,n_t}), \ \text{for } k_t < l < n_t.
\end{aligned}$$

Here w^+ is the probability weighting function for gains and w^- that for losses, satisfying $w^+(0) = w^-(0) = 0$, $w^+(1) = w^-(1) = 1$, both monotonically increasing.

The overall value of the prospect, which we label V, is then the sum of all time-discounted PT values, thus $V = \sum_{t=0}^{T-1} \delta(t) V_t$.

2.3 Present-value Method

The present-value method evaluates an intertemporal prospect by calculating the present values of the possible streams of outcomes in utility terms, and then weighting the probabilities of these present values. A stream of outcomes, is a sequence of outcomes that arrives with positive probability. We denote the different possible streams of outcomes by $o_1, o_2, \ldots$. An outcome stream o_j thus consists of T outcomes, one in each time period, $o_j = (o_{j,0}, \ldots, o_{j,T-1})$, with $o_{j,t}$ denoting the outcome in period t. Figure 3 illustrates the outcome streams of the example given in Figure 1 (for o_2, for instance, $o_{2,0} = z_{0,1}$, $o_{2,1} = z_{1,2}$, and $o_{2,2} = z_{2,2}$).

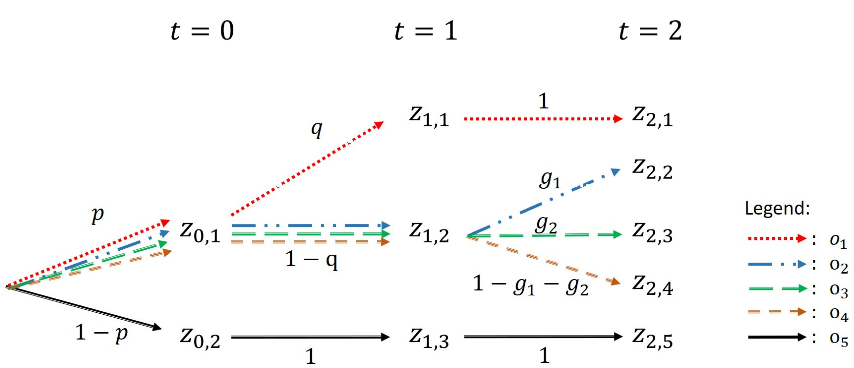


Figure 3: Illustration of the present-value method

The present value of outcome stream o_j, in utility terms, is given by $PV(o_j) = \sum_{t=0}^{T-1} \delta(t)v(o_{j,t})$. The present value of outcome stream o_1 in Figure 3 is, for instance, given by $v(z_{0,1}) + \delta(1)v(z_{1,1}) + \delta(2)v(z_{2,1})$, with $\delta(0) = 1$ as usually. We denote the number of distinct present values by m and rank-order them, so that $PV_1 > ... > PV_k > 0 > PV_{k+1} > ... > PV_m$. The probability of receiving the outcome stream with present value PV_j is denoted by q_j.

The overall value of the intertemporal prospect, W, is the sum of all present values multiplied by their decision weights as in atemporal prospect theory,

$$W = \sum_{j=1}^{k} \pi_j^+ PV_j + \sum_{h=k+1}^{m} \pi_h^- PV_h,$$

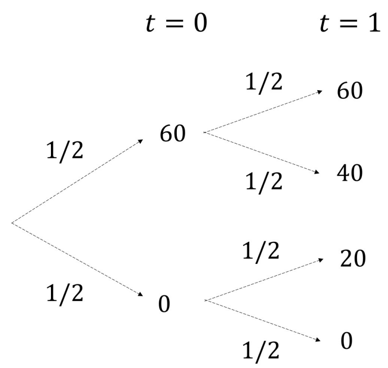
with the decision weights for gains and losses defined by

$$\begin{aligned}
\pi_1^+ &= w^+(q_1)\ ,\ \pi_m^- = w^-(q_m), \\
\pi_j^+ &= w^+(q_1 + ... + q_j) - w^+(q_1 + ... + q_{j-1}),\ \text{for } 1 < j \leq k, \\
\pi_h^- &= w^-(q_h + ... + q_m) - w^-(q_{h+1} + ... q_m),\ \text{for } k < h < m.
\end{aligned}$$

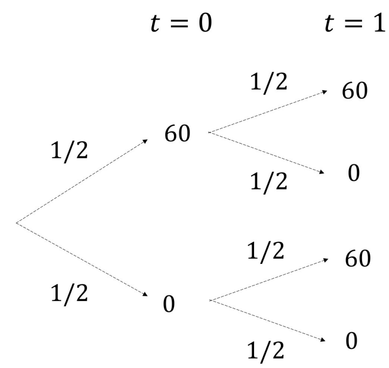
2.4 Differences in Evaluations

In general, both application methods yield different evaluations of prospects, assuming non-linear weighting of probabilities (with linear weighting, both methods always yield the same evaluation). However, for a subset of lotteries, both methods give the same PT value. Specifically, the methods give the same PT value for a lottery if any outcome stream of the lottery contains only non-negative or non-positive outcomes and if, in addition, the lottery posses a property that we call rank-order stability. We provide a formal definition in Appendix A and a proof that lotteries with these properties are indeed evaluated equally. Intuitively, rank-order stability means that if one outcome is larger than another outcome within one period, then, for any other period, the outcomes that may arrive in addition to the larger outcome (i.e., outcomes that are in one of the outcome streams containing this larger outcome) must be at least as large as the outcomes that may arrive in addition to the smaller outcome. Rank-order stability therefore implies that the best outcome stream bundles the best outcomes of each period.

Figure 4 displays two explanatory prospect. Both prospects contain only non-negative payouts. The first prospect is rank-order stable and its evaluation is, thus, identical for both application methods. The second prospect is not rank order stable. An arrival of the good outcome 60 at $t = 0$ can be followed by the arrival of 0 at $t = 1$, whereas the arrival of the outcome 0 at $t = 0$ can be followed by the arrival of 60 at $t = 1$.



(a) Rank-oder stable



(b) Not rank-order stable

Figure 4: Rank Order Stability (Examples)

Notes: This figure shows two examples of intertemporal prospects, one which is rank-order stable and one which is not rank-order stable.

For most economic or financial applications the evaluations between the methods will differ. For instance, if a good outcome in one period can be followed by a bad outcome in another period, rank-order stability does not hold. In addition, even if rank-order stability holds, the evaluations can differ if some outcome streams contain both positive and negative outcomes.

3 Experimental Design and Procedures

3.1 Decision Tasks

Each subject makes 48 decisions. For each decision, a subject is presented a lottery (framed neutrally as a risky option) with a time horizon of three periods ($t = 0, 1, 2$). The first period ($t = 0$) refers to the time of the experiment, the second period ($t = 1$) to three months after the experiment, and the third period ($t = 2$) to six months after the experiment. 16 of the lotteries only contain positive outcomes (gain lotteries), 16 only contain negative outcomes (loss lotteries), and 16 contain positive and negative outcomes (mixed lotteries).

For each lottery, we elicit a subject's switching point from the lottery to a safe option that yields three certain and identical payouts, the first received at the time of the experiment, the second three months later, and the third six months later. We elicit the switching point by asking subjects to position a slider that indicates for which payout amounts of the lottery they prefer the lottery over the safe option (and, vice versa, for which payout amounts of the lottery they prefer the safe option).

Figure 5 shows an example of such a decision task. At the top, subjects see the intertemporal lottery of this decision task (with non-trivial probabilities illustrated with a wheel of fortune). Below that, subjects see an illustration of the safe option (without the monetary values filled in, because these depend on subjects' choices in the task). Below that, subjects see a short description of the meaning of the position of the slider and the slider which they can position. When subjects move the slider (which is initially always positioned at the leftmost position) they increase the amount that they need to be offered in the safe option (three times) in order to prefer the safe option over the lottery. The bold numbers in the row above the slider change when the slider is moved. The difference between these two bold numbers is always one euro (in the analysis, we take the midpoint between these two number as the certainty equivalent).[5]

3.2 Lotteries

All lotteries used in the experiment, with a description of why they have been chosen, can be found in Appendix B. Here, we describe the structure behind the lottery choices more generally and show a few examples.

The following design features hold for all lotteries and aim at keeping the experiment comprehensible for participants: (1) each branching in the trees describing the intertemporal lotteries leads to at most two distinct outcomes in the following period; (2) in the last

[5]The limits of the slider (i.e., the leftmost and rightmost positions) are the minimal possible sum of payouts across the three time periods (without time discounting) of the intertemporal lottery in the decision task and the maximal possible sum of payouts.

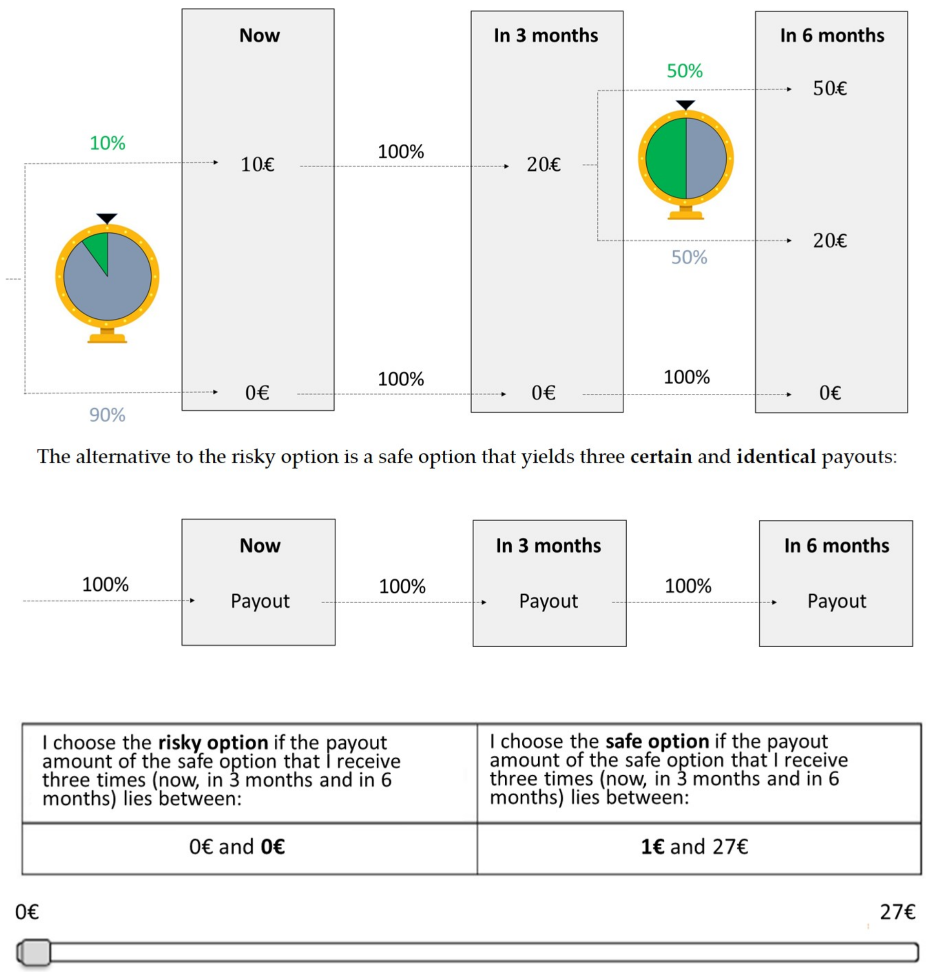


Figure 5: Example of a decision task

Notes: This figure shows a decision task. On top is the decision task's intertemporal lottery, followed by an illustration of the safe option (without payout values). The bottom part shows a description of the slider and the slider to elicit the certainty equivalent of the intertemporal lottery.

period, the number of distinct outcomes is limited to four (this is achieved by relying on outcomes that arrive with probability one in one time period); (3) outcomes appear only with probabilities 1, 9/10, 2/3, 1/2, 1/3, or 1/10.

The 48 lotteries of the experiment can be categorized in 6 sets, each containing 8 lotteries. The first 3 sets consist of low-stake lotteries: low-stake gain lotteries (Set 1), low-stake loss lotteries (Set 2), and low-stake mixed lotteries (Set 3). The other three sets consist of high-stake gain lotteries (Set 4), high-stake loss lotteries (Set 5), and high stake mixed lotteries (Set 6). The sets are summarized in Table 1. The different sets appear in random order. Lotteries within each set similarly appear in random order.

Table 1: Overview lottery sets

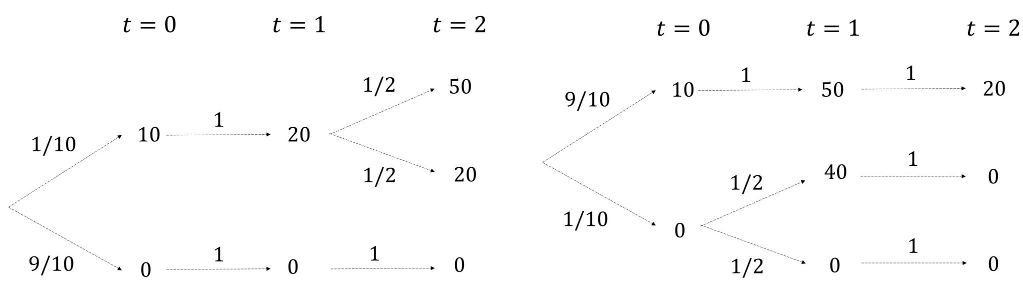
	Gains	Losses	Gains & Losses
Low stakes	Set 1	Set 2	Set 3
High stakes	Set 4	Set 5	Set 6

In general, the loss lotteries (Sets 2 and 5) can be obtained by multiplying all outcomes of the gain lotteries (Sets 1 and 4) by -1. The high-stake lotteries (Sets 4 to 6) can be obtained by multiplying the low-stake lotteries (Sets 1 to 3) by 20. The 8 lotteries of each set consist of 6 calibration and 2 test lotteries, implying a total of 36 calibration and 12 test lotteries. Below we provide further information on the calibration and test lotteries.

Calibration Lotteries: We use the certainty equivalents reported for the calibration lotteries to estimate the parameters of value function, weighting functions, and time-discount function jointly. The calibration lotteries of the sets ensure that all functions can be estimated on a relatively dense grid, for the value functions including small, intermediate and large positive and negative values.

The outcome streams of the calibration lotteries contain only non-negative or non-positive outcomes and the calibration lotteries are rank-order-stable. As explained in Section 2.4, this implies that for any given combination of value function, probability-weighting functions, and time-discount function, the evaluations under the time-separation method and the present-value method are identical. The parameter estimates obtained on the calibration lotteries are, thus, the same for both methods.

Figure 6 shows two explanatory calibration lotteries from Set 1 (low-stake gain lotteries). All outcomes are in euros. The outcome of Lottery 1 at $t = 0$ is either 10 with a low probability or 0 with a high probability. An arrival of the good outcome at $t = 0$ implies the arrival of good outcomes at $t = 1$ (20 with certainty as compared to 0 with certainty) and also at $t = 2$ (50 and 20 equally likely, compared to 0 with certainty). Lottery 1, hence, contains several unlikely good outcomes. Lottery 3, in contrast, contains several likely good outcomes. The variation in outcomes contributes to the value function being estimated on several small positive values. Similarly the variation in good outcomes' arrival probabilities contributes to the probability weighting function for gains being estimated on several values on the interval $[0, 1]$.



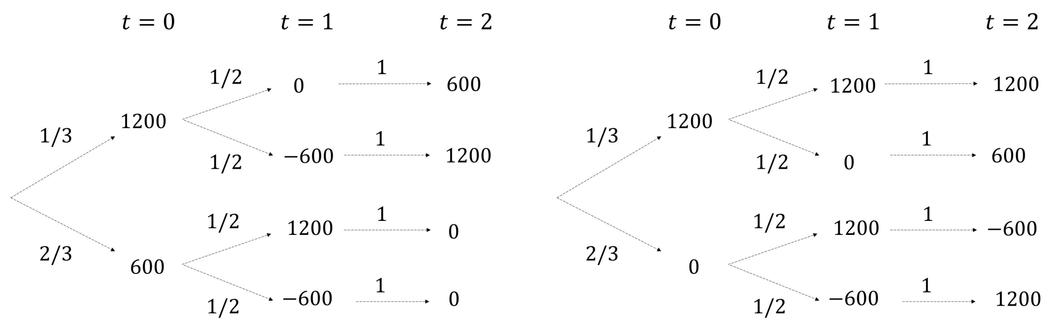
(a) Lottery 1 (b) Lottery 3

Figure 6: Calibration lottery examples, Lottery Set 1 (low-stake gain lotteries)

Test Lotteries: The test lotteries are designed such that the evaluations under the two methods give (sizably) different results. The test lotteries can therefore be used to compare the out-of-sample explanatory power of the two methods. In each of the six sets, one of the two test lotteries is chosen such that the time-separation method leads to a higher (multi-period) certainty equivalent than linear probability weighting if parameters are similar to the ones usually found for atemporal prospect theory, while the present-value method leads to a lower certainty equivalent. The opposite holds for the other test lottery. In order to assure a fair comparison between the two methods we choose the test lotteries such that the degree to which one method over- and the other method undervalues the lottery are similar (assuming typical atemporal calibrations).

Figure 7 shows the two test lotteries from Set 6 (high-stake mixed lotteries). In Lottery 47, the time-separation method would simultaneously underweight the most negative outcome at $t = 1$ (-600 arrives with probability 1/2) and overweight the most positive outcome at $t = 2$ (1200 arrives with probability 1/6). The time-separation method would therefore overvalue the lottery. For the present-value method, note that the lottery contains three positive outcome streams: $(1200, -600, 1200)$, $(1200, 0, 600)$, $(600, 1200, 0)$. The present value of these streams should be close to each other. The probability that one of these three streams arrives is 2/3. The present-value method would therefore underweight these positive streams and, thus, undervalue the lottery.

In Lottery 48, at $t = 1$ and at $t = 2$, the outcome 1200 arrives with probability 1/2. The time-separation method would underweight the positive outcome 1200 in these two periods and therefore undervalue the lottery. For the present-value method, note that the outcome stream $(1200, 1200, 1200)$ arrives with probability 1/6. The present-value method therefore overweights this stream and, hence, overvalues the lottery.



(a) Lottery 47

(b) Lottery 48

Figure 7: Test lottery examples, Lottery Set 6 (high-stake mixed lotteries)

In order to quantify the over- and undervaluation (that was just explained intuitively with the examples), we calculate the certainty equivalent of each lottery using typical atemporal specifications of the weighting and value functions (omitting time discounting). Note that, in line with our experimental design, this certainty equivalent would be received three times. Table 2 displays the results for two lotteries as illustration, showing that the degrees of over- and undervaluation are similar. Column (1) shows the certainty equivalent that is calculated with linear probability weighting. Column (2) displays the certainty equivalent that is calculated with the time-separation method, and Column (3) its relative deviation from the evaluation with linear probability weighting. Column (4) shows the certainty equivalent implied by an evaluation with the present-value method, and Column (5) its relative deviation from the evaluation with linear probability weighting. As shown in Table B.1 in Appendix B, these characteristics are shared by all 12 test lotteries enabling an a-priori fair comparison of the two application methods.

Table 2: Overview test lotteries

	No weighting	Time-separation		Present-value	
	CE	CE	Rel. Deviation	CE	Rel. Deviation
	(1)	(2)	(3)	(4)	(5)
Lottery 47	220.3	265.4	+20%	171.8	−22%
Lottery 48	195.9	141.1	−28%	248.8	+27%

Notes: To calculate the certainty equivalents we use the parametric specifications suggested by Tversky and Kahneman (1992) i.e, $v(x) = \mathbb{1}(x \geq 0)x^{0.88} - \mathbb{1}(x < 0)2.25(-x)^{0.88}$, and $w^{+}(p) = w^{-}(p) = \frac{p^{0.69}}{(p^{0.69}+(1-p)^{0.69})^{1/0.69}}$ (without time discounting).

3.3 Hypothetical and Incentivized Lotteries

Treatment comparisons are not the main focus of the paper. We implement two treatments in which the remuneration of participants varies. For participants in treatment T1, all decisions are hypothetical and participants receive a fixed payment only (15 euros). For participants in treatment T2 (who also receive the fixed payment of 15 euros), the gain lotteries are incentivized. Each participant in the experiment has a 25% chance to be assigned to the incentivized treatment T2. These two treatments serve as robustness check of whether choices are consistent in both treatments. In the incentivized treatment, the payouts from the three different time periods are also transferred to the participants' bank accounts at three different points in time (the payouts from the first time period are transferred jointly with the payment for participation in the experiment, the other payouts are transferred three and six months later).

For subjects in T2, one gain lottery is randomly selected for payment at the end of the experiment. There is a 99% chance that a low-stake lottery is selected and a 1% chance that a high-stake lottery is selected. Then a standard payment mechanism is employed. A random number is drawn from a uniform distribution, defined over the interval from $y^{-}/3$ to $y^{+}/3$, where y^{+} is the maximal and y^{-} the minimal non-discounted sum of payouts of the lottery (these are the limits of the sliders in the decision tasks, as described in Footnote 5). The terms y^- and y^+ are divided by three, because the certainty equivalent is paid thee times. If the drawn random number is greater than or equal to the subject's stated switch point, the subject receives the random number three times. If the random number is lower, the subject keeps the lottery and the reward is determined by a realization of the lottery.

3.4 Pre-Analysis Plan

The study was pre-registered at the Open Science Foundation (https://osf.io/7zyp4) with a detailed pre-analysis plan (https://osf.io/fv8a2). We follow the pre-analysis closely for our main analysis. Minor existing deviations from the pre-analysis plan are explicitly mentioned.

3.5 Procedures

The data was collected through CentERpanel on a sample consisting of about 9000 Dutch households that is representative for the Dutch population.[6] Data collection and data analysis were entirely separated. The experiment was conducted in Dutch. Instructions and other material was translated from English to Dutch by CentERpanel. We made sure in several iterations that the translations are accurate (one of us speaks Dutch and we checked with an experimental economist who is a native speaker). Before starting the decision tasks, subjects had to correctly answer a set of comprehension test questions. The English version of the instructions and comprehension test questions can be found in Appendix C. All subjects who completed the experiment received 15 euros for participating. Subjects in the incentivized treatment received on average 84 euros in addition to this.

In a post-experimental questionnaire, we inquired about the comprehension of the tasks and the attention paid in the experiment:

- During the decision tasks we asked you to position a slider. Was it clear for you what the position of the slider meant?

[6] Two pilots with about 25 subjects each were conducted online with the student subject pool of the Behavioral Lab of the University of St. Gallen. These pilots served to make the elicitation procedure and the instructions as comprehensible as possible.

- During the decision tasks we asked you to evaluate risky options with uncertain payouts at three different points in time. How complicated were these decision tasks?
- Please honestly report your attention during the questionnaire.

Each of these questions only had three answer possibilities. As outlined in our pre-analysis plan, we exclude all subjects from the analysis that answer at least one of these questions with the 'worst' answer possibility (indicating that it was unclear what the slider meant, that the decision tasks were too complicated, or that a subject only paid attention in a few parts of the experiment or not at all). In addition (and as pre-registered), we exclude all subjects that completed the tasks with a median decision time of less than 15 seconds. We use such strict exclusion criteria to make sure that we only use data from subjects who understood the tasks and took them seriously (so that the temptation to click through as fast as possible to collect the participation fee in a short period of time does not drive our results).

391 subjects started the decision tasks, 13 of them did not finish the experiment.[7] Of the remaining 378 subjects, we exclude 192 participants based on the answers to the post-experimental questionnaire and an additional 86 participants that finished the decision tasks with a median time per task of less than 15 seconds. Following our pre-registered data exclusion rules, we therefore drop about 74% of our sample.[8] Out of the exactly 100 subjects that remain in our main analysis, 27 are in the incentivized treatment and 73 in the hypothetical treatment. Demographic variables in the 100 subjects of the main analysis are similar to the demographic variables among all participants starting the experiment and also among the Center Panel subject pool (details can be found in the appendix, in Table D.4). Additional information on responses in the post-game survey, the distribution of median decision times, and demographic characteristics of participants and Panel subject pool can be found in Appendix D.

Exclusion rates for studies that estimate preference parameters on a representative sample

[7] Our pre-registration specifies that the target number of participants was 240. If 240 had been reached by the end of September 2020, the experiment should have stopped, otherwise continued in October. At the end of September, 205 subjects had completed the experiment. We did not receive the data at this point, only summary information on payments and on the responses from the post-experimental questionnaire. Based on this, we asked CentERpanel to continue the data collection in October to obtain in total between 350 and 400 participants. Data collection then stopped at the end of October and we obtained the data from CentERpanel in December.

[8] Using the full sample of 378 subjects would not affect our main conclusion of the superiority of the present-value method over the time-separation method.

are in general only slightly below ours, despite the fact that most other studies have less complex tasks without time and risk dimensions simultaneously. See, for instance, Booij et al. (2010), who also use the CentERPanel, with an exclusion rate of 61%, and Dohmen et al. (2005) and Guiso and Paiella (2008), using representative samples in Germany and Italy, who drop 61% and 57% of their observations, respectively (the reasons for exclusion differ between the studies).

4 Estimation Procedure

We first introduce the parametric specifications that we consider for the value, probability-weighting, and time-discount functions. Then we discuss the maximum-likelihood procedure with which we estimate the model parameters on the calibration set and how we measure prediction performance on the test set. The described approach is exactly as in the pre-analysis plan.

4.1 Parametric Specification

We consider two value functions, three weighting functions and two time-discount functions. These are briefly discussed below and shown in Table 3.

As common in the literature (e.g., Sugden, 2003; Köbberling and Wakker, 2005; and Abdellaoui and Kemel, 2013), we assume that the value function is composed of a loss aversion coefficient λ and basic utility functions for gains and losses. We focus on the two most common specifications, power utility (forcing the power coefficients of gains and losses to be identical, as is common to obtain an interpretable loss aversion coefficient; e.g., Köbberling and Wakker, 2005; Harrison and Rutström, 2008; Post et al., 2008; and Tanaka et al., 2010) and exponential utility.

We focus on three common parametric specifications of the probability weighting functions. The same type of function is always used for gains and for losses, but we allow the parameters to differ. The first option we consider is the function used by Tversky and Kahneman (1992), with only one parameter. The second weighting function considered in this study is the two-parameter specification proposed by Prelec et al. (1998). The third weighting

Table 3: Function specifications

Specification	Parameters
Value functions	
Power utility: $v(x) = \mathbb{1}_{x\geq 0}x^{\alpha} - \mathbb{1}_{x<0}\lambda\ (-x)^{\alpha}$	α, λ
Exponential utility: $v(x) = \mathbb{1}_{x\geq 0}\frac{1-\exp(-\alpha x)}{\alpha} - \mathbb{1}_{x<0}\lambda\frac{1-\exp(\beta x)}{\beta}$	α , β, λ
Probability weighting functions (gains and losses)	
Tversky and Kahneman (1992): $w(p) = \frac{p^{\gamma}}{(p^{\gamma}+(1-p)^{\gamma})^{1/\gamma}}$	γ^{+}, γ^{-}
Prelec (1998): $w(p) = \exp(-\eta(-ln(p))^{\gamma})$	η^{+}, η^{-}, γ^{+}, γ^{-}
Goldstein Einhorn (1987): $w(p) = \frac{\eta p^{\gamma}}{\eta p^{\gamma}+(1-p)^{\gamma}}$	η^{+}, η^{-}, γ^{+}, γ^{-}
Time-discount functions	
Exponential discounting: $\delta(t) = \exp(-rt)$	r
Quasi-Hyperbolic discounting: $\delta(t) = \mathbb{1}_{t=0} + \mathbb{1}_{t>0}k\exp(-rt)$	k, r

Notes: This table shows the formulas for the considered value functions, probability weighting functions (for gains and losses), and time-discount functions. For each function the table lists the parameters that have to be estimated.

function considered in this study is the two-parameter specification proposed by Goldstein and Einhorn (1987).

The time-discount function captures the discounting of future outcomes. We focus on the two most common specifications, exponential discounting and quasi-hyperbolic discounting (which discounts all future outcomes when compared to rewards at time $t = 0$, in addition to exponential discounting).

As displayed in Table 4, we analyze each possible combination of these functions (there are 12 such combinations) for the time-separation method as well as for the present-value method.

Table 4: Overview function combinations

		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

Notes: This table shows the twelve combinations of the two types of value functions, three types of probability weighting functions, and two types of time-discount functions.

4.2 Maximum-Likelihood Estimation

Each participant reports a multi-period certainty equivalent for a total of 48 lotteries, consisting of 36 calibration and 12 test lotteries. We use the multi-period certainty equivalents reported for the calibration lotteries to estimate the model parameters.

Consider model Combination C1 with parameters $\alpha, \lambda, \gamma^{+}, \gamma^{-}$ and r, where the parameters have the meanings as above. That is, the model corresponds to power utility, a Tversky-Kahnemann probability weighting function, and exponential discounting (the other model specifications are handled accordingly). As discussed before, we allow the parameters of the weighting function to differ for gains and losses. These parameters are labeled as γ^{+} and γ^{-}.

Denote by $V(L_j|\alpha, \lambda, \gamma^{+}, \gamma^{-}, r)$ the PT value of the j-th lottery (L_j), arising with one given parametric specification. The multi-period certainty equivalent of the lottery, ce_j, is a certain payout that arises three times (at $t = 0$, at $t = 1$, and at $t = 2$), so that the three payments jointly have a PT value equal to the PT value of the lottery. The PT value of the certainty equivalent is $v(ce_j) + \exp(-r)v(ce_j) + \exp(-2r)v(ce_j)$, independently of which application method of prospect theory is used (the payment is certain, there is thus no probability weighting). Indifference between the certainty equivalent and the lottery implies that the certainty equivalent of a lottery can be calculated as $ce_j(\alpha, \lambda, \gamma^{+}, \gamma^{-}, r) = v^{-1}\left(\frac{1}{1+\exp(-r)+\exp(-2r)} V(L_j|\alpha, \lambda, \gamma^{+}, \gamma^{-}, r)\right)$, with v^{-1} denoting the inverse of the value function.

Following existing literature (e.g., Hey et al., 2009; Bruhin et al., 2010), we write the observed certainty equivalent of participant i, $CE_{i,j}$, as $CE_{i,j} = ce_j + \epsilon_{i,j}$ (note that we use capital letters for observed data). The term $\epsilon_{i,j}$ reflects a variety of sources (among others hurrying or inattentiveness). We assume that $\epsilon_{i,j}$ is normally distributed with mean 0 and standard deviation $\sigma_{i,j}$. In addition, we allow for two sources of heteroskedasticity. First, as the employed set of lotteries has different payout ranges, we assume that the error variance is proportional to the payout range of the lottery. The payout range of lottery j, w_j, is calculated as $w_j = |L_{j,max} - L_{j,min}|$, where $L_{j,max}$ denotes the maximal and $L_{j,min}$ the minimal non-discounted sum of payouts of lottery j. Second, as participants may be heterogeneous (e.g., with respect to their attention span), we allow the error variance to differ by participant. This yields the form $\sigma_{i,j} = \epsilon_i w_j$ for the standard deviation of the error term, where ϵ_i denotes a participant-specific parameter.

Given our assumptions on the distribution of the error term, the contribution of participant i to the likelihood function L can be expressed as

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

where the vector $CE_i = (CE_{i,1}, ..., CE_{i,36})$ contains the certainty equivalents participant i reports for the 36 calibration lotteries, and the vector $\theta = (\alpha, \lambda, \gamma^+, \gamma^-, r)$ contains the parameters of the model. Furthermore, ϕ denotes the density of the standard normal distribution. Using data from all n participants, the log-likelihood function is given by

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

where the vector $CE = (CE_1, ..., CE_n) = (CE_{1,1}, ..., CE_{1,36}, ..., CE_{n,1}, ..., CE_{n,36})$ contains the certainty equivalents reported for the calibration lotteries by all n participants, and the vector $\epsilon = (\epsilon_1, ..., \epsilon_n)$ contains the participant-specific error variance parameters.

For the estimation of the parameters, we maximize the log-likelihood function using standard iterative algorithms. With this estimation strategy we simultaneously estimate the model parameters $(\alpha, \lambda, \gamma^+, \gamma^-, r)$ and the participant-specific error variance parameters ϵ_1,...,ϵ_n.

4.3 Measurement of Prediction Performance on the Test Set (Main Outcome Variable)

We use the multi-period certainty equivalents reported for the 12 test lotteries to measure the (out-of-sample) performance of the two methods. For each of the two methods we use the model parameters that were estimated on the calibration set (e.g., $\widehat{\alpha}, \widehat{\lambda}, \widehat{\gamma^{+}}, \widehat{\gamma^{-}}, \widehat{r}$) to predict choices of the multi-period certainty equivalents in the test set,

$$\widehat{ce_j} = v^{-1}\left(\frac{1}{1+\exp(-\widehat{r})+\exp(-2\widehat{r})} V(L_j|\widehat{\alpha}, \widehat{\lambda}, \widehat{\gamma^{+}}, \widehat{\gamma^{-}}, \widehat{r})\right).$$

As measure of prediction performance, we use the (weighted) mean squared errors. The (weighted) mean squared error of participant i is calculated as

$$MSE_i = \frac{1}{12}\sum_{j=1}^{12}\left(\frac{1}{w_j}\left(CE_{i,j} - \widehat{ce_j}\right)\right)^2,$$

where $w_j = |L_{j,max} - L_{j,min}|$, as before. The weighting ensures that the error is not disproportionately driven by lottery choices involving large (absolute) payout amounts. Using data from all n participants, we define our main outcome variable as

$$MSE = \frac{1}{n}\sum_{i=1}^{n} MSE_i.$$

Out of all 24 combinations of application method and functional specification, we consider the combination best that leads to the lowest MSE.

The predicted certainty equivalents $\widehat{ce_j}$ and thus MSE_i and MSE depend on the choice of the application method. We denote by MSE_{TS} and MSE_{PV} the MSE resulting from the evaluation under the time-separation method and under the present-value method, respectively (given a combination of value, probability-weighting, and time-discount functions). We denote the difference between these two values by

$$\triangle MSE := MSE_{TS} - MSE_{PV}.$$

For each of the 12 functional specifications, $\triangle MSE$ indicates whether one application method describes decisions better than the other.

4.4 Tests and Standard Errors

To test whether the two application methods differ with respect to their prediction performance, we investigate the 12 functional specifications in isolation. For each specification we test whether the difference between the two (weighted) mean squared errors, i.e. $\triangle MSE$, is statistically significant different from 0. The null hypothesis H_0 is $\triangle MSE = 0$ and the (two-sided) alternative hypothesis H_a is $\triangle MSE \neq 0$.

Therefore, we conduct (non-parametric) paired bootstrap tests. We draw from our sample with replacement to obtain bootstrap samples of the same size as our original sample (resampling is done at the level of participants). For each bootstrap sample we follow the procedure described in Sections 4.2 and 4.3. We first estimate the parameters on the calibration lotteries and then calculate the two (weighted) mean squared errors (one for each application method) on the test lotteries. We denote by $MSE_{b,TS}$ and $MSE_{b,PV}$ the two MSEs that are estimated on bootstrap sample b and by $\triangle MSE_b$ their difference, $\triangle MSE_b = MSE_{b,TS} - MSE_{b,PV}$. We reject the null hypothesis if $\triangle MSE_b \geq 0$ in less than 2.5% of the bootstrap samples or if $\triangle MSE_b \leq 0$ in less than 2.5% of the bootstrap samples.

We also use these bootstrapped samples to provide standard errors of the parameter estimates (estimated on the calibration lotteries of the full sample).

4.5 Final Re-estimation

In a final step, we pool the data on the calibration and test lotteries (thus making use of the full 48 certainty equivalents per participant) to reestimate the parameters, at least for the most successful combination of application method and functional specification (to provide a complete model benchmark of how to best apply intertemporal prospect theory). We again provide bootstrapped standard errors of parameters (as for the parameters estimated on the calibration set).

5 Results

We first focus on the main result, which is the comparison of the prediction performance of the two application methods. Thereafter, we re-estimate the parameters on the pooled training and test set. Analyses that are pre-registered as additional analyses can be found thereafter in Sections 5.3 and 5.4. In Section 5.3, we analyze which components of PT are essential for good prediction performance. In Section 5.4, we analyze an additional possible application method of PT, a version of the present-value method. In Section 5.5 (not pre-registered), we show that our results are robust to allowing reference points to differ from zero. As specified in the pre-analysis plan, we analyze the data from both treatments jointly, as they are sufficiently similar according to the pre-defined critera (details can be found in Section 5.6). The number of observations for the analyses is 100.

5.1 Time-separation vs. Present-value Method

As discussed in Section 4, we first estimate the parameters for each of the 12 combinations separately on the calibration lotteries. We then use these parameters to compare the prediction performance of the models for the test lotteries. The parameter estimates on the calibration set (which are identical for both application methods) are not very different from the estimates on the whole sample (shown below in Section 5.2), therefore we do not discuss them here (they can be found in Appendix D). Note that the probability-weighting functions are clearly non-linear in all cases, which is necessary for the application methods to lead to different predictions on the test set.

The main result of the comparison of the different application methods can be found in Figure 8. This figure shows the mean squared errors on the test set for both application methods and all 12 combinations of function specifications. It can be seen that the present-value method always performs better than the time-separation method. This difference is sizable and highly statistically significant: Figure 9 shows the difference between the mean squared errors jointly with bootstrapped 95%-confidence intervals.[9]

[9] The pre-registration states that we also provide a test in which we compare the best performing of the 12 combinations for the present-value method with the best performing combination for the time-separation method. This difference is naturally also significant, because the performance of the function combinations

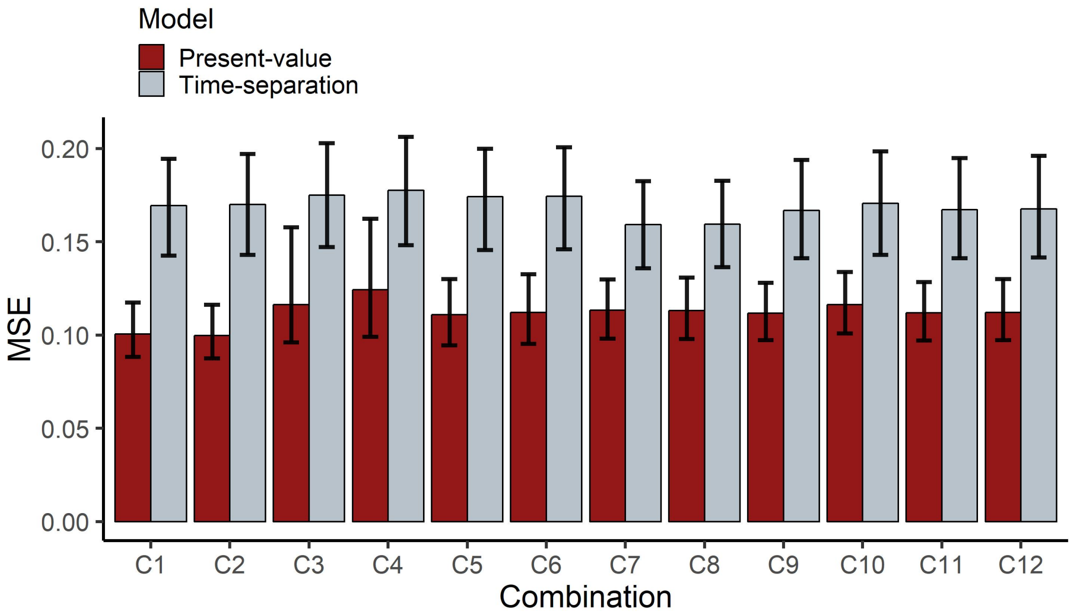


Figure 8: MSE for both application methods (all function combinations)

Notes: This figure shows mean squared errors of present-value and time-separation method on the test set for all 12 function combinations (with bootstrapped 95% confidence intervals).

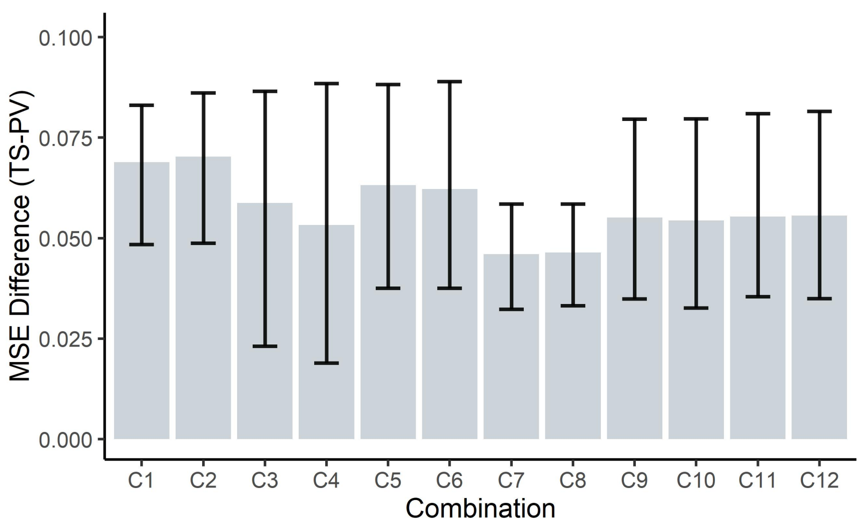


Figure 9: Difference in MSE between the application methods (all function combinations)

Notes: This figure shows the average differences in mean squared errors (MSE of time-separation method minus MSE of present-value method) with bootstrapped 95% confidence intervals.

The difference in prediction performance is not driven by a particular subset of the test lotteries. For each of the twelve combinations, the present-value method has a lower MSE than the time-separation method for each single test lottery (Table D.6 in the appendix contains the MSE per lottery for each function combination). To illustrate the magnitude of the differences in prediction performance, Table 5 shows a lottery-by-lottery comparison for all test lotteries for combination C1, with the mean absolute prediction errors instead of mean squared errors for easier interpretation (Table D.5 in the appendix shows a by-lottery comparison of all twelve combinations).

Table 5: 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	5.6	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.

One can also illustrate the difference by classifying participants as time-separation or present-value types (for illustration; this classification is not mentioned in the pre-analysis plan). If the MSE of one participant for one method is greater than the MSE for the other with a difference of at least one standard error (of the MSE difference between the two methods), a participant is classified as a time-separation or a present-value type (otherwise the person remains unclassified). Table 6 displays the results. For all twelve combinations, the vast majority of subjects is classified as present-value type. The share of present-value types ranges from 81 to 89 percent, with 8 to 18 percent unclassified and 1 to 10 percent time-separation types. Figure D.1 in the Appendix, displays the distribution of the difference between the two MSEs showing that these are sizable for a majority of the participants.

C1-C12 within one application method is very similar (which can be seen from Figures 8 and 9).

Note that in a few cases the predictions of the time-separation are so extreme that they fall outside of the slider limits. In these cases we replace the predictions by the slider limits (not adjusting the prediction would lead to an even larger difference in MSE between the two methods).

Table 6: 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	83	81	89	88	87	87	86	89	88	88
Unclassified	9	11	16	18	10	11	8	8	11	10	10	11

Notes: Participant i is classified as time-separation type if $MSE_i^{PV} - SE(\triangle MSE) > MSE_i^{TS}$ or as present-value type if $MSE_i^{TS} - SE(\triangle MSE) > MSE_i^{PV}$. $SE(\triangle MSE)$ denotes the standard error of the difference in MSE between the two methods.

5.2 Parameter Estimates

Here, we present and discuss the parameter estimates from the reestimation on the full set of lotteries. Because the present-value method is the superior evaluation method, we only present the parameter estimates resulting from an evaluation under this method. The results for the time-separation method can be found in Appendix D.

Which utility function is used (power or exponential) and which of the two-parameter probability-weighting functions is used (Prelec et al. (1998) or Goldstein and Einhorn (1987)) has only minor implications for the shapes of the calibrated functions. We therefore only show the parametrizations of the four combinations that employ a power utility and either one-parameter weighting functions as suggested by Tversky and Kahneman (1992) (C1 and C2) or two-parameter weighting functions suggested by Goldstein and Einhorn (1987) (C5 and C6). C1 and C5 use exponential time-discounting, C2 and C6 quasi-hyperbolic discounting. The estimates are shown in Table 7 (with standard errors in parentheses). The estimates for all twelve model combinations can be found in Appendix D.

Value function: The value functions consist of a basic utility functions for gains and losses and a loss aversion parameter (λ). The basic utility functions are in general close to linear, with utility in the gain domain mostly slightly convex and in the loss domain slightly concave (for the four combinations C9 to C12 that employ an exponential utility and two-parameter probability weighting functions, the estimated utility is slightly concave for gains and convex for losses, see the full Table D.8). In the atemporal literature the utility for gains is generally found to be concave and the utility for losses is found to be convex or linear.[10]

[10]However, similar shapes have already been observed in the atemporal literature. Bruhin et al. (2010),

Table 7: Overview parameter estimates on full sample (48 lotteries)

	Value			Weighting				Discounting	
Combination	α	β	λ	$\gamma+$	$\gamma-$	$\nu+$	$\nu-$	r	k
C1	1.195		1.069	0.535	0.551			0.117	
	(0.068)		(0.096)	(0.045)	(0.06)			(0.036)	
C2	1.172		1.065	0.547	0.565			0.016	0.863
	(0.061)		(0.121)	(0.043)	(0.057)			(0.025)	(0.043)
C5	1.091		1.037	0.426	0.448	0.766	0.786	0.070	
	(0.035)		(0.127)	(0.066)	(0.073)	(0.074)	(0.071)	(0.028)	
C6	1.085		1.101	0.423	0.451	0.786	0.796	0.0012	0.884
	(0.03)		(0.163)	(0.07)	(0.073)	(0.073)	(0.068)	(0.016)	(0.045)

Notes: This table shows the parameter estimates on the full set of lotteries (calibration plus test set) for the present-value method, for four selected function combinations. Estimates for all twelve function combinations and for the time-separation method can be found in Appendix D.

A potential explanation for this relates to complexity. The present-value method aggregates the outcomes of each stream into a single present value. A simple form of aggregation is the (time-discounted) sum of all outcomes (monetary values). When subjects aggregate outcomes, they may, due to the complexity of the task, evaluate outcomes in payout terms rather than in (non-linear) utility terms. Note that Abdellaoui, Bleichrodt, l'Haridon and Paraschiv (2013), who estimate intertemporal utility without risk, similarly find that utility is mildly convex for gains and mildly concave for losses.

The second component of the value function is the loss aversion parameter λ. The loss aversion parameters are on average slightly above one. This is lower than what is usually observed in the atemporal literature. According to a recent meta study (Brown et al., 2021), the median loss aversion parameter among 286 aggregate level studies is 1.5. However, the few studies that estimate loss aversion in an intertemporal context (without risk) also find loss aversion coefficients closer to one (Abdellaoui, l'Haridon and Paraschiv, 2013; Abdellaoui, Bleichrodt, l'Haridon and Paraschiv, 2013). Following the complexity explanation from above, subjects may evaluate outcomes in monetary values for simplicity rather than in utility terms and thereby not distinguish between losses and gains (as these do not appear in

for instance, calibrate a combination of Goldstein-Einhorn weighting functions and power utility, as our C5 and C6, with very similar estimates of weighting and utility functions, including basic utility that is convex for gains and concave for losses.

pure monetary values). A second potential explanation is that participants may aggregate losses and gains rather than evaluating these in isolation. Suppose, for instance, that an outcome stream yields 60 at $t = 1$ and -40 at $t = 2$. If subjects aggregate gains and losses, the evaluation of the stream should (neglecting time discounting) be close to the evaluation of a payout of 20 at $t = 1$ and no payout at $t = 2$. The present-value method (as also the time-separation method), however, transforms each outcome to a utility term before aggregating. The evaluation in isolation is only consistent to the aggregation of gains and losses if the loss aversion parameter is close to 1. A version of the present-value method (pre-registered as additional analysis) calculating present values in monetary terms before entering them in a utility function is treated below in Section 5.4.

Probability weighting functions: The calibrations of the probability-weighting functions turn out to be very similar for gains and for losses for all function combinations, suggesting that one can use the same functions in the gain and in the loss domain (reducing the number of model parameters). For all specifications, an inverse S-shape is found, as common in the atemporal literature. The shapes of the weighting functions are also similar across the different specifications, with the two-parameter weighting functions ascending faster for low values than the one-parametric Tversky-Kahneman specification. Figure 10 displays the estimated weighting functions for gains for model combinations C1 and C5 as illustrations. In general, our estimates of probability weighting functions are very similar to those estimated in atemporal contexts (e.g., Stott, 2006; Booij et al., 2010; Bruhin et al., 2010).

Time-discount function: Exponential discounting and quasi-hyperbolic discounting lead to very similar prediction performances in our setting. The estimates for exponential discounting reach from zero to about 12% per quarter. Estimates for quasi-hyperbolic discounting reach from no discounting at all to discounting all future payments by about 15% (with very small additional exponential discounting when all future outcomes are already discounted). These discounting patterns are similar to those found in risk-less intertemporal decision tasks, as reported in the recent meta study by Havránek et al. (2021).

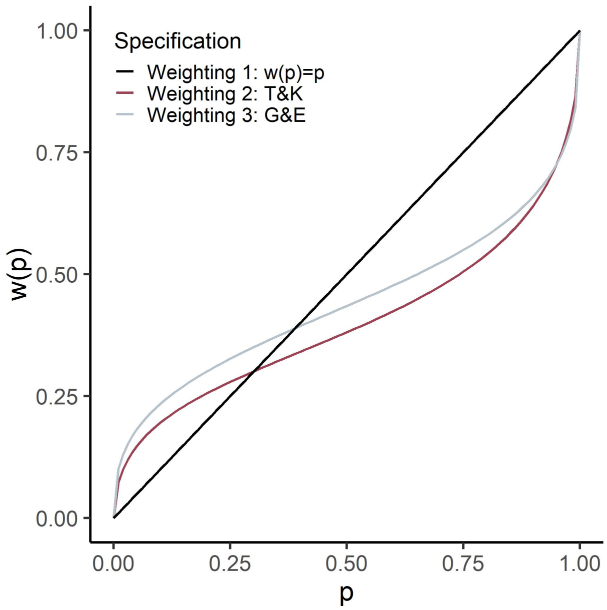


Figure 10: Estimated probability-weighting functions

Notes: This figure shows the estimated weighting functions for gains for function combinations C1 and C5 (for the present-value method). C1 and C5 both use a power value function and exponential discounting.

5.3 The Importance of the Components of PT in Intertemporal Settings

As mentioned in the pre-analysis plan, we also investigate the importance of PT's components. To do so, we show how the performance deteriorates when the probability weighting functions are linear, when there is no loss aversion, or when the utility functions are linear (with all free parameters estimated anew). We also compare the methods to discounted expected utility maximization. In this section we focus on the results for the present-value method, an investigation for the time-separation method can be found in Appendix D.

The analyses in Sections 5.3 and 5.4 are described as additional analyses in the pre-analysis plan (with little detail; the analysis is kept as similar to the main analysis as possible). The core design feature of our calibration set is that the calibrations are the same for time-separation and present-value method. In addition, the test lotteries are such that the two methods over- or underestimate the lotteries to a similar extent with typical atemporal calibrations. These features ensure that our design does not favor one of the two methods. Both features, however, do not hold when also considering other methods (including expected discounted utility). Therefore, the comparison here differs as follows.

We randomly select six of the eight lotteries from each set, implying a total of 36 lotteries, as the new calibration lotteries. We repeat this random selection process 10 times. For each of these ten repetitions the 36 calibration lotteries are used to calibrate a models' parameters. The other twelve lotteries (two per set) are used as test lotteries to compare the predictive performance, measured by the (weighted) MSE. The numbers of calibration and test lotteries are therefore the same as in the main analysis. We denote by MSE_{S_i} the (weighted) MSE of the random selection S_i, $i \leq 10$. The main outcome variable is the mean of these ten values,

$$MSE = \frac{1}{10} \sum_{i=1}^{10} MSE_{S_i}$$

Bootstrapped confidence intervals for the MSE are obtained as in the main analyses.

The results are shown in Figure 11. The prediction performances of models with linear utility for gains and losses ('PT: Linear Utility'; this model allows for a loss aversion coefficient different from one) or no loss aversion ('PT: No Loss Aversion') are almost identical to the performance of the full PT application ('PT: All Components'). The same holds if utility is not only required to be linear for gains and losses, but on the whole domain, including the absence of loss aversion ('PT: Linear Utility and No Loss Aversion'). This implies that utility curvature and loss aversion can be omitted if a more parsimonious model is desired. However, forcing linear-probability weighting ('PT: Linear Probability Weighting') predicts decisions significantly worse.[11] These findings are consistent with the parametrizations discussed in Section 5.2, which exhibit almost linear utility with a loss aversion parameter close to one and highly non-linear probability weighting functions.

We also compare the prediction performance of PT models to standard expected discounted utility models with linear probability weighting, no loss aversion, and no hyperbolic discounting ('EDU'). Figure 11 shows that EDU predicts decisions significantly worse than

[11]Several of the MSE bars in Figure 11 are identical across the function combinations. For instance, when utility is assumed to be linear there is no more distinction between power utility (C1 to C6) and exponential utility (C7 to C12). For EDU, there are even only two specifications: one with power utility (C1 to C6) and one with exponential utility (C7 to C12), as there is neither probability weighting nor hyperbolic discounting.

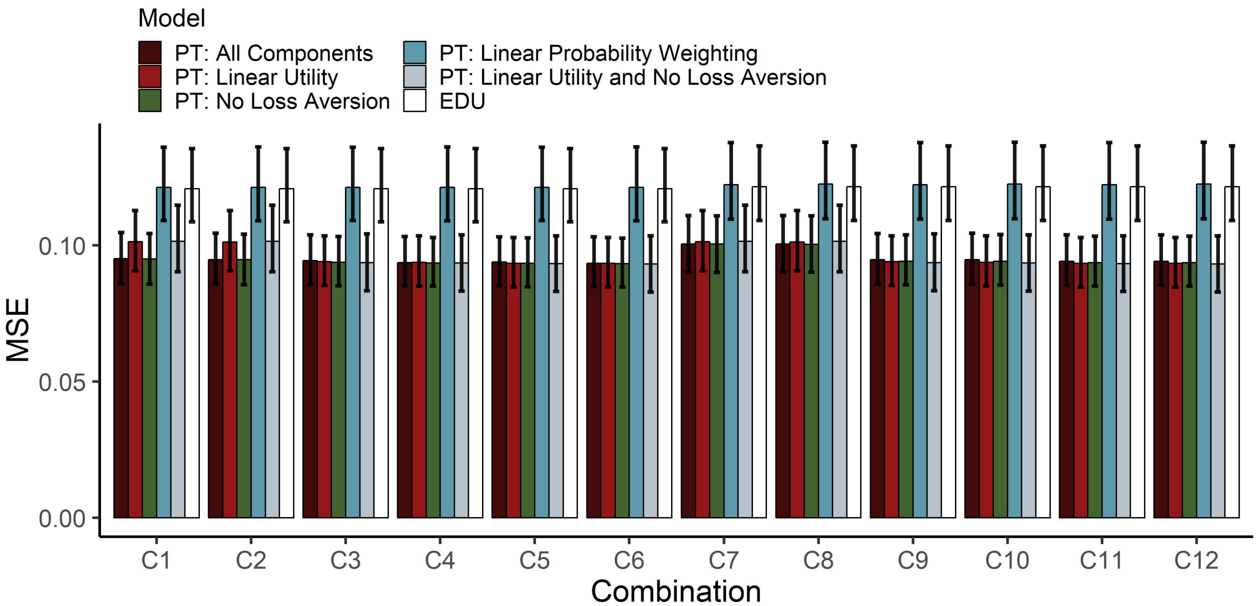


Figure 11: MSE, PT components (all function combinations)

Notes: This figure compares the prediction performance of the present-value method to versions of it with some components absent and to expected utility theory (parameters are estimated anew for each method). The versions are (from left to right for each of the 12 function combinations) the present-value method with all components ('PT: All Components'), with linear utility for gains and losses still including a loss-aversion parameter ('PT: Linear Utility'), with a loss-aversion parameter of 1 ('PT: No Loss Aversion'), with linear-probability weighting ('PT: Linear Probability Weighting'), with linear utility for gains and losses and a loss-aversion parameter of 1 ('PT: Linear Utility and No Loss Aversion'); on the right expected discounted utility is shown ('EDU').

PT, including the sparser versions that still allow for non-linear probability weighting. The prediction performance of EDU and the PT version with linear probability weighting is, however, almost identical. The calibration of EDU utility functions is almost linear.

5.4 A Third Way of Applying Prospect Theory in Intertemporal Situations

In addition to the two main methods, we also analyze the performance of a third application method of prospect theory (as described in the pre-analysis plan). This method is in spirit similar to the present-value method. Whereas the regular present-value method calculates the present value of an outcome stream (before probability weighting plays a role) as $PV(o_j) = \sum_{t=0}^{T-1} \delta(t) v(o_{j,t})$, the third application method calculates it as $PV(o_j) = v\left(\sum_{t=0}^{T-1} \delta(t) o_{j,t}\right)$. That is, instead of calculating the present value in utility terms by discounting the utility obtained each period, it first calculates a present value in monetary terms and then transforms this into utility (this happens again before probability weighting plays a role). We call the third application method monetary present-value method.[12]

We find the method intuitively appealing, because the outcomes of the lotteries are of monetary nature and money can be transferred from one period to another (in contrast, we find the regular present-value more intuitive if the outcomes are not of monetary nature but, for example, final consumption that cannot be transferred from one period to another). One may also argue that aggregating outcomes in monetary terms is less demanding than aggregating outcomes in utility terms, so that the method may correspond more closely to simple decision making by participants.

Figure 12 shows the predictive power of the monetary present-value method and the regular present-value method. The comparison is again based on 10 random draws of calibration and test lotteries, as in Section 5.3 (the degrees of freedom are always identical for a given function combination). It can be seen that the predictive power of the two present-value methods is virtually identical.

We re-calibrate the functions for the monetary present-value method on the full set of lotteries. As the choice of the utility function (power or exponential) and also the choice of the two-parameter probability weighting function (Prelec et al., 1998, or Goldstein and Einhorn, 1987) has again only minor implications for the parametrizations, we only show combinations C1, C2, C5, and C6 in Table 8, as before (for estimates of all combinations, see Appendix D).

The calibrations are overall similar to those of the regular present-value method (Table 7). There are some differences, however. The loss-aversion coefficient is on average slightly higher for the monetary present-value method. This makes intuitively sense if one considers loss aversion to be present at the "overall" level rather than in each period: the loss-aversion coefficient in the monetary present-value method captures aversion to losses when the payouts

[12] We are not aware that this method has, in this form, been explicitly mentioned in the literature. Note, however, that the "time-first" specification in Rohde and Yu (2020), which is kept general, can be restricted to become the regular or the monetary present-value method.

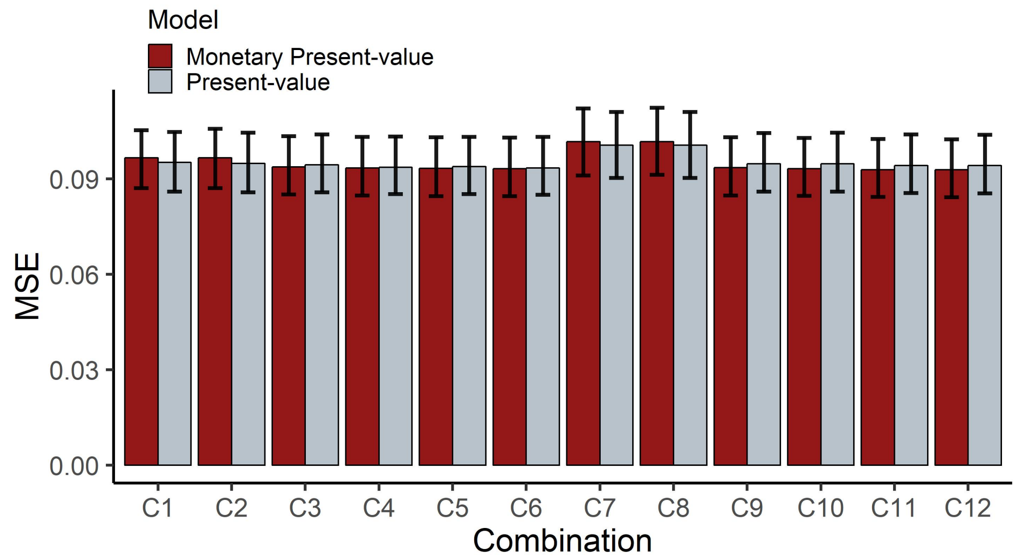


Figure 12: MSE for the present-value method and the monetary present-value method

Notes: This figure shows mean squared errors of the monetary present-value and the (regular) present-value method on the test set for all 12 function combinations (with bootstrapped 95% confidence intervals).

Table 8: Calibration monetary present-value method

	Value			Weighting				Discounting	
Combination	α	β	λ	$\gamma+$	$\gamma-$	$\nu+$	$\nu-$	r	k
C1	1.179		1.133	0.579	0.592			0	
	(0.075)		(0.227)	(0.04)	(0.058)			(0)	
C2	1.182		1.141	0.576	0.576			0	1
	(0.072)		(0.224)	(0.041)	(0.058)			(0)	(0.001)
C5	0.915		1.033	0.403	0.406	1.036	1.05	0.041	
	(0.049)		(0.223)	(0.068)	(0.075)	(0.097)	(0.112)	(0.023)	
C6	0.915		1.087	0.406	0.421	1.03	1.05	0.001	0.938
	(0.049)		(0.200)	(0.069)	(0.074)	(0.096)	(0.118)	(0.009)	(0.032)

Notes: This table shows the parameter estimates on the full set of lotteries (calibration plus test set) for the monetary present-value method, for four selected function combinations. Estimates for all twelve function combinations can be found in Appendix D.

over the periods are summed up (after discounting), while the loss-aversion coefficient in the regular present-value method captures aversion to losses within each period. While the utility is still mildly convex for gains and mildly concave for losses for the combinations that employ the one-parameter weighting function the utility is now mildly concave for gains and mildly convex for losses for combinations that employ a two-parameter weighting function.

For all combinations, the curvature parameters of the probability weighting functions (γ^{+} and γ^{-}) are almost identical to those estimated for the regular present-value method. All functions, thus, again have an inverse-s shape. However, for the two-parametric probability weighting functions, the elevation (measured by $\nu+$ and $\nu-$) is greater, implying more optimism for gains and more pessimism for losses, when compared to the estimates for the regular present-value method. Figure 13 shows an example of the different weighting functions (estimated weighting function for gains for combination C5).

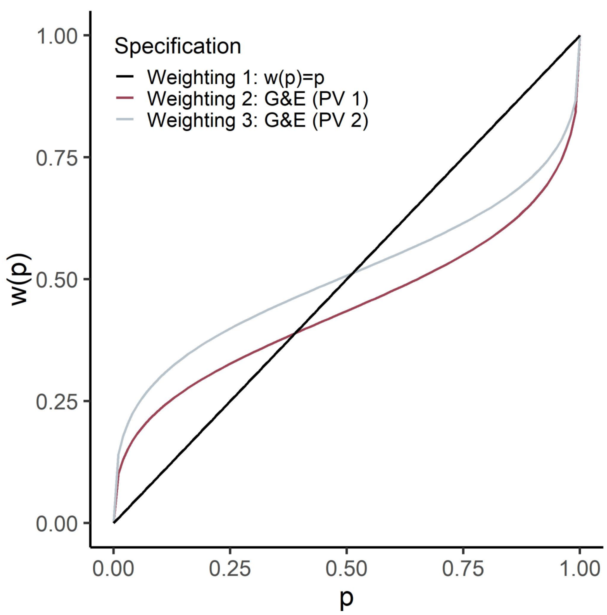


Figure 13: Probability-weighting function for gains for the regular (PV 1) and the monetary (PV 2) present-value methods

Notes: This figure shows estimated Goldstein-Einhorn probability weighting functions for gains for the present-value method (PV1) and the monetary present-value method (PV2) for function combination C5 (this function combination uses a power value function and exponential discounting.

Estimates of time discounting are somewhat lower for the monetary present-value method than for the regular one. For exponential discounting, estimated discounting per quarter reaches from zero to about 7%. For quasi-hyperbolic discounting, all future outcomes are estimated to be discounted by zero to about 12% (with very little additional exponential discounting).

We repeat the analysis of Section 5.3 for the monetary present value method to find out which components of PT are most important for the prediction performance. As for the regular present-value method, probability weighting is by far most important. Relying on linear utility for gains and losses (with or without loss aversion) makes no noticeable difference for prediction performance. The results can be found in Figure D.3 in the appendix.

5.5 Allowing for Non-zero Reference Points

We designed the experiment such that reference points of zero are natural. Here, we provide evidence that our results do not depend on this assumption (we conducted this analysis in response to feedback that we received after the experiment, it is thus not pre-registered). To do so, we conduct the analysis described in Sections 5.1 and 5.2 again, with the reference points of the three time periods as additional parameters to be estimated. The results are extremely similar to the results obtained when assuming reference points of zero. Figure 14 shows the comparison of the time-separation and present-value methods, similar to Figure 8 (the two figures are basically indistinguishable).

The results are almost identical, because estimated reference points are always very close to zero. Considering the reference points estimated on the full set of lotteries (considering those estimated on the calibration set would give a similar picture), all of the 72 reference points are in absolute terms smaller than one and a half euros, most of them are even smaller than half a euro (72 reference points, because there are three time periods, two application methods, and 12 function combinations). The estimated parameters can be found in Tables D.11 and D.12 in the appendix. These results confirm the interpretation in Section 5.3: the key model feature in explaining the valuation of intertemporal prospects is probability weighting, not utility curvature or loss aversion (with or without varying reference points).

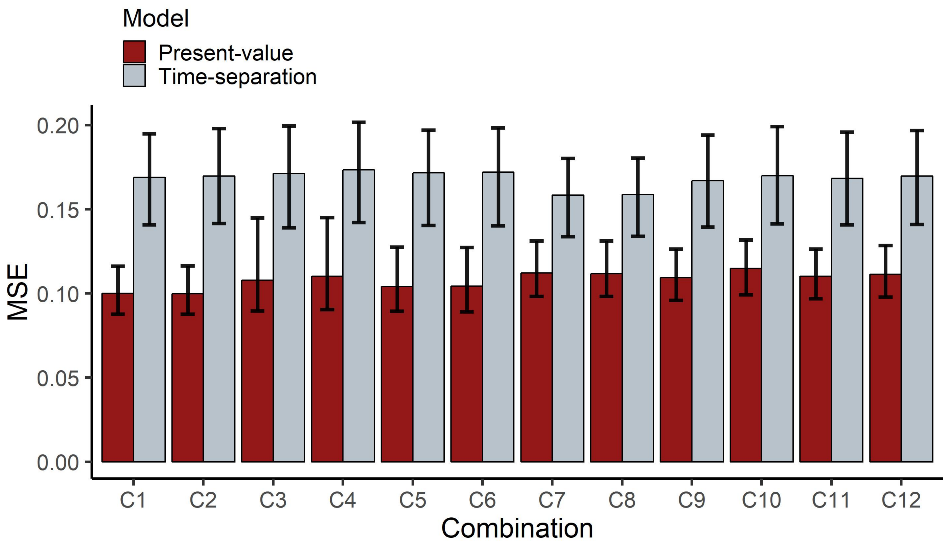


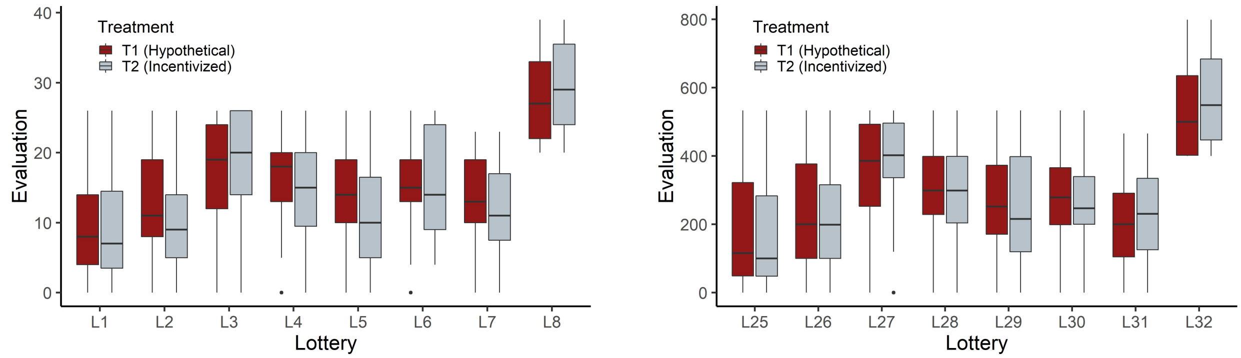
Figure 14: MSE for both application methods (all function combinations), estimated reference points

Notes: This figure shows mean squared errors of present-value and time-separation method on the test set for all 12 function combinations (with bootstrapped 95% confidence intervals).

5.6 Treatment Comparison

To test whether decisions differ between the two treatments, we proceed as specified in our pre-analysis plan. For each of the 16 gain lotteries, we test with exact tests for comparing means (Schlag, 2008) whether the means of the stated certainty equivalents between the two treatments differ. To counteract the problem of multiple testing, we correct the p-values using the Bonferroni-Holm method (as described in the pre-analysis plan). None of these tests are significant (using t-tests or Wilcoxon-Mann-Whitney tests instead also yields no significant differences; even without corrections for multiple testing none of the p-values for any single lottery are below 0.05.). Details on these test results can be found in Appendix D. Figure 15 contains boxplots of choices for all gain lotteries, showing that choices are very similar in both treatments.

These results confirm earlier findings in the literature. In the elicitation of risk and time preferences (situations in which lies by participants would neither increase their monetary rewards nor their social image or self image; these situations are naturally different from



(a) Set 1 (low-stake gain) (b) Set 4 (high-stake gain)

Figure 15: Evaluations (stated certainty equivalents) in T1 and T2

Notes: This figure shows boxplots of reported certainty equivalents for all gain lotteries, split up according to whether the decisions are hypothetical or with monetary incentives.

game-theoretic settings), it is predominantly found that there are no considerable differences between hypothetical and incentivized choices (e.g., Abdellaoui et al., 2011; Brañas Garza et al., 2020; Hackethal et al., 2020).

6 Conclusion

Our pre-registered experiment on a representative sample of the Dutch population tests the performance of applications of prospect theory in intertemporal settings. We find that the present-value method (aggregating first over time) performs much better than the time-separation method (aggregating first over risk). The method also performs much better than expected discounted utility theory.

While the choice of present-value method or time-separation method has a sizable impact on the explanatory power, the functional form of value, probability-weighting, and time-discounting functions has only a negligible effect on the explanatory power. We provide parameter estimations for twelve combinations of functional forms. There may be theoretical considerations for preferring some combinations over others. If one is only interested in the predictive power (possibly favoring a combination with few parameters), we recommend using the following specifications of a power utility and Tversky-Kahneman probability-weighting functions for gains and losses, combined with an exponential discounting of 12% per quarter: $v(x) = \mathbb{1}(x \geq 0)x^{1.20} - \mathbb{1}(x < 0)1.07(-x)^{1.20}$, and $w^{+}(p) = w^{-}(p) = \frac{p^{0.54}}{(p^{0.54}+(1-p)^{0.54})^{1/0.54}}$ (labeled combination C1 in Table 4 above).[13] As alternative, if one prefers to use a two-parametric probability-weighting function to allow for more flexibility (for theoretical arguments in favor of a two-parametric function, see Fehr-Duda and Epper, 2012), we recommend using the following specifications of a power utility and Goldstein-Einhorn weighting functions, combined with exponential discounting of 7% per quarter: $v(x) = \mathbb{1}(x \geq 0)x^{1.09} - \mathbb{1}(x < 0)1.04(-x)^{1.09}$, and $w^{+}(p) = w^{-}(p) = \frac{0.78p^{0.44}}{0.78p^{0.44}+(1-p)^{0.44}}$ (C5). This specification has a prediction performance that is almost as good as the specification mentioned earlier and the parameter values seem intuitively more natural and are more closely aligned with the atemporal literature (they are, for instance, very similar to the estimates of Bruhin et al., 2010).

We also compare these two application methods to a third method, which is similar to the present-value method, but which calculates present values in monetary terms before assigning them a utility value instead of calculating present values directly in utility terms. We consider this a natural way to apply prospect theory when the outcomes are monetary gains and losses (which they usually are; however, prospect theory can also be used with other outcomes, such as consumption levels). For an application of this method, we recommend using the following specifications of a power utility and Goldstein-Einhorn weighting functions, combined with exponential discounting of 4% per quarter: $v(x) = \mathbb{1}(x \geq 0)x^{0.92} - \mathbb{1}(x < 0)1.03(-x)^{0.92}$, and $w^{+}(p) = w^{-}(p) = \frac{1.04p^{0.40}}{1.04p^{0.40}+(1-p)^{0.40}}$ (C5). This specification has very high explanatory power and intuitive parameter estimates.

References

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.

[13]This is the combination with with the best out-of-sample explanatory power. It is also the combination with the lowest number of parameters. The calibration stems from the whole set of lotteries – in the exact calibration (Table 7), the parameters for the weighting functions for gains and losses are 0.535 and 0.551, respectively. As these values are so close, we believe that it is reasonable to use the same parameter in the gain and loss domains (the same argument for using identical specifications of weighting functions for gains and losses also applies to the other examples in this conclusion).

Abdellaoui, M. and Kemel, E. (2013). Eliciting prospect theory when consequences are measured in time units: "Time is not money", *Management Science* **60**(7): 1844–1859.

Abdellaoui, M., l'Haridon, O. and Paraschiv, C. (2011). Experienced vs. described uncertainty: Do we need two prospect theory specifications?, *Management Science* **57**(10): 1879–1895.

Abdellaoui, M., l'Haridon, O. and Paraschiv, C. (2013). Sign dependence in intertemporal choice, *Journal of Risk and Uncertainty* **47**(3): 225–253.

Andersen, S., Harrison, G. W., Lau, M. I. and Rutström, E. E. (2018). Multiattribute utility theory, intertemporal utility, and correlation aversion, *International Economic Review* **59**(2): 537–555.

Andreoni, J., Feldman, P. and Sprenger, C. (2017). A stream of prospects or a prospect of streams: On the evaluation of intertemporal risks. NBER Working Paper No. 24075.

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

Barberis, N. C. (2013). Thirty years of prospect theory in economics: A review and assessment, *Journal of Economic Perspectives* **27**(1): 173–96.

Barberis, N. and Huang, M. (2006). The loss aversion/narrow framing approach to the equity premium puzzle. NBER Working Paper No. 12378.

Barberis, N. and Huang, M. (2009). Preferences with frames: a new utility specification that allows for the framing of risks, *Journal of Economic Dynamics and Control* **33**(8): 1555–1576.

Barberis, N. and Xiong, W. (2009). What drives the disposition effect? An analysis of a long-standing preference-based explanation, *The Journal of Finance* **64**(2): 751–784.

Benartzi, S. and Thaler, R. H. (1995). Myopic loss aversion and the equity premium puzzle, *The Quarterly Journal of Economics* **110**(1): 73–92.

Berkelaar, A. B., Kouwenberg, R. and Post, T. (2004). Optimal portfolio choice under loss aversion, *Review of Economics and Statistics* **86**(4): 973–987.

Blavatskyy, P. R. (2016). A monotone model of intertemporal choice, *Economic Theory* **62**(4): 785–812.

Bleichrodt, H. and Eeckhoudt, L. (2006). Survival risks, intertemporal consumption, and insurance: The case of distorted probabilities, *Insurance: Mathematics and Economics* **38**(2): 335–346.

Booij, A. S., Van Praag, B. M. and Van De Kuilen, G. (2010). A parametric analysis of prospect theory's functionals for the general population, *Theory and Decision* **68**(1–2): 115–148.

Brañas Garza, P., Estepa Mohedano, L., Jorrat, D., Orozco, V. and Rascon-Ramirez, E. (2020). To pay or not to pay: Measuring risk preferences in lab and field. MPRA Working Paper No. 103088.

Brown, A. L., Imai, T., Vieider, F. and Camerer, C. (2021). Meta-analysis of empirical estimates of loss-aversion. CESifo Working Paper No. 8848.

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

Chew, S. H. and Epstein, L. G. (1990). Nonexpected utility preferences in a temporal framework with an application to consumption-savings behaviour, *Journal of Economic Theory* **50**(1): 54–81.

De Giorgi, E. G. and Legg, S. (2012). Dynamic portfolio choice and asset pricing with narrow framing and probability weighting, *Journal of Economic Dynamics and Control* **36**(7): 951–972.

DellaVigna, S. (2009). Psychology and economics: Evidence from the field, *Journal of Economic literature* **47**(2): 315–72.

Deng, L. and Pirvu, T. A. (2019). Multi-period investment strategies under cumulative prospect theory, *Journal of Risk and Financial Management* **12**(2): 83.

Dimmock, S. G. and Kouwenberg, R. (2010). Loss-aversion and household portfolio choice, *Journal of Empirical Finance* **17**(3): 441–459.

Dohmen, T. J., Falk, A., Huffman, D., Sunde, U., Schupp, J. and Wagner, G. G. (2005). Individual risk attitudes: New evidence from a large, representative, experimentally-validated survey. IZA Discussion Paper No. 1730.

Drouhin, N. (2015). A rank-dependent utility model of uncertain lifetime, *Journal of Economic Dynamics and Control* **53**: 208–224.

Easley, D. and Yang, L. (2015). Loss aversion, survival and asset prices, *Journal of Economic Theory* **160**: 494–516.

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.

Fehr-Duda, H. and Epper, T. (2012). Probability and risk: Foundations and economic implications of probability-dependent risk preferences, *Annual Review of Economics* **4**(1): 567–593.

Goette, L., Huffman, D. and Fehr, E. (2004). Loss aversion and labor supply, *Journal of the European Economic Association* **2**(2-3): 216–228.

Goldstein, W. M. and Einhorn, H. J. (1987). Expression theory and the preference reversal phenomena, *Psychological Review* **94**(2): 236.

Gomes, F. J. (2005). Portfolio choice and trading volume with loss-averse investors, *The Journal of Business* **78**(2): 675–706.

Guiso, L. and Paiella, M. (2008). Risk aversion, wealth, and background risk, *Journal of the European Economic Association* **6**(6): 1109–1150.

Guo, J. and He, X. D. (2017). Equilibrium asset pricing with Epstein-Zin and loss-averse investors, *Journal of Economic Dynamics and Control* **76**: 86–108.

Hackethal, A., Kirchler, M., Laudenbach, C., Razen, M. and Weber, A. (2020). On the (ir)relevance of monetary incentives in risk preference elicitation experiments. SAFE Working Paper No. 286.

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

Harrison, G. W. and Rutström, E. E. (2008). Risk aversion in the laboratory, *Research in Experimental Economics* **12**(8): 41–196.

Havránek, T., Irsova, Z. and Matousek, J. (2021). Individual discount rates: A meta-analysis of experimental evidence. CEPR Discussion Paper No. 15688.

Henderson, V. (2012). Prospect theory, liquidation, and the disposition effect, *Management Science* **58**(2): 445–460.

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

Hlouskova, J., Fortin, I. and Tsigaris, P. (2017). The consumption–investment decision of a prospect theory household: A two-period model, *Journal of Mathematical Economics* **70**: 74–89.

Köbberling, V. and Wakker, P. P. (2005). An index of loss aversion, *Journal of Economic Theory* **122**(1): 119–131.

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

Kyle, A. S., Ou-Yang, H. and Xiong, W. (2006). Prospect theory and liquidation decisions, *Journal of Economic Theory* **129**(1): 273–288.

Lampe, I. and Würtenberger, D. (2020). Loss aversion and the demand for index insurance, *Journal of Economic Behavior & Organization* **180**: 678–693.

Meng, J. and Weng, X. (2017). Can prospect theory explain the disposition effect? A new perspective on reference points, *Management Science* **64**(7): 3331–3351.

Odean, T. (1998). Are investors reluctant to realize their losses?, *The Journal of finance* **53**(5): 1775–1798.

Post, T., Van den Assem, M. J., Baltussen, G. and Thaler, R. H. (2008). Deal or no deal? Decision making under risk in a large-payoff game show, *American Economic Review* **98**(1): 38–71.

Prelec, D. et al. (1998). The probability weighting function, *Econometrica* **66**: 497–528.

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

Schlag, K. (2008). A new method for constructing exact tests without making any assumptions. Department of Economics and Business Working Paper No. 1109, Universitat Pompeu Fabra.

Stott, H. P. (2006). Cumulative prospect theory's functional menagerie, *Journal of Risk and Uncertainty* **32**(2): 101–130.

Sugden, R. (2003). Reference-dependent subjective expected utility, *Journal of Economic Theory* **111**(2): 172–191.

Sydnor, J. (2010). (Over)insuring modest risks, *American Economic Journal: Applied Economics* **2**(4): 177–99.

Tanaka, T., Camerer, C. F. and Nguyen, Q. (2010). Risk and time preferences: Linking experimental and household survey data from Vietnam, *American Economic Review* **100**(1): 557–71.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty* **5**(4): 297–323.

van Bilsen, S. and Laeven, R. J. (2020). Dynamic consumption and portfolio choice under prospect theory, *Insurance: Mathematics and Economics* **91**: 224–237.

van Bilsen, S., Laeven, R. J. and Nijman, T. E. (2020). Consumption and portfolio choice under loss aversion and endogenous updating of the reference level, *Management Science* **66**: 3799–4358.

Wakker, P., Thaler, R. and Tversky, A. (1997). Probabilistic insurance, *Journal of Risk and Uncertainty* **15**(1): 7–28.

Weber, M. and Camerer, C. F. (1998). The disposition effect in securities trading: An experimental analysis, *Journal of Economic Behavior & Organization* **33**(2): 167–184.

Online Appendix to "Intertemporal Prospect Theory"

Immanuel Lampe Matthias Weber

A Differences in Evaluation and Rank-order Stability

This section provides our formal definition of rank-order stability and a proof of a proposition that summarizes the relationship between rank-order stability and identical evaluations of the two application methods.

We start with the formal definition of rank-order stability.

Definition Rank-order-stability: Consider an intertemporal prospect with potential payouts at T different points in time. Denote by $X_{\tilde{t}|x_t=x_{t,k_t}}$ the set of outcomes that may arrive in period $\tilde{t} \leq T-1$ given that the outcome in period $t \leq T-1$ is x_{t,k_t}. An intertemporal prospect is called rank-order stable if for any two outcomes x_{t,j_t} and x_{t,k_t} of any period $t \leq T-1$ with $x_{t,j_t} > x_{t,k_t}$ it holds that $\min(X_{\tilde{t}|x_t=x_{t,j_t}}) \geq \max(X_{\tilde{t}|x_t=x_{t,k_t}})$ for all $\tilde{t} \leq T-1$.

We next state three lemmas that are required to proof the core proposition of this section.

Lemma 1: Denote by $o_j = (o_{j,0}, ..., o_{j,T-1})$ and $o_k = (o_{k,0}, ..., o_{k,T-1})$ two outcome streams. Rank-order stability implies that if for any period t the outcome of stream j is larger than the outcome of stream k, $o_{j,t} > o_{k,t}$, then the present-value of stream j is larger then the present value of stream k, $PV(o_j) > PV(o_k)$.

Proof: Under Rank-order stability $o_{j,t} > o_{k,t}$ implies $o_{j,\tilde{t}} \geq o_{k,\tilde{t}}$ for all $\tilde{t} \neq t$ and therefore $PV(o_j) > PV(o_k)$.

□

Lemma 2: Denote by $o_1, o_2, ...$ the outcome streams and by $PV_1 > ... > PV_n > 0$ their ordered distinct present values. in addition, denote by $q_1, ..., q_n$ the associated probabilities. The present-value method transforms the objective probabilities $q_1, ..., q_n$ into subjective decision weights $\pi_1, ..., \pi_n$ employing a probability weighting function w^+. The following calculation rules hold for these subjective decision weights:

- A1: $\pi_1 + ... + \pi_k = w^+(q_1 + ... + q_k)$, for all $k \leq n$.
- A2: $\pi_k + ... + \pi_j = w^+(q_1 + ... + q_j) - w^+(q_1 + ... + q_{k-1})$, for all $k > 1$ and $j > k$.

Proof:

$$\begin{aligned}
\pi_1 + \pi_2 + \pi_3 + ... + \pi_k &= w^+(q_1) + [w^+(q_1 + q_2) - w(q_1)] \\
&+ [w^+(q_1 + q_2 + q_3) - w^+(q_1 + q_2)] + ... \\
&+ [w^+(q_1 + ... + q_k) - w^+(q_1 + ... + q_{k-1})] \\
&= w^+(q_1) - w^+(q_1) + w^+(q_1 + q_2) - w^+(q_1 + q_2) + ... + w^+(q_1 + ... + q_k) \\
&= 0 + ... + 0 + w^+(q_1 + ... + q_k) = w^+(q_1 + ... + q_k)
\end{aligned}$$

$$\begin{aligned}
\pi_k + \pi_{k+1} + ... + \pi_j &= [w^+(q_1 + ... + q_k) - w^+(q_1 + ... + q_{k-1})] \\
&+ [w^+(q_1 + ... + q_{k+1}) - w^+(q_1 + ... + q_k)] + ... \\
&+ [w^+(q_1 + ... + q_j) - w^+(q_1 + ... + q_{j-1})] \\
&= -w^+(q_1 + ... + q_{k-1}) + w^+(q_1 + ... + q_k) - w^+(q_1 + ... + q_k) + ... + w^+(q_1 + ... + q_j) \\
&= -w^+(q_1 + ... + q_{k-1}) + 0 + ... + 0 + w^+(q_1 + ... + q_j) \\
&= w^+(q_1 + ... + q_j) - w^+(q_1 + ... + q_{k-1})
\end{aligned}$$

□

Lemma 3: Both application methods imply an evaluation of the form $\sum_{t=0}^{T-1} \delta(t) \sum_{i=1}^{k_t} \pi_{t,i} v(x_{t,i})$.

Proof: The evaluation of an intertemporal prospect by the time-separation method is given by

$$V = \sum_{t=0}^{T-1} \delta(t) V_t = \sum_{t=0}^{T-1} \delta(t) \sum_{i=1}^{k_t} \pi_{t,i}^{TS} v(x_{t,i}).$$

The outcome $x_{t,i}$ thus, receives the weight $\pi_{t,i}^{TS}$.

The present value method assigns decision weights to outcome streams. Suppose the outcome stream o_j with present value $\sum_{t=1}^{T}\delta(t)v(o_{j,t})$ is assigned the decision weight π_j^{PV}. This implies that each outcome of the outcome stream is assigned the weight π_j^{PV} ($\pi_j^{PV}\sum_{t=0}^{T-1}\delta(t)v(o_{j,t}) = \sum_{t=0}^{T-1}\delta(t)\pi_j^{PV}v(o_{j,t})$). Multiple outcome streams may yield the outcome $x_{t,i}$ in period t. The total weight outcome $x_{t,i}$ receives under the present-value method is, thus, the sum of the weights that the streams containing this outcome receive. Denote the this sum by $\pi_{t,i}^{PV}$. If, for instance, the only two streams that contain outcome $x_{t,i}$ are stream o_j and o_k, then $\pi_{t,i}^{PV} = \pi_j^{PV} + \pi_k^{PV}$, with π_j^{PV} and π_k^{PV} denoting the weight assigned to stream o_j and o_k. The evaluation of an intertemporal prospect as suggested by the present-value method can therefore be written as

$$W = \sum_{j=1}^{k}\pi_j PV_j = \sum_{t=0}^{T-1}\delta(t)\sum_{i=1}^{k_t}\pi_{t,i}^{PV}v(x_{t,i}).$$

□

Comparing the two evaluations V and W from the proof above, it is obvious that if $\pi_{t,i}^{TS} = \pi_{t,i}^{PV}$ holds for all outcomes $x_{t,i}$, then the two methods yield the same evaluation. The proposition below summarizes the relationship between rank-order stability and identical evaluations.

Proposition 1: If each outcome stream of an intertemporal prospect contains only non-negative or non-positive outcomes and the prospect is rank-order stable, then the evaluations under the two application methods are identical.

Proof: The proof contains two steps.

Step 1: We start the proof by assuming that the prospect only contains non-negative payouts. Denote by $x_{t,1}$ the best outcome of period t and by $p_{t,1}$ its objective probability. The assigned weight under the time-separation method is

$$\pi_{t,1}^{TS} = w^{+}(p_{t,1}).$$

As displayed in Table A.1, denote by $o_1 ... o_{n_{t,1}}$ the $n_{t,1}$ outcome streams that contain the outcome $x_{t,1}$ and by $q_1 ... q_{n_{t,1}}$ their objective probabilities. Note that $q_1 + ... + q_{n_{t,1}} = p_{t,1}$. Lemma 1 implies that $o_1, ..., o_{n_1}$ are placed in the first $n_{t,1}$ positions if streams are ordered by their present values. Using Lemma 2 A2, the sum of the weights assigned to these streams is therefore equal to

$$w^{+}(q_1 + \ldots + q_{n_{t,1}}) = w^{+}(p_{t,1}).$$

The outcome $x_{t,1}$ is therefore multiplied by the same factor under both methods.

Table A.1: Intertemporal prospect (for proof)

Stream	Probability	Payout in period t
o_1	q_1	$x_{t,1}$
o_2	q_2	$x_{t,1}$
$\vdots$	$\vdots$	$\vdots$
$o_{n_{t,1}}$	$q_{n_{t,1}}$	$x_{t,1}$
$o_{n_{t,1}+1}$	$q_{n_{t,1}+1}$	$x_{t,2}$
o_2	q_2	$x_{t,2}$
$\vdots$	$\vdots$	$\vdots$
$o_{n_{t,1}+n_{t,2}}$	$q_{n_{t,1}+n_{t,2}}$	$x_{t,2}$
$\vdots$	$\vdots$	$\vdots$

Denote by $x_{t,2}$ the second best outcome of period t and by $p_{t,2}$ its objective probability. The assigned weight under the time-separation method is

$$\pi^{TS}_{t,2} = w^{+}(p_{t,1} + p_{t,2}) - w^{+}(p_{t,1}).$$

As displayed in Table A.1 denote by $o_{n_{t,1}+1} \ldots o_{n_{t,1}+n_{t,2}}$ the $n_{t,2}$ outcome streams that contain the outcome $x_{t,2}$ and by $q_{n_{t,1}+1}$ ... $q_{n_{t,1}+n_{t,2}}$ their objective probabilities. Note that $q_{n_{t,1}+1} + \ldots + q_{n_{t,1}+n_{t,2}} = p_{t,2}$. Lemma 1 implies that $o_{n_{t,1}+1} \ldots o_{n_{t,1}+n_{t,2}}$ are placed in the $n_{t,1} + 1$ to $n_{t,1} + n_{t,2}$ positions if streams are ordered by their present values. Using Lemma 2 A2 the sum of the decision weights assigned to these streams is therefore given by

$$w^{+}(q_1 + \ldots + q_{n_{t,1}+n_{t,2}}) - w^{+}(q_1 + \ldots + q_{n_{t,1}}) = w^{+}(p_{t,1} + p_{t,2}) - w^{+}(p_{t,1})$$

The outcome $x_{t,2}$ is therefore multiplied by the same factor under both methods.

The same reasoning as above can also be applied to the other outcomes of period t and also to outcomes of other periods. It therefore holds that $\pi^{PV}_{t,i} = \pi^{TS}_{t,i}$ for all $i \le k_t$ and $t \le T - 1$.

Because both methods assign the same weights, the evaluations are identical.

Step 2: The same reasoning as in Step 1 can also be applied to prospects that only contain negative payouts. For mixed prospects whose outcome streams either contain only non-negative payouts or non-positive payouts the evaluation can be separated into two parts. The evaluation of the non-negative part yields V^+ under the time-separation method and W^+ under the present-value method. The evaluation of the non-positive part yields V^- and W^-. With the same reasoning as above $V^+ = W^+$ and $V^- = W^-$ holds and, thus, also $V^+ + V^- = W^+ + W^-$.

$\square$

B Lotteries in the Experiment

This section summarizes the lotteries of Set 1 (low-stake gain lotteries) and Set 3 (low-stake mixed lotteries). The loss lotteries (Sets 2 and 5) can be obtained by multiplying all outcomes of the gain lotteries (Sets 1 and 4) by −1. The high-stake lotteries (Sets 4 to 6) can be obtained by multiplying the low-stake lotteries (Sets 1 to 3) by 20.

Figure B.1 displays the 8 lotteries of Set 1. All outcomes are in euros. Lotteries 1 to 6 are calibration lotteries, lotteries 7 and 8 test lotteries. The outcome of Lottery 1 at $t = 0$ is either 10 with a low probability or 0 with a high probability. An arrival of the good outcome at $t = 0$ implies the arrival of good outcomes at $t = 1$ (20 with certainty as compared to 10 with certainty) and also at $t = 2$ (50 and 20 equally likely compared to 0 with certainty). Lottery 1, hence, contains two unlikely good outcome streams. Lottery 2 is identical to Lottery 1 except that probabilities at $t = 0$ change. For Lotteries 3 and 4 an arrival of the good outcome at $t = 0$ again implies the arrival of good outcomes at $t = 1$ and also at $t = 2$. Note that the good outcome streams are now likely to arrive. Lottery 5 is a variation of Lottery 1 in which the order of the first two periods and also probabilities at $t = 0$ change. Lottery 6 is a variation of Lottery 5 in which the uncertainty is shifted to the lower (bad) outcome stream. In general, the large number of low-stake outcomes of lotteries 1,2,3,4,5 and 6 ensures that the value function is estimated on a dense grid of small positive values. The large number of distinct outcomes and the variation in probabilities ensures that the probability weighting function is estimated on a dense grid spanning the whole domain.

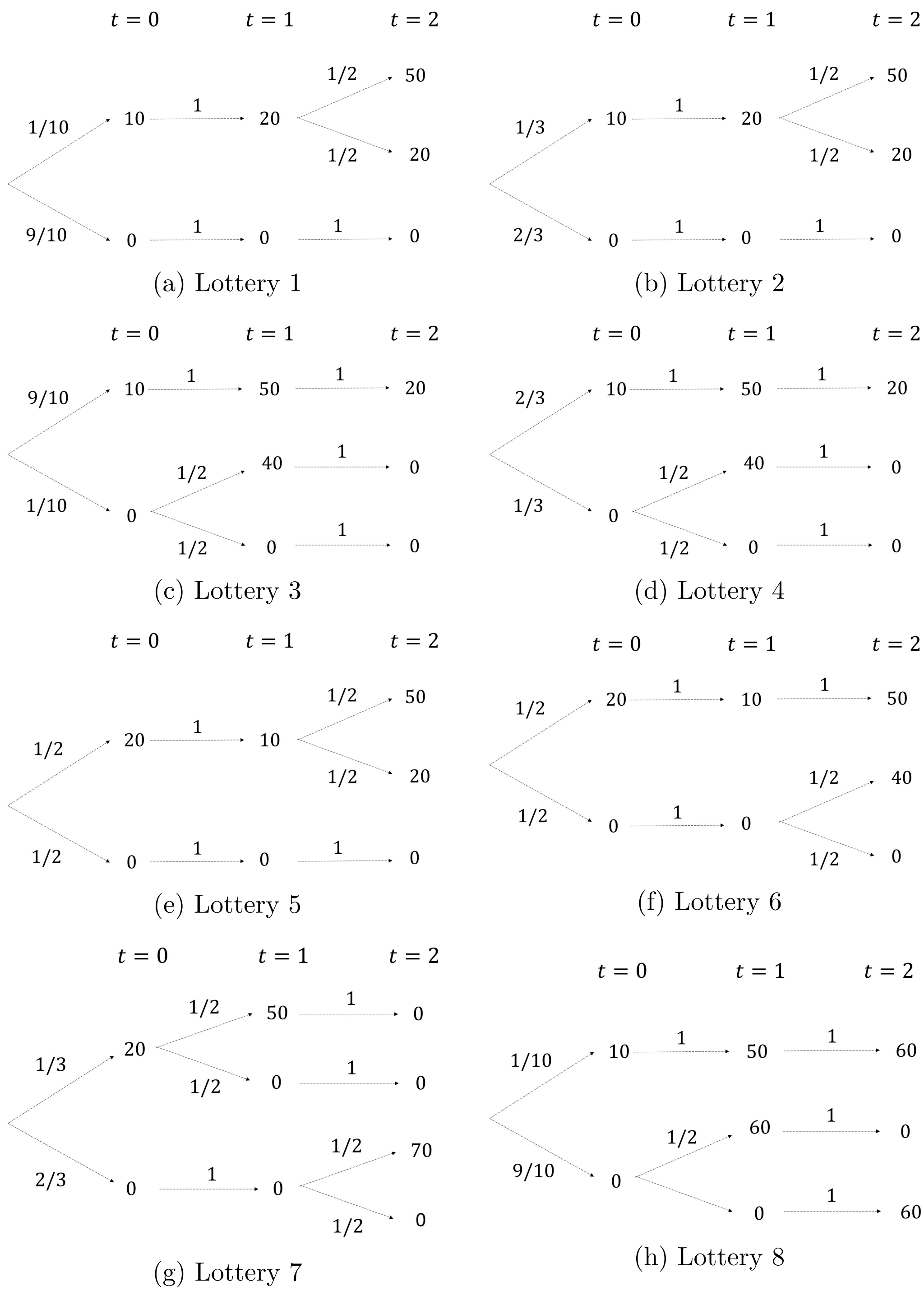
For the first test lottery, Lottery 7, note that at $t = 0$ and also at $t = 2$ the good outcome (20 at $t = 0$ and 70 at $t = 2$) arrives with probability 1/3. With a probability weighting function as found in atemporal prospect theory, the time-separation method would assess these probabilities approximately correct (common specifications of the weighting function suggest that a probability of 1/3 is assessed approximately correct). At $t = 1$, the good outcome 50 arrives with probability 1/6. With a typical atemporal weighting function, the time-separation method would overweight the good outcome 50 and therefore overestimate the lottery. This is different for the present-value method. The present values of the two best outcome streams $(20, 50, 0)$ and $(0, 0, 70)$ should be similar. The probability that one of these two outcome streams arrives is 1/2. With a typical atemporal weighting function, the present-value method would underweight these two streams and therefore underestimate the lottery.

For the second test lottery, Lottery 8, note that at $t = 1$ and also at $t = 2$ the outcome 60 arrives with a probability close or equal to 1/2. With a typical atemporal weighting function, the time-separation method would underweight the good outcome 60 and, thus, underestimate the lottery. For the present-value method note that the best outcome stream $(10, 50, 60)$ arrives with probability 1/10. With a typical atemporal weighting function, the present-value method would overweight the best outcome stream and therefore overestimate the lottery.

The lotteries of Set 3 are displayed in Figure B.2. Lotteries 17 to 22 are calibration lotteries, lotteries 23 and 24 are test lotteries. For the calibration lotteries note that these are replicates of the calibration lotteries of Set 1 (Lotteries 1 to 6 as displayed in Figure B.1) except that outcomes change. The upper half of each lottery exclusively consists of positive outcomes and the lower half exclusively consists of negative outcomes. The arrival of positive outcomes is unlikely in Lottery 17 and 18. Lottery 19 and 20 are variations of Lottery 17 and 18 in which outcomes are multiplied by -1 and the order of period 2 and 3 changes. Lottery 21 is a variation of Lottery 17 in which the order of the first two periods as well as the probabilities at $t = 0$ change. Lottery 22 is a variation of Lottery 21 in which outcomes are multiplied by -1. These six calibration lotteries again ensure that the value function is estimated on a dense grid of small values and that the probability weighting function is estimated on a dense grid spanning the whole domain.

For Lottery 23 note that at $t = 1$ the outcome -30 arrives with probability 1/2; and at $t = 2$ the outcome 60 arrives with probability 1/6. The time-separation method would underweight the negative outcome -30 and overweight the positive outcome 60, and therefore overestimate the lottery. The situation is different for the present-value method. The lottery contains three positive outcome streams: $(60, -30, 60)$, $(60, 0, 30)$, $(30, 60, 0)$. The present value of these streams should be close to each other. The probability that one of these three streams arrives is 2/3. The present-value method would therefore underweight these positive streams and, thus, underestimate the lottery.

For Lottery 24, note that at $t = 1$ and also at $t = 2$ the outcome 60 arrives with probability 1/2. The time-separation method would underweight the positive outcome 60 and therefore underestimate the lottery. For the present-value method note the outcome stream $(60, 60, 60)$



(a) Lottery 1 (b) Lottery 2

(c) Lottery 3 (d) Lottery 4

(e) Lottery 5 (f) Lottery 6

(g) Lottery 7 (h) Lottery 8

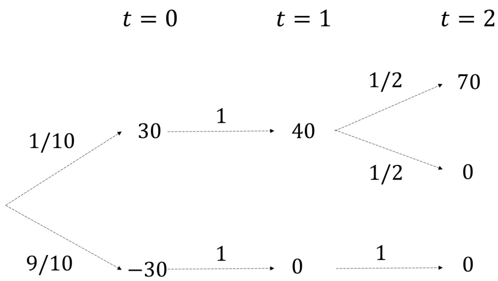
Figure B.1: Lottery Set 1 (low-stake gain lotteries) arrives with probability 1/6. The present-value method therefore overweights this stream and, hence, overestimates the lottery.

In general, the test lotteries are set up so that the evaluations under the two methods give (sizably) different results. In order to assure a fair comparison between the two methods we choose the test lotteries such that the degree to which one method over- and the other method underestimates the lottery are similar (assuming typical atemporal calibrations). Table B.1 summarizes these features of out test lotteries.

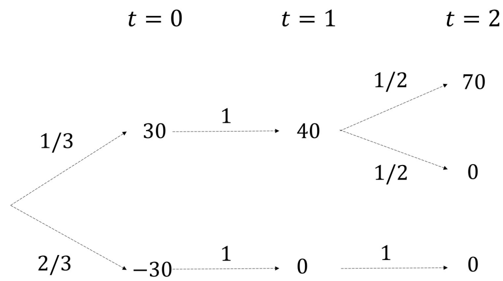
Table B.1: Overview test lotteries (12 Lotteries)

	No weighting	Time-separation		Present-value	
	CE	CE	Rel. deviation	CE	Rel. deviation
	(1)	(2)	(3)	(4)	(5)
Lottery 7	10.6	12	0.13	9.2	-0.13
Lottery 8	19.3	17	-0.09	20.8	0.08
Lottery 15	211.2	240.4	0.14	184.3	-0.13
Lottery 16	386.1	340.2	-0.1	416.6	0.08
Lottery 23	-10.6	-12	-0.13	-9.2	0.13
Lottery 24	-19.3	-17	0.09	-20.8	-0.08
Lottery 31	-211.2	-240.4	-0.14	-184.3	0.13
Lottery 32	-386.1	-340.2	0.12	-416.6	-0.08
Lottery 39	11	13.3	0.21	8.6	-0.22
Lottery 40	9.8	7.1	-0.28	12.4	0.27
Lottery 47	220.3	265.4	0.2	171.8	-0.22
Lottery 48	195.9	141.1	-0.28	248.8	0.27

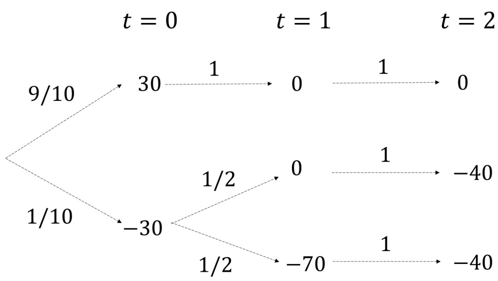
Notes: To calculate the certainty equivalents we use the parametric specifications suggested by Tversky and Kahneman (1992) i.e, $v(x) = \mathbb{1}(x \geq 0)x^{0.88} - \mathbb{1}(x < 0)2.25(-x)^{0.88}$, and $w^{+}(p) = w^{-}(p) = \frac{p^{0.69}}{(p^{0.69}+(1-p)^{0.69})^{1/0.69}}$ (without time discounting).



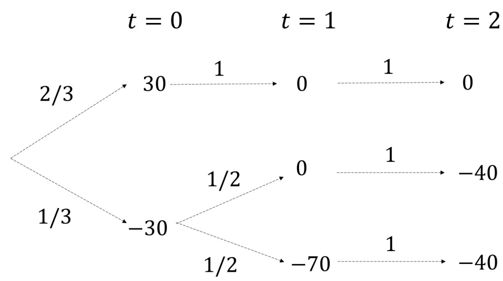
(a) Lottery 17



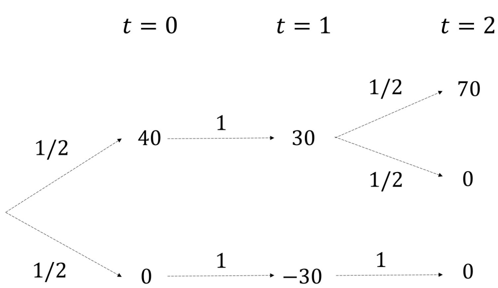
(b) Lottery 18



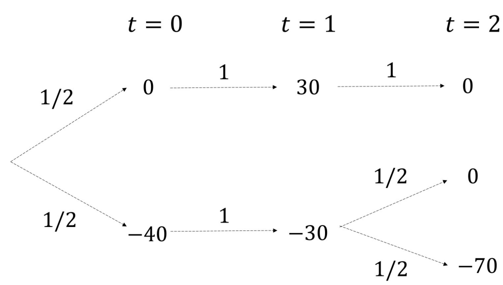
(c) Lottery 19



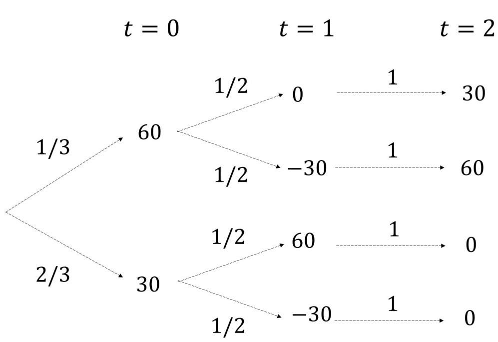
(d) Lottery 20



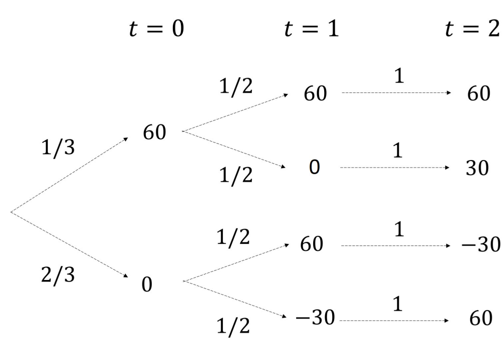
(e) Lottery 21



(f) Lottery 22



(g) Lottery 23



(h) Lottery 24

Figure B.2: Lottery Set 3 (low-stake mixed lotteries)

C Experimental Instructions and Comprehension Test Question

In Section C.1, we reproduce the complete experimental instructions that subjects receive before the experiment. Section C.2 contains the comprehension test questions that participants have to answer correctly before starting with the decision tasks. Section C.3 contains the information that participants receive during the experiment before starting each of the six blocks of decision tasks (that appear in random order). Section C.4 contains the brief post-experimental questionnaire. All texts are in English translation.

We provide the information for both treatments jointly. If nothing is indicated, texts and figures are shown in both treatments. Where there is a difference, this is indicated by highlighted brackets. The first part in the brackets applies to the hypothetical treatment T1, the latter part to the incentivized treatment T2. For instance, [example text / different example text] signifies that participants in T1 see "example text", while participants in T2 see "different example text". [NO TEXT / example text] signifies that participants in T1 do not see any text for this part, while participants in T2 see "example text".

C.1 Experimental Instructions

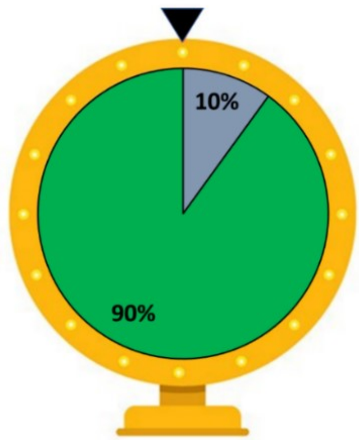
The aim of this research is to better understand how people make decisions. The research is carried out by researchers from the University of St. Gallen in Switzerland.

The experiment consists of 6 parts. In each part you will complete eight decision tasks, thus 48 decision tasks in total. Please read the following explanation carefully.

[The decisions you make in this questionnaire do not affect your payments for this questionnaire. Please make your decisions as if they involved real monetary gains or losses. / In addition to the regular payment for completing this questionnaire, there will be an additional payment depending on your decisions. This will be explained later. Not all decisions will affect your earnings, but please make all decisions as if they involved real monetary gains or losses.]

General Instructions

Probabilities (expressed in percentages) play an important role in this research. We will use a wheel of fortune to display probabilities.



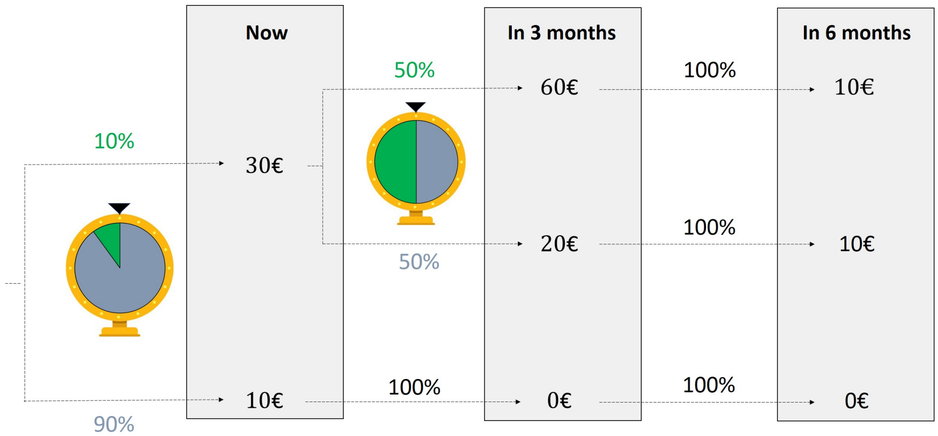
The probability that the black indicator (on the top of the wheel) points at a grey part after spinning the wheel is equal to 10%, because 1/10 (one tenth) of the wheel is grey. The probability that the black indicator points at a green part after spinning the wheel is thus equal to 90%, because 9/10 (nine tenths) of the wheel are green.

Time also plays an important role in this research. As displayed by the grey areas in the table below, payments can be made ”now”, ”in 3 months” or ”in 6 months”.

Now (Time of payment for participation)			**In 3 months**			**In 6 months**
First payout from your choice			Second payout from your choice			Third payout from your choice

Each decision task will involve a *safe option* and a *risky option*. A risky option yields monetary payouts with certain probabilities. Each risky option may have payouts now, in 3 months and in 6 months. As will be explained below, the payout in 3 months depends on the payout now. The payout in 6 months depends on the payout in 3 months.

We explain how the risky option works with the example below. In the chart, you see two wheels of fortune that show illustrate the probabilities.



The payout now is either 30 euro or 10 euro. The probability that 30 euro are paid is 10% and the probability that 10 euro are paid is 90%, as the arrows in the first block (Now) show.

In addition to the payout now, the risky option yields a payout in 3 months and a payout in 6 months. The additional payout in 3 months depends on the payout now and is 60 euro, 20 euro, or 0 euro.

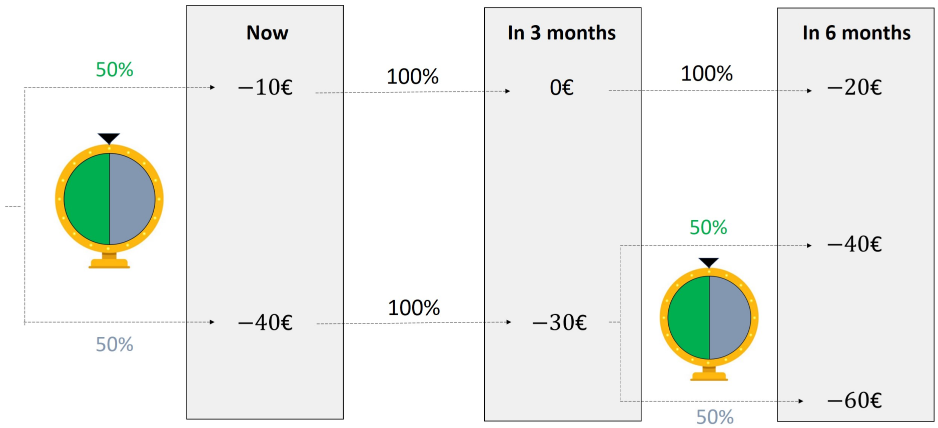
- If the payout now is 30 euro, then the payout in 3 months is 60 euro with probability 50% and 20 euro with probability 50%, as indicated by the arrows in the second block (in 3 months).
- If the payout now is 10 euro, then the payout in 3 months is 0 euro, as the arrow between the 10 euro in the first block (now) and the 0 euro in the second block (in 3 months) shows.

The additional payout in 6 months is either 10 euro or 0 euro depending on the payout in 3 months.

- If the payout in 3 months is 20 euro or 60 euro, then the payout in 6 months is 10 euro with certainty.
- If the payout in 3 months is 0 euro, then the payout in 6 months is also 0 euro.

Here is another example. Note that there are only negative payouts in this example.

For negative payouts you should imagine that you have to pay money to the researcher, thus that you lose money. In this example, in the block ?now?, is the probability 50% that the payout is -10 euro and the probability is 50% that the payout is -40 euro. The additional

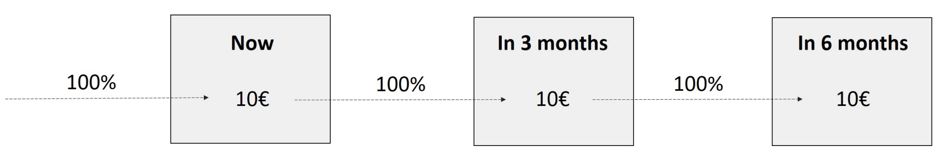


payout in 3 months is either 0 euro or -30 euro and depends on the payout now. If, for example, the payout now is -10 euro, then the payout in 3 months is 0 euro. The additional payout in 6 months is then -20 euro, -40 euro, or -60 euro and depends on the payout in the second block (in 3 months). If the payout there is -30 euro, then the probability is 50% that the payout in 6 months is -40 euro and 50% that the payout is -60 euro.

The two examples that you saw gave only positive or only negative payouts. In the research, there is also a third type of risky option, which contains a mix of positive and negative payouts.

Decisions

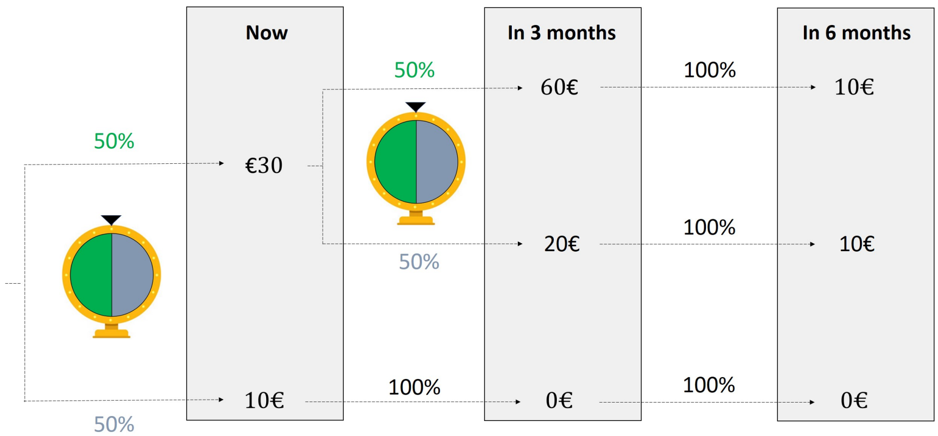
In each decision task there is one risky option. You can choose between the risky option and a safe option. Both options (risky or safe) yield payouts at three different points in time (now, in 3 months, and in 6 months). In contrast to the payouts of the risky option, the payouts of the safe option are **certain** and they are **always the same**. As displayed in the figure below, the safe option can, for example, yield a certain payout sequence of 10 euro now, 10 euro in three months, and 10 euro in six months.



Each decision screen contains three elements. The first element is a chart that displays the payout structure of the risky option. The second element is a chart that displays the payouts of the safe option. The third element is a "slider" that you can position in order to indicate for which amounts paid in the safe option you would rather have the risky option than the safe option. An example of the three elements together is displayed below.

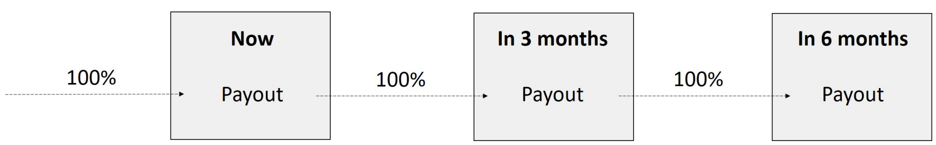
<u>Element 1</u>

Consider the risky option for this decision task.



<u>Element 2</u>

The alternative to the risky option is a safe option that yields three certain and identical payouts:



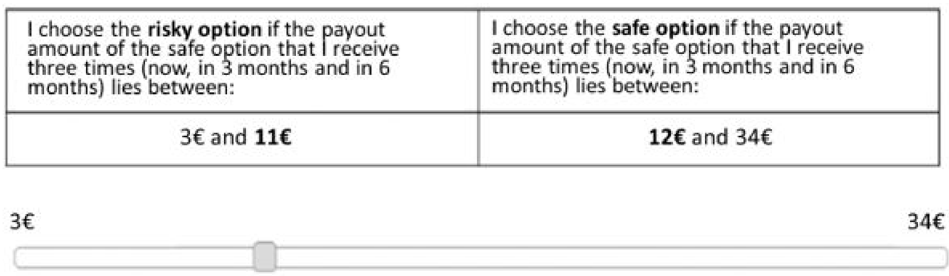
<u>Element 3</u>

Below you can see the slider that you can move.

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:
3€ and **3€**	**4€** and 34€

You can move the slider to the left and to the right to indicate your choice. There are no right or wrong choices: the position of the slider indicates what you prefer and preferences may be different for different people.

Example: Assume that you prefer the risky option as long as the certain amount of the safe option that you would receive three times (once now, once in 3 months, once in 6 months) is 11 euro or less. However, if the certain amount of the safe option that you would receive three times is 12 euro or more (thus in total 36 euro or more), you prefer the safe option. Then you have to position the slider as displayed below.



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:
3€ and **11€**	**12€** and 34€

In this example, the risky option only contains **positive** payouts. But you will also see risky options that only contain **negative** payouts (you do not receive money, but you have to pay) and risky options that contain a mix of positive and negative payouts

Imagine: You consider a risky option with only negative payouts the lesser evil when the amount in the safe option that you would lose three times (once now, once in 3 months, once in 6 months) is 22 euro or more (thus in total 66 euro or more). You, however, perceive the risky option as the greater evil when the amount that you would lose three times in the safe option is 21 euro or less (which means that you would lose in total at most 63 euro). Then you have to position the slider as shown here:

Thus, each decision task thus has two options to choose from:

- a risky option where the payouts at three different points in time (now, in 3 months, and in 6 months) are uncertain.
- a safe option that yields three payouts (now, in 3 months and in 6 months) that are the same.

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:
-30€ and **-22€**	**-21€** and 0€

You are asked to position a slider to indicate for which amounts that you receive three times in the safe option, you would prefer the risky option over the safe option.

[NO TEXT / **Payments**

You receive a compensation for completing this questionnaire. But you can also earn an extra payment, depending on your choices and on luck. This extra payment is calculated as follows. The research consists of six parts. In each part, there are eight decision tasks, thus 48 in total. In two of the six parts, there are risky options that only contain positive payouts. The eight risky options in one of these two parts yield small to medium payouts, whereas the eight risky options in the other part yield medium to high payouts. One of these 16 decision tasks will be randomly selected for payment at the end of the questionnaire. There is a 1% chance that the selected decision task is from the part with medium to large payouts, and there is a 99% chance that the selected decision task is from the part with small to medium payouts. The decisions you make in the other four parts cannot be selected for payment.

In the example below, we describe how the payment for the randomly selected decision task is calculated. Assume that you positioned the slider in the selected decision task as displayed below:

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:
3€ and **11€**	**12€** and 34€

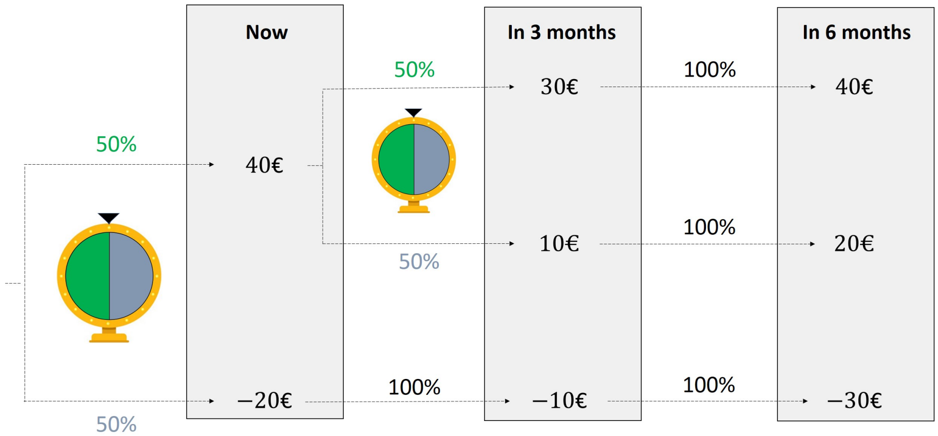
The computer then randomly selects an amount between the lower and upper limits of the slider, in this example thus between 3 euro and 33 euro. All amounts on the slider are equally likely to be selected. You will then receive the option that you chose for the randomly selected amount. In this example:

- If the randomly selected amount is 12 euro or more, the slider shows that you prefer receiving this amount three times (now, in 3 months, in 6 months) over the three uncertain payouts of the risky option. If, for example, the randomly selected amount is 13 euro, you receive 13 euro now, 13 euro in 3 months, and 13 euro in 6 months.
- If the randomly selected amount 11 euro less, the slider shows that you prefer the risky option. Your payment will then be determined through the risky option. The computer will then randomly determine the outcomes, and you will be told the results at the end of the questionnaire. Also here, the payouts will take place now, in 3 months, and in 6 months.]

Before the first part of the research starts, we would like to ask you a few test questions to make sure that you fully understand the instructions.

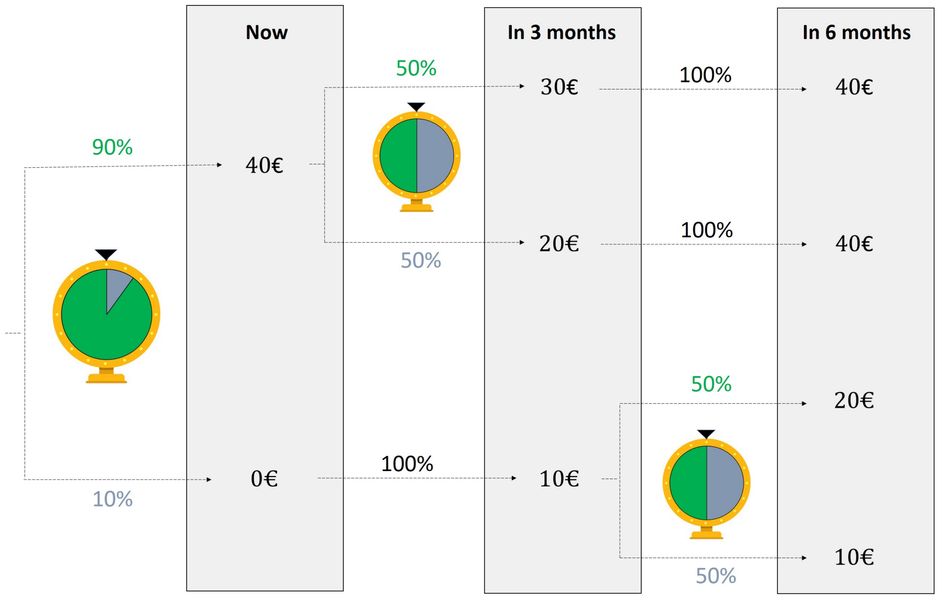
Please click "Next" to proceed to the questions.

C.2 Comprehension Test Questions



Question 1: Consider the risky option above and indicate for each statement whether it is true or false.

- The payout now is either 40 euro with probability 50% or -20 euro with probability 50%.
- If the payout now is -20 euro (you lose of 20 euro), then the payout in 3 months is -10 euro (you lose 10 euro) with certainty, and the payout in 6 months is -30 euro (you lose 30 euro) with certainty.



Question 2: Consider the risky option above and indicate for each statement whether it is true or false.

- The payout now is 0 euro with probability 10%. If the payout now is 0 euro, then the payout in 3 months is 10 euro. If the payout is 10 euro in three months, then the payout in 6 months is 20 euro with probability 50%. The probability that the payouts are 0 euro now, 10 euro in 3 months, and 20 euro in six months is, thus, 5%, because 50% of 10% is 5%.
- The payout in three months can only be 30 euro if the payout now is 40 euro. The payout now is 40 euro with probability 90%. After a payout of 40 euro now, the payout in 3 months is 30 euro with probability 50%. The payout in three months is, thus, 30 euro with probability 45%, because 50% of 90% is 45%.

[

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:
3€ and **22€**	**23€** and 34€

Question 3: Consider the slider displayed above and indicate whether the following statement is true or false.

- The position of the slider shows that if you are offered certain payouts of three times 20 euro (20 euro now, 20 euro in 3 months and 20 euro in 6 months), then you would reject these certain payouts and instead prefer the risky option.

/

Part of your payment for the decision tasks depends on luck and your decisions in two of the six parts of the questionnaire. For this, only one decision task will be randomly selected. Assume that you positioned the slider in the randomly selected decision task as shown here:

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:
3€ and **22€**	**23€** and 34€

The computer randomly selects an amount between 3 euro and 33 euro. All amounts are equally likely to be selected. Assume that the randomly selected amount is 20 euro.

Question 3: Please indicate whether the following statement is true or false.

- The position of the slider shows that if you are offered certain payouts of three times 20 euro (20 euro now, 20 euro in 3 months and 20 euro in 6 months), then you would reject these certain payouts and instead choose the risky option. The risky option will then determine the payments. The payouts take place now, in 3 months, and in 6 months. The money will be thus be paid at three different points in time.

]

C.3 Information Before Each Block of Decision Tasks During the Experiment

Information before starting the decision tasks of block 1:

In the eight decisions tasks of this part you will see risky options that yield small-to-medium positive payouts (gains).

[NO TEXT / **As explained earlier, only one decision task will be randomly selected for payment at the end of the questionnaire. There is a 99% chance that one of the eight decision tasks from this part will be selected for payment.**]

Please click on the "Next" button to proceed to the first decision task of this part.

Information before starting the decision tasks of block 2:

In the eight decisions tasks of this part you will see risky options that yield medium-to-large positive payouts (gains).

[NO TEXT / **As explained earlier, only one decision task will be randomly selected for payment at the end of the questionnaire. There is a 1% chance that one of the eight decision tasks from this part will be selected for payment.**]

Please click on the "Next" button to proceed to the first decision task of this part.

Information before starting the decision tasks of block 3:

In the eight decisions tasks of this part you will see risky options that yield small-to-medium negative payouts (losses).

Please click on the "Next" button to proceed to the first decision task of this part.

Information before starting the decision tasks of block 4:

In the eight decisions tasks of this part you will see risky options that yield medium-to-large negative payouts (losses).

Please click on the "Next" button to proceed to the first decision task of this part.

Information before starting the decision tasks of block 5:

In the eight decisions tasks of this part you will see risky options that yield small-to-medium positive payouts (gains) but also small-to-medium negative payouts (losses).

Please click on the “Next” button to proceed to the first decision task of this part.

Information before starting the decision tasks of block 6:

In the eight decisions tasks of this part you will see risky options that yield medium-to-large positive payouts (gains) but also medium-to-large negative payouts (losses).

Please click on the “Next” button to proceed to the first decision task of this part.

C.4 Post-experimental Questionnaire

You have now completed all six parts with decision tasks. There are only a few questions remaining.

During the decision tasks we asked you to position a slider. Was it clear for you what the position of the slider meant?

- Very clear
- Rather clear
- Unclear

During the decision tasks we asked you to evaluate risky options with uncertain payouts at three different points in time. How complicated were these decision tasks?

- Okay
- Complicated
- Too complicated

Please honestly report your attention during the questionnaire.

- I paid attention in all six parts.
- I paid attention in most of the six parts.
- I only paid attention in a few parts or not at all.

When making your decisions, did you take into consideration that the payouts of the risky option arrive at three different points in time?

- Always
- Sometimes
- Never

D Additional Data Analysis

D.1 Exclusion Criteria

Table D.1 summarizes subjects' responses in the post-game survey. The relative share of subjects that gave a specific answer is added in parentheses. Subjects are excluded if they answer at least one of these questions with the 'worst' answer possibility. Table D.2 summarizes the distribution of 'worst' answers. Note that the same subject may give the 'worst' answer to several questions.

The second exclusion criteria is the median decision time. Table D.3 provides an overview of the median decision times among i) the subjects that started the experiment (n = 391), ii) the subjects that were excluded due to their responses in the post-game survey or because they did not finish the experiment (n = 205) , iii) the subjects that finished the experiment and were not excluded due to their responses in the post-game survey (n = 186).

Table D.1: Summary post-game survey responses

Q1: Was it clear what the position of the slider meant?			
Answer	Very clear	Rather clear	Unclear
n	106 (28%)	153 (40%)	119 (31%)
Q2: How complex were the decision tasks?			
Answer	Okay	Complicated	Too complicated
n	101 (26%)	159 (42%)	118 (31%)
Q3: In how many parts did you pay attention?			
Answer	All parts	Most parts	Few parts
n	103 (27%)	130 (34%)	145 (38%)

Table D.2: Exclusion by post-game survey

Worst response to	n
Only Q1	17
Only Q2	17
Only Q3	33
Q1 and Q2	13
Q1 and Q3	24
Q2 and Q3	23
Q1 and Q2 and Q3	65

Table D.3: Overview median decision time

Median time per decision (in seconds)	[0-5)	[5 - 15)	[15 - 30)	[30 - 45)	> 45
Full sample	133	125	66	33	34
Excluded due to survey	95	77	25	7	1
Not excluded due to survey	38	48	41	26	33

D.2 Demographic Variables of Participants

Table D.4 presents summary statistics of main demographic variables of our sample and of the representative subject pool. Column (1) displays the statistics of our main sample ($n = 100$). Columns (2) and (3) display summary statistics of the excluded subjects ($n = 291$) and of the entire representative subject pool from which the participants of our experiment were drawn ($n = 8976$) Comparing these samples we only observe statistically significant differences for i) the share of participants with academic education and ii) the share of participants with an income larger than 4500 Euros (which is only marginally significant when comparing our final sample to the original pool).

Table D.4: Demographic variables of participants

	Final (1)	Excluded (2)	Pool (3)	t-tests (1) vs. (2)	(1) vs. (3)
Sex (1 = Female)	0.48	0.50	0.51		
Age ≤ 24	0.11	0.08	0.09		
Age 25 – 44	0.28	0.26	0.31		
Age 45 – 64	0.30	0.37	0.33		
Age ≥ 65	0.31	0.30	0.27		
Income ≤ 1500	0.28	0.30	0.30		
Income 1501 – 3000	0.27	0.30	0.33		
Income 3001 – 4500	0.26	0.20	0.19		
Income > 4500	0.16	0.11	0.10		*
Primary education	0.05	0.09	0.06		
Secondary education	0.23	0.30	0.30		
Vocational education	0.51	0.51	0.50		
Academic education	0.21	0.11	0.14	**	*

Notes: * and ** indicate significant differences of t-tests at the 10%-level and 5%-level.

D.3 Prediction Error by-Lottery

Table D.5 displays the mean absolute prediction error of the time-separation (TS) and present-value method (PV). The column names refer to the 12 test lotteries and the row names refer to the model combinations as in Table 4. The payout ranges of the lotteries are displayed in parentheses. Table D.6 shows the weighted mean-squared errors of the time-separation and present-value methods per lottery.

Table D.5: Mean absolute prediction error (by lottery)

Comb.	Meth.	Low-stake lotteries L7 (23)	L8 (20)	L15 (23)	L16 (20)	L23 (30)	L24 (50)	High-stake lotteries L31 (467)	L32 (400)	L39 (467)	L40 (400)	L47 (600)	L48 (1000)
C1	TS	6.0	8.1	7.3	5.6	7.9	16.7	144.1	148.6	146	124.9	180	319.9
	PV	5.2	5.6	6.1	5.6	7.2	14.6	108.5	118	114.4	115	150.7	276.8
C2	TS	6.0	8.1	7.3	5.6	8.2	16.3	144.0	148.1	145.9	125.4	188.1	312.3
	PV	5.2	5.6	6.1	5.6	7.0	14.5	108.5	117.7	114.4	114.0	150.6	275.3
C3	TS	6.1	8.1	7.7	5.5	7.5	18.0	149.3	148.6	155.9	123.1	170.4	347.1
	PV	5.7	6.5	6.1	5.3	7.5	14.6	104.5	124.3	116.0	110.3	153.9	279.1
C4	TS	6.2	8.1	7.6	5.5	8.0	17.7	150.6	148.6	153.9	125.4	183.8	339.3
	PV	5.7	7.1	6.1	5.4	7.1	14.6	104.5	132.2	116.0	113.2	150.6	277.6
C5	TS	6.2	8.0	7.9	5.6	7.6	17.6	152.2	148.1	162.4	117.9	174.0	339.2
	PV	5.5	6.2	6.0	5.3	7.4	14.6	104.8	121.2	114.3	109.9	152.1	278.6
C6	TS	6.1	8.1	7.6	5.6	7.6	17.8	149.4	148.6	152.8	123.2	172.6	341.5
	PV	5.5	6.3	6.0	5.3	7.5	14.7	104.7	122.4	115.2	110.1	153.1	279.9
C7	TS	5.3	8.1	6.4	5.6	7.1	17.7	122.4	146.6	128.3	125.4	161	340.7
	PV	5.6	6.2	6.1	5.3	7.5	15.1	104.6	120.8	114.2	109.9	153.4	288.9
C8	TS	5.3	8.1	6.4	5.6	7.1	17.7	122.4	148.6	128.3	125.4	161.0	340.7
	PV	5.6	6.2	6.1	5.3	7.5	15.1	104.6	120.8	114.2	109.9	153.4	288.9
C9	TS	5.3	8.1	6.4	5.6	7.1	17.7	122.4	144.3	128.3	125.4	161.0	340.7
	PV	5.6	6.2	6.1	5.3	7.5	15.1	104.6	120.8	114.2	109.9	153.4	288.9
C10	TS	6.4	7.6	7.8	5.5	7.9	16.7	128.6	141.2	131.4	125.4	158.3	362.6
	PV	5.3	5.7	6.0	5.3	7.2	14.6	104.5	127.5	116.3	111.3	158.6	290.4
C11	TS	6.4	7.9	7.9	5.6	8.8	16.1	129.5	148.6	134.9	121.4	170.1	348.4
	PV	5.3	6.0	6.0	5.3	6.9	14.5	104.5	135.2	115.6	114.0	150.7	286.9
C12	TS	6.6	7.4	8.0	5.4	8.1	16.5	132.7	148.6	136.0	123.4	161.4	355.4
	PV	5.2	5.7	6.0	5.3	7.1	14.5	104.5	128.4	115.7	111.6	156.8	288.1

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. Payout ranges are shown in parentheses below the lottery number.

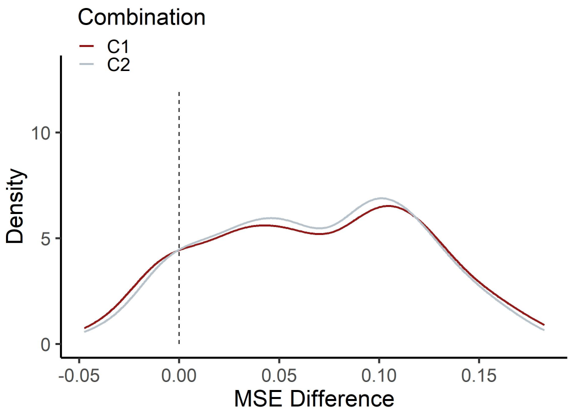
Table D.6: MSE (by lottery)

		Low-stake lotteries						High-stake lotteries					
Comb.	Meth.	L7	L8	L15	L16	L23	L24	L31	L32	L39	L40	L47	L48
C1	TS	0.10	0.27	0.15	0.18	0.11	0.19	0.13	0.25	0.15	0.21	0.14	0.17
	PV	0.08	0.11	0.10	0.10	0.09	0.13	0.08	0.11	0.09	0.11	0.09	0.12
C2	TS	0.10	0.27	0.15	0.18	0.12	0.18	0.13	0.25	0.15	0.21	0.15	0.17
	PV	0.08	0.11	0.10	0.10	0.08	0.12	0.08	0.12	0.09	0.11	0.09	0.12
C3	TS	0.10	0.27	0.16	0.18	0.10	0.22	0.14	0.25	0.16	0.20	0.13	0.20
	PV	0.09	0.17	0.10	0.12	0.10	0.13	0.08	0.16	0.09	0.13	0.09	0.12
C4	TS	0.10	0.27	0.16	0.18	0.12	0.21	0.14	0.25	0.16	0.21	0.15	0.19
	PV	0.09	0.20	0.10	0.15	0.08	0.13	0.08	0.20	0.09	0.16	0.09	0.12
C5	TS	0.11	0.27	0.17	0.18	0.11	0.21	0.15	0.25	0.18	0.18	0.13	0.19
	PV	0.09	0.15	0.10	0.11	0.09	0.13	0.08	0.15	0.09	0.12	0.09	0.12
C6	TS	0.10	0.27	0.16	0.18	0.10	0.21	0.14	0.25	0.16	0.20	0.13	0.20
	PV	0.09	0.16	0.10	0.12	0.09	0.13	0.08	0.15	0.09	0.13	0.09	0.12
C7	TS	0.08	0.27	0.11	0.18	0.09	0.21	0.10	0.25	0.12	0.21	0.11	0.20
	PV	0.09	0.15	0.10	0.12	0.09	0.15	0.08	0.15	0.09	0.12	0.09	0.14
C8	TS	0.08	0.27	0.11	0.18	0.09	0.21	0.10	0.25	0.12	0.21	0.11	0.20
	PV	0.09	0.15	0.10	0.12	0.09	0.15	0.08	0.15	0.09	0.12	0.09	0.14
C9	TS	0.11	0.24	0.17	0.17	0.12	0.19	0.11	0.25	0.12	0.21	0.11	0.22
	PV	0.08	0.12	0.10	0.10	0.09	0.13	0.08	0.18	0.09	0.14	0.10	0.14
C10	TS	0.11	0.26	0.17	0.18	0.14	0.18	0.11	0.25	0.13	0.21	0.13	0.20
	PV	0.08	0.14	0.10	0.11	0.08	0.12	0.08	0.21	0.09	0.16	0.09	0.13
C11	TS	0.12	0.23	0.18	0.16	0.12	0.18	0.11	0.25	0.13	0.21	0.11	0.21
	PV	0.08	0.12	0.10	0.10	0.09	0.12	0.08	0.18	0.09	0.15	0.10	0.14
C12	TS	0.12	0.23	0.18	0.16	0.12	0.18	0.11	0.25	0.13	0.21	0.11	0.21
	PV	0.08	0.13	0.10	0.10	0.08	0.12	0.08	0.18	0.09	0.15	0.09	0.13

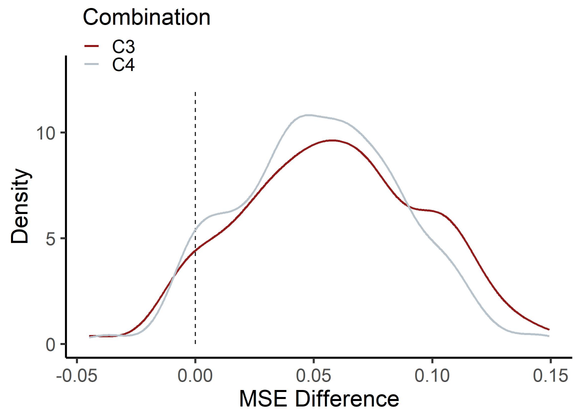
Notes: This table contains the (weighted) MSE for each test lottery and each function combination separately.

D.4 Prediction Error by-Participant

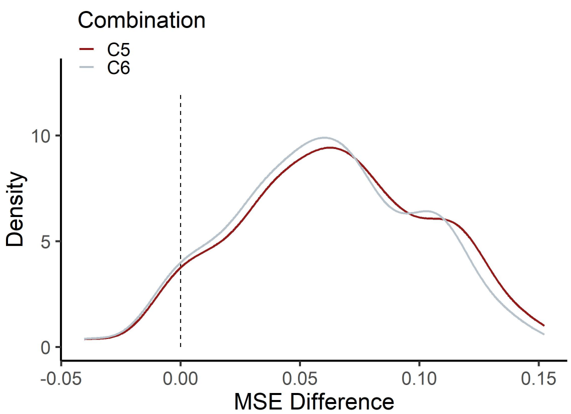
Figure D.1 shows the distribution of the MSE difference ($MSE_{TS} - MSE_{PV}$) on the participant level. The MSE difference is calculated for each participant using the aggregate level parameter estimates displayed in Table D.7 below (these are the parameter estimates on the calibration set). The labels C1 to C12 refer to the combinations as in Table 4. For each of the 12 combinations the MSE difference is (sizable) larger than 0 for a majority of the participants.



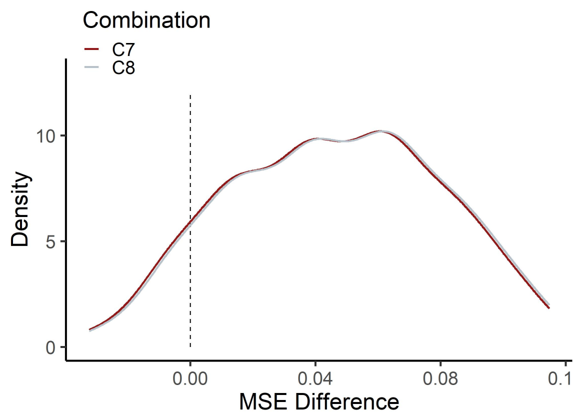
(a) Combinations C1 and C2



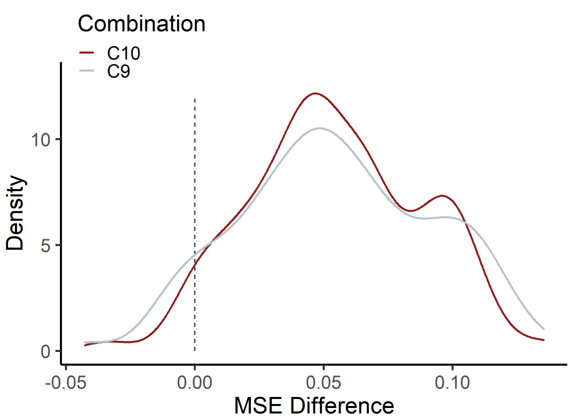
(b) Combinations C3 and C4



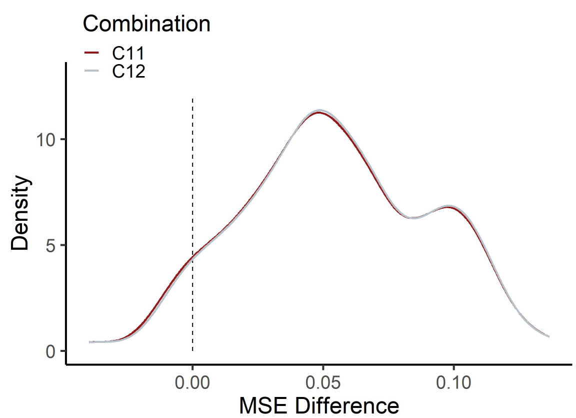
(c) Combinations C5 and C6



(d) Combinations C7 and C8



(e) Combinations C9 and C10



(f) Combinations C11 and C12

Figure D.1: Densities of by-participant MSE differences ($MSE_{TS} - MSE_{PV}$)

Notes: This figure shows the empirical density functions of differences in MSE per participant (MSE_{TS} – MSE_{PV}). Always two function combinations are shown in one panel, only differing by the type of time discounting (exponential or quasi hyperbolic).

D.5 Parameter Estimates on the Calibration Set

Table D.7 shows the parameters estimated on the calibration set (with bootstrapped standard errors in parentheses). These estimates are identical for the time-separation and the present-value method. The labels C1 to C12 refer to the combinations as in Table 4. Combinations C1 to C6, thus, summarize the parameter estimates for the model combinations that employ power utilities as basic utility functions. Combinations C7 to C12 summarize the parameters for the model combinations that employ exponential utilities as basic utility functions. The coefficients for exponential utility (α and β) in C7 to C12 are multiplied by a factor of 1000.

Table D.7: Overview parameter estimates on calibration set (36 lotteries)

	Value				Weighting			Discounting	
Combination	α	β	λ	$\gamma+$	$\gamma-$	$\nu+$	$\nu-$	r	k
C1	1.172		1.176	0.530	0.550			0	
	(0.072)		(0.275)	(0.041)	(0.061)			(0)	
C2	1.182		1.089	0.536	0.557			0	0.943
	(0.081)		(0.295)	(0.04)	(0.061)			(0)	(0.069)
C3	0.817		1.062	0.471	0.481	0.729	0.701	0.001	
	(0.126)		(0.254)	(0.07)	(0.092)	(0.123)	(0.151)	(0.012)	
C4	0.801		1.072	0.488	0.508	0.681	0.657	0	0.845
	(0.123)		(0.206)	(0.072)	(0.094)	(0.119)	(0.145)	(0.003)	(0.063)
C5	0.896		1.116	0.432	0.443	1.052	1.115	0	
	(0.061)		(0.244)	(0.06)	(0.072)	(0.112)	(0.168)	(0.006)	
C6	0.891		1.142	0.425	0.437	1.050	1.101	0	0.990
	(0.062)		(0.243)	(0.06)	(0.072)	(0.108)	(0.167)	(0.005)	(0.037)
C7	-0.016	-0.120	1.137	0.617	0.636			0	
	(0.150)	(0.1)	(0.210)	(0.033)	(0.042)			(0)	
C8	-0.016	-0.120	1.137	0.617	0.636			0	1
	(0.149)	(0.096)	(0.212)	(0.032)	(0.042)			(0)	(0)
C9	0.418	0.402	1.136	0.455	0.476	0.870	0.850	0.002	
	(0.126)	(0.115)	(0.219)	(0.062)	(0.081)	(0.047)	(0.04)	(0.014)	
C10	0.406	0.382	1.09	0.474	0.494	0.841	0.812	0	0.843
	(0.124)	(0.117)	(0.184)	(0.062)	(0.082)	(0.052)	(0.046)	(0.003)	(0.060)
C11	0.40	0.390	1.123	0.441	0.456	0.961	1.013	0.004	
	(0.128)	(0.119)	(0.217)	(0.06)	(0.073)	(0.070)	(0.076)	(0.010)	
C12	0.392	0.376	1.105	0.443	0.456	0.964	1.016	0	0.979
	(0.131)	(0.119)	(0.226)	(0.061)	(0.073)	(0.074)	(0.08)	(0.005)	(0.037)

Notes: This table shows the parameter estimates on the calibration set, for four selected function combinations. These estimates are identical for time-separation and present-value methods.

D.6 Parameter Estimates on All Lotteries

Table D.8 shows the parameter estimates (with standard errors in parentheses) for the present-value method, estimated on all lotteries. Table D.9 shows the same estimates for the time-separation method. The labels C1 to C12 refer to the combinations as in Table 4. Combinations C1 to C6, thus, summarize the parameter estimates for the model combinations that employ power utilities as basic utility functions. Combinations C7 to C12 summarize the parameters for the model combinations that employ exponential utilities as basic utility functions. The coefficients for exponential utility (α and β) are multiplied by a factor of 1000.

Table D.8: Overview parameter estimates on full sample (48 lotteries)

	Value				Weighting			Discounting	
Combination	α	β	λ	$\gamma+$	$\gamma-$	$\nu+$	$\nu-$	r	k
C1	1.195		1.069	0.535	0.551			0.117	
	(0.068)		(0.096)	(0.045)	(0.06)			(0.036)	
C2	1.172		1.065	0.547	0.565			0.016	0.863
	(0.061)		(0.121)	(0.043)	(0.057)			(0.025)	(0.043)
C3	1.092		1.010	0.406	0.427	0.992	0.988	0.076	
	(0.04)		(0.150)	(0.067)	(0.082)	(0.065)	(0.057)	(0.032)	
C4	1.080		1.104	0.425	0.453	0.976	0.975	0	0.855
	(0.032)		(0.182)	(0.068)	(0.08)	(0.059)	(0.054)	(0.003)	(0.037)
C5	1.091		1.037	0.426	0.448	0.766	0.786	0.070	
	(0.035)		(0.127)	(0.066)	(0.073)	(0.074)	(0.071)	(0.028)	
C6	1.085		1.101	0.423	0.451	0.786	0.796	0.0012	0.884
	(0.03)		(0.163)	(0.07)	(0.073)	(0.073)	(0.068)	(0.016)	(0.045)
C7	-0.115	-0.116	0.960	0.622	0.639			0	
	(0.066)	(0.065)	(0.095)	(0.035)	(0.042)			(0)	
C8	-0.122	-0.112	0.919	0.624	0.638			0	1
	(0.065)	(0.065)	(0.096)	(0.036)	(0.043)			(0)	(0)
C9	0.078	0.200	0.925	0.452	0.446	0.897	0.861	0	
	(0.049)	(0.060)	(0.113)	(0.07)	(0.086)	(0.05)	(0.036)	(0.013)	
C10	0.07	0.166	1	0.442	0.450	0.879	0.855	0	0.953
	(0.05)	(0.061)	(0.116)	(0.069)	(0.086)	(0.049)	(0.036)	(0)	(0.034)
C11	0.064	0.170	0.943	0.428	0.422	0.913	0.974	0	
	(0.050)	(0.059)	(0.106)	(0.07)	(0.077)	(0.064)	(0.067)	(0.007)	
C12	0.066	0.179	0.933	0.428	0.423	0.912	0.975	0	1
	(0.048)	(0.06)	(0.118)	(0.07)	(0.077)	(0.065)	(0.067)	(0)	(0.023)

Notes: This table shows the parameter estimates on the full set of lotteries (calibration plus test set) for the present-value method.

Table D.9: Overview parameter estimates time-separation method (48 lotteries)

	Value			Weighting				Discounting	
Combination	α	β	λ	$\gamma+$	$\gamma-$	$\nu+$	$\nu-$	r	k
C1	1.037		1.255	0.791	0.76			0	
	(0.032)		(0.134)	(0.086)	(0.090)			(0)	
C2	1.032		1.224	0.788	0.764			0	0.96
	(0.034)		(0.131)	(0.085)	(0.090)			(0)	(0)
C3	1.077		1.245	0.73	0.699	1.038	1.025	0	
	(0.046)		(0.189)	(0.113)	(0.130)	(0.046)	(0.050)	(0)	
C4	1.049		1.116	0.751	0.694	1.003	0.978	0	1
	(0.045)		(0.140)	(0.110)	(0.130)	(0.043)	(0.045)	(0)	(0)
C5	1.047		1.164	0.745	0.686	0.893	0.895	0	
	(0.031)		(0.135)	(0.108)	(0.128)	(0.057)	(0.072)	(0)	
C6	1.033		1.118	0.754	0.687	0.913	0.912	0	1
	(0.032)		(0.140)	(0.108)	(0.127)	(0.057)	(0.071)	(0)	(0)
C7	-0.012	-0.034	1.1468	0.828	0.806			0	
	(0.043)	(0.056)	(0.096)	(0.065)	(0.071)			(0)	
C8	-0.014	-0.034	1.155	0.829	0.808			0	1
	(0.042)	(0.057)	(0.097)	(0.064)	(0.071)			(0)	(0)
C9	0.1	0.15	1.109	0.807	0.758	0.921	0.897	0	
	(0.08)	(0.053)	(0.093)	(0.087)	(0.120)	(0.049)	(0.033)	(0)	
C10	0.0916	0.1508	1.104	0.802	0.755	0.922	0.897	0	1
	(0.079)	(0.054)	(0.105)	(0.086)	(0.120)	(0.050)	(0.033)	(0)	(0)
C11	0.09	0.148	1.138	0.78	0.7305	1.025	1.0417	0	
	(0.074)	(0.054)	(0.106)	(0.099)	(0.125)	(0.044)	(0)	(0)	
C12	0.0986	0.1486	1.098	0.7736	0.7128	1.021	1.041	0	1
	(0.074)	(0.053)	(0.091)	(0.098)	(0.124)	(0.045)	(0)	(0)	(0)

Notes: This table shows the parameter estimates on the full set of lotteries (calibration plus test set) for the time-separation method.

D.7 Parameter Estimates for the Monetary Present-value Method (All Lotteries)

Table D.10 shows the parameter estimates (with standard errors in parentheses) for the monetary-present value method. The labels C1 to C12 refer to the combinations as in Table 4. Combinations C1 to C6, thus, summarize the parameter estimates for the model combinations that employ power utilities as basic utility functions. Combinations C7 to C12 summarize the parameters for the model combinations that employ exponential utilities as basic utility functions. The coefficients for exponential utility (α and β) are multiplied by a factor of 1000.

Table D.10: Calibration monetary present-value method (all combinations)

	Value				Weighting			Discounting	
Combination	α	β	λ	$\gamma+$	$\gamma-$	$\nu+$	$\nu-$	r	k
C1	1.179		1.133	0.579	0.592			0	
	(0.075)		(0.227)	(0.040)	(0.058)			(0)	
C2	1.182		1.141	0.576	0.576			0	1
	(0.072)		(0.224)	(0.041)	(0.058)			(0)	(0.001)
C3	0.886		1.05	0.415	0.432	0.78	0.774	0.073	
	(0.065)		(0.206)	(0.070)	(0.082)	(0.071)	(0.097)	(0.039)	
C4	0.889		1.075	0.427	0.449	0.782	0.789	0	0.885
	(0.047)		(0.192)	(0.068)	(0.082)	(0.058)	(0.063)	(0.002)	(0.036)
C5	0.915		1.033	0.403	0.406	1.036	1.05	0.041	
	(0.049)		(0.223)	(0.068)	(0.075)	(0.097)	(0.112)	(0.023)	
C6	0.915		1.087	0.406	0.421	1.03	1.05	0.001	0.938
	(0.049)		(0.200)	(0.069)	(0.074)	(0.096)	(0.118)	(0.009)	(0.032)
C7	-0.092	-0.096	1.115	0.634	0.646			0	
	(0.075)	(0.066)	(0.224)	(0.034)	(0.043)			(0)	
C8	-0.074	-0.0852	1.11	0.628	0.639			0	1
	(0.079)	(0.062)	(0.232)	(0.033)	(0.048)			(0)	(0)
C9	0.174	0.208	1.147	0.427	0.452	0.834	0.841	0.038	
	(0.048)	(0.054)	(0.2)	(0.068)	(0.083)	(0.046)	(0.035)	(0.021)	
C10	0.174	0.222	1.138	0.428	0.456	0.829	0.836	0	0.918
	(0.049)	(0.053)	(0.175)	(0.068)	(0.082)	(0.046)	(0.036)	(0.001)	(0.029)
C11	0.196	0.226	1.151	0.429	0.441	0.996	1.023	0.018	
	(0.054)	(0.050)	(0.186)	(0.070)	(0.074)	(0.061)	(0.071)	(0.019)	
C12	0.198	0.218	1.12	0.41	0.429	0.993	1.0246	0	0.95
	(0.054)	(0.051)	(0.187)	(0.067)	(0.074)	(0.063)	(0.071)	(0.005)	(0.029)

Notes: This table shows the parameter estimates on the full set of lotteries (calibration plus test set) for the monetary present-value method.

D.8 Components of PT (Time-separation Method)

Figure D.2 displays the prediction performance of different Time-Separation models that either i) include all PT components, ii) force linear probability weighting, iii) force no loss aversion, or iv) force a linear utility. The prediction performances of models with a linear utility or no loss aversion are almost identical to the performance of full PT models. Also models that force linear-probability weighting predict decisions only a bit worse than full PT models.

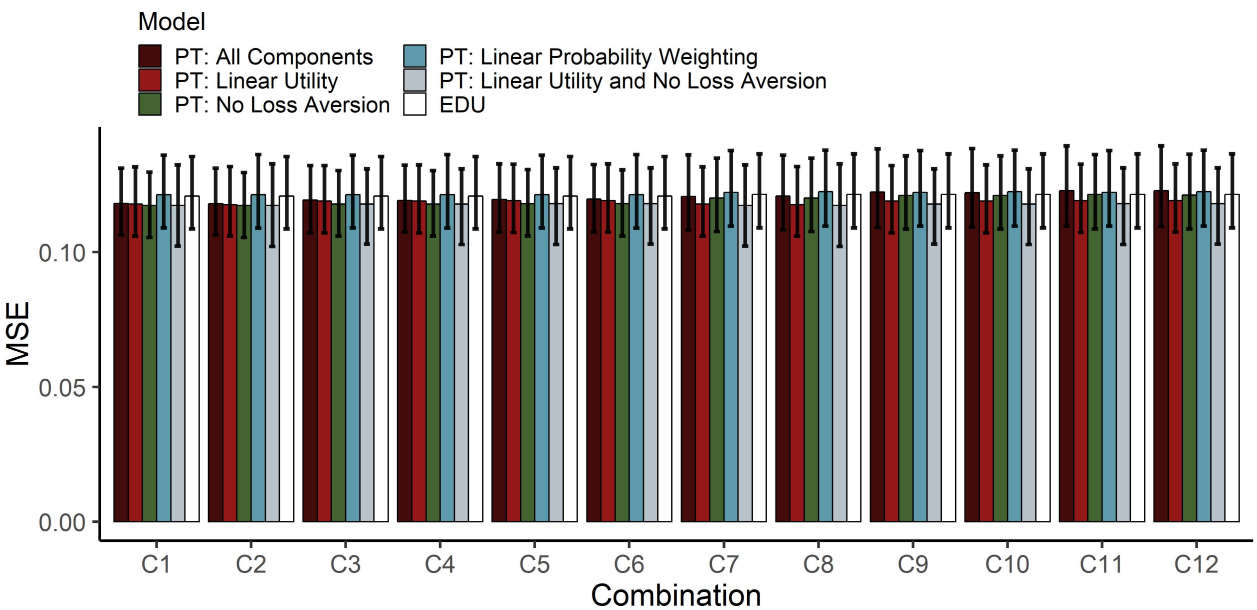


Figure D.2: MSE time-separation method, PT components (all function combinations)

Notes: This figure compares the prediction performance of the time-separation method to versions of it with some components absent and to expected utility theory (parameters are estimated anew for each method). The versions are (from left to right for each of the 12 function combinations) the monetary present-value method with all components ('PT: All Components'), with linear utility for gains and losses still including a loss-aversion parameter ('PT: Linear Utility'), with a loss-aversion parameter of 1 ('PT: No Loss Aversion'), with linear-probability weighting ('PT: Linear Probability Weighting'), with linear utility for gains and losses and a loss-aversion parameter of 1 ('PT: Linear Utility and No Loss Aversion'); on the right expected discounted utility is shown ('EDU').

D.9 Components of PT (Monetary Present Value Method)

Figure D.3 displays the prediction performance of different Monetary-Present-value Models that either i) include all PT components, ii) force linear probability weighting, iii) force no loss aversion, or iv) force a linear utility. The prediction performances of models with a linear utility or no loss aversion are almost identical to the performance of full PT models. However, for models that force linear-probability weighting the prediction performance becomes a lot worse.

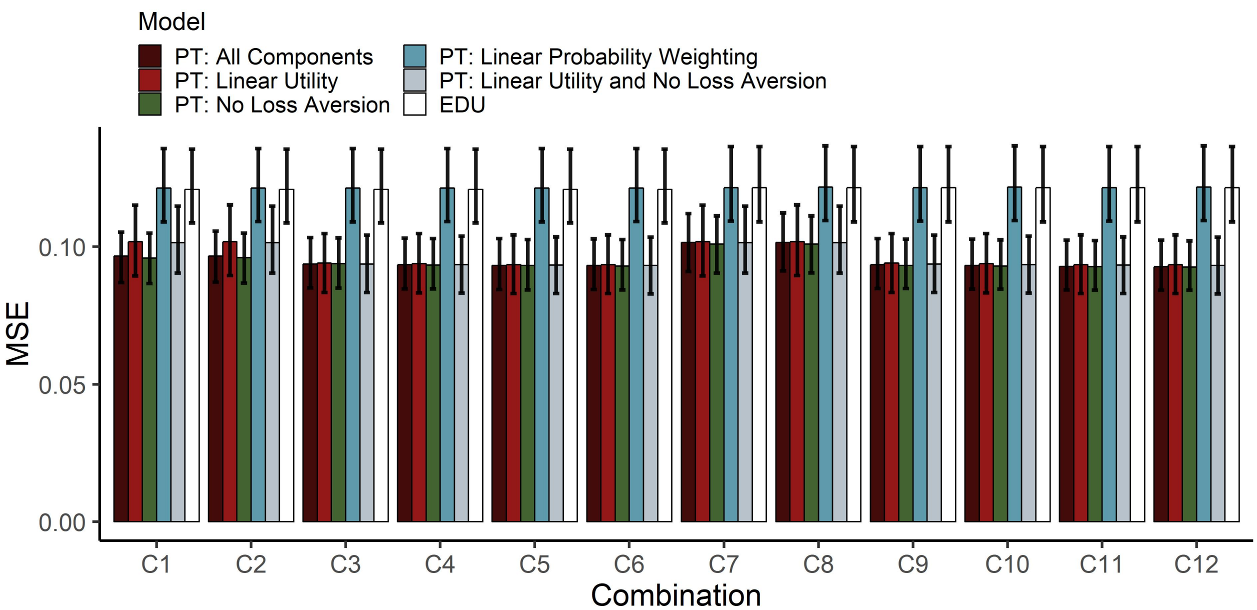


Figure D.3: MSE monetary present-value method, PT components (all function combinations)

Notes: This figure compares the prediction performance of the monetary present-value method to versions of it with some components absent and to expected utility theory (parameters are estimated anew for each method). The versions are (from left to right for each of the 12 function combinations) the monetary present-value method with all components ('PT: All Components'), with linear utility for gains and losses still including a loss-aversion parameter ('PT: Linear Utility'), with a loss-aversion parameter of 1 ('PT: No Loss Aversion'), with linear-probability weighting ('PT: Linear Probability Weighting'), with linear utility for gains and losses and a loss-aversion parameter of 1 ('PT: Linear Utility and No Loss Aversion'); on the right expected discounted utility is shown ('EDU').

D.10 Allowing for Non-zero Reference Points

Table D.11 shows the parameter estimates (with standard errors in parentheses) for the present-value method, with estimated reference points r_0 (now), r_1 (in three months), and r_2 (in six months). Table D.12 shows the same estimates for the time-separation method. The labels C1 to C12 refer to the combinations as in Table 4. Combinations C1 to C6, thus, summarize the parameter estimates for the model combinations that employ power utilities as basic utility functions. Combinations C7 to C12 summarize the parameters for the model combinations that employ exponential utilities as basic utility functions. The coefficients for exponential utility (α and β) are multiplied by a factor of 1000.

Table D.11: Overview parameter estimates present-value method (48 lotteries), estimated reference points

	Value					Weighting					Discounting	
Combination	r_0	r_1	r_2	α	β	λ	$\gamma+$	$\gamma-$	$\nu+$	$\nu-$	r	k
C1	-1.006	-1.285	-1.333	1.204		1.085	0.518	0.551			0.121	
	(1.4)	(1.249)	(1.321)	(0.069)		(0.113)	(0.073)	(0.056)			(0.038)	
C2	-1.019	-1.394	-1.228	1.17		1.109	0.532	0.567			0.004	0.841
	(1.338)	(1.222)	(1.299)	(0.076)		(0.13)	(0.07)	(0.058)			(0.039)	(0.056)
C3	-0.239	0.194	0.207	1.111		1.131	0.406	0.428	0.978	1.02	0.103	
	(1.126)	(1.835)	(1.003)	(0.048)		(0.139)	(0.065)	(0.079)	(0.095)	(0.056)	(0.036)	
C4	-0.323	-0.215	-0.23	1.064		1.189	0.422	0.459	0.935	0.96	0	0.855
	(1.139)	(1.829)	(1.693)	(0.037)		(0.136)	(0.066)	(0.078)	(0.088)	(0.049)	(0.015)	(0.038)
C5	-0.499	-0.737	-0.795	1.07		1.055	0.418	0.435	0.805	0.814	0.054	
	(1.15)	(1.82)	(1.367)	(0.047)		(0.133)	(0.065)	(0.073)	(0.111)	(0.07)	(0.033)	
C6	-0.382	-0.351	-0.374	1.07		1.156	0.419	0.442	0.827	0.799	0.019	0.926
	(1.162)	(1.84)	(1.487)	(0.04)		(0.153)	(0.068)	(0.072)	(0.104)	(0.073)	(0.025)	(0.045)
C7	0.125	-0.236	-0.221	-0.13	-0.102	0.934	0.617	0.644			0	
	(0.306)	(0.245)	(0.228)	(0.071)	(0.067)	(0.084)	(0.039)	(0.041)			(0)	
C8	0.125	-0.236	-0.221	-0.13	-0.102	0.934	0.617	0.644			0	1
	(0.307)	(0.262)	(0.243)	(0.071)	(0.067)	(0.075)	(0.04)	(0.042)			(0)	(0)
C9	0.084	-0.136	-0.122	0.068	0.216	0.933	0.441	0.448	0.901	0.846	0	
	(0.41)	(0.536)	(0.502)	(0.08)	(0.094)	(0.096)	(0.066)	(0.083)	(0.086)	(0.065)	(0.012)	
C10	0.186	-0.235	-0.218	0.047	0.205	0.935	0.435	0.448	0.905	0.832	0	0.958
	(0.375)	(0.505)	(0.472)	(0.076)	(0.089)	(0.107)	(0.067)	(0.083)	(0.081)	(0.058)	(0)	(0.034)
C11	0.034	-0.09	-0.078	0.089	0.17	1	0.425	0.422	0.928	0.967	0	
	(0.33)	(0.478)	(0.445)	(0.071)	(0.082)	(0.094)	(0.07)	(0.072)	(0.1)	(0.11)	(0.006)	
C12	0.034	-0.09	-0.078	0.089	0.17	1	0.425	0.422	0.928	0.967	0	1
	(0.319)	(0.475)	(0.442)	(0.071)	(0.083)	(0.1)	(0.07)	(0.072)	(0.098)	(0.108)	(0)	(0.02)

Notes: This table shows the parameter estimates on the full set of lotteries (calibration plus test set) for the present-value method, including estimates of the reference points.

Figure D.4 illustrates the difference in MSE between the time-separation and the present-value methods, when reference points are estimated. The confidence intervals show that all

Table D.12: Overview parameter estimates time-separation method (48 lotteries), estimated reference points

	Value					Weighting					Discounting	
Combination	r_0	r_1	r_2	α	β	λ	$\gamma+$	$\gamma-$	$\nu+$	$\nu-$	r	k
C1	-0.523	-0.34	-0.375	1.031		1.142	0.802	0.773			0	
	(1.813)	(1.613)	(1.253)	(0.033)		(0.135)	(0.105)	(0.103)			(0)	
C2	-0.523	-0.34	-0.375	1.031		1.142	0.802	0.773			0	1
	(1.845)	(2.137)	(2.279)	(0.033)		(0.134)	(0.104)	(0.104)			(0)	(0)
C3	-0.223	0.658	0.673	1.102		1.203	0.724	0.661	1.047	1.07	0	
	(1.898)	(2.558)	(2.716)	(0.096)		(0.196)	(0.144)	(0.15)	(0.103)	(0.101)	(0)	
C4	-0.223	0.658	0.673	1.102		1.203	0.724	0.661	1.047	1.07	0	1
	(1.932)	(2.608)	(2.765)	(0.1)		(0.213)	(0.145)	(0.149)	(0.107)	(0.107)	(0)	(0)
C5	-0.316	0.541	0.543	1.035		1.172	0.76	0.674	0.947	0.872	0	
	(1.216)	(1.914)	(1.072)	(0.086)		(0.221)	(0.135)	(0.141)	(0.153)	(0.147)	(0.001)	
C5	-0.316	0.541	0.543	1.035		1.172	0.76	0.674	0.947	0.872	0	1
	(1.255)	(1.902)	(1.058)	(0.09)		(0.229)	(0.134)	(0.14)	(0.157)	(0.153)	(0)	(0)
C7	-0.279	0.134	0.132	-0.003	-0.033	1.077	0.847	0.803			0	
	(0.324)	(0.288)	(0.28)	(0.041)	(0.063)	(0.084)	(0.079)	(0.089)			(0)	
C8	-0.279	0.134	0.132	-0.003	-0.033	1.077	0.847	0.803			0	1
	(0.325)	(0.287)	(0.279)	(0.041)	(0.062)	(0.08)	(0.079)	(0.089)			(0)	(0)
C9	-0.105	0.003	0.002	0.105	0.116	1.057	0.804	0.743	0.922	0.902	0	
	(0.611)	(0.682)	(0.668)	(0.096)	(0.152)	(0.082)	(0.105)	(0.115)	(0.051)	(0.076)	(0)	
C10	-0.077	-0.026	-0.028	0.102	0.12	1.035	0.804	0.745	0.929	0.899	0	0.99
	(0.625)	(0.675)	(0.661)	(0.099)	(0.153)	(0.078)	(0.105)	(0.115)	(0.052)	(0.076)	(0)	(0)
C11	0.514	-0.651	-0.641	0.011	0.282	1.045	0.77	0.711	0.944	1.148	0	
	(0.517)	(0.577)	(0.561)	(0.089)	(0.125)	(0.133)	(0.103)	(0.121)	(0.066)	(0.093)	(0)	
C12	0.571	-0.709	-0.698	0.006	0.288	1.05	0.767	0.714	0.938	1.153	0	0.99
	(0.502)	(0.58)	(0.564)	(0.088)	(0.125)	(0.125)	(0.102)	(0.12)	(0.065)	(0.092)	(0)	(0)

Notes: This table shows the parameter estimates on the full set of lotteries (calibration plus test set) for the time-separation method, including estimates of the reference points.

differences are strongly significant.

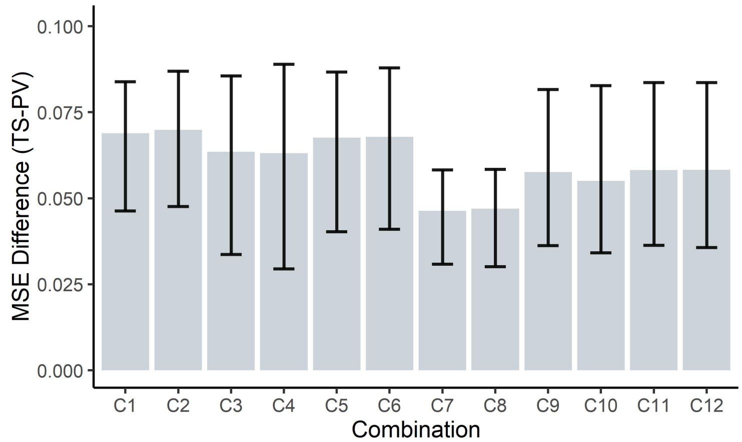


Figure D.4: Difference in MSE between the application methods (all function combinations), estimated reference points

Notes: This figure shows the average differences in mean squared errors (MSE of time-separation method minus MSE of present-value method) with bootstrapped 95% confidence intervals.

D.11 Treatment Comparison

Table D.13 summarizes mean evaluations for lotteries of Set 1 and Set 4 (the two sets of gain lotteries) of participants in the incentivized and hypothetical treatment. Different statistical test (exact Schlag tests, as pre-registered, as well as t-tests and Wilcoxon-Mann-Whitney tests) do not show any significant differences in evaluations.

Table D.13: Evaluation of gain lotteries in T1 and T2

Lottery	Mean evaluation		Schlag-tests		t-tests		WMW-Tests	
	T1 (Hyp.)	T2 (Inc.)	p	Adj. p	p	Adj. p	p	Adj. p
L1	10.18	10.26	1	1	0.97	0.97	0.79	1
L2	13.07	10.37	0.71	1	0.12	1	0.09	1
L3	17.07	18.78	1	1	0.33	1	0.21	1
L4	16.12	15.04	1	1	0.48	1	0.39	1
L5	14.55	12	0.78	1	0.14	1	0.06	0.9
L6	15.47	15.89	1	1	0.79	1	0.87	0.87
L7	13.48	12.04	1	1	0.34	1	0.24	1
L8	27.75	28.93	1	1	0.45	1	0.5	1
L33	188.32	175.41	1	1	0.73	1	0.75	1
L34	239.64	206.78	1	1	0.31	1	0.52	1
L35	349.58	390.44	0.94	1	0.2	1	0.23	1
L36	293.27	299.52	1	1	0.83	1	0.86	1
L37	262.08	249.22	1	1	0.72	1	0.47	1
L38	279.15	274.56	1	1	0.88	1	0.57	1
L39	208.48	233.63	1	1	0.4	1	0.41	1
L40	540.44	570.7	0.99	1	0.34	1	0.26	1

Notes: This table shows mean evaluations for lotteries of Set 1 and Set 4 (the two sets of gain lotteries) of participants in the incentivized and hypothetical treatments. Simple p-values are denoted by p, Bonferroni-Holm-adjusted p-values are denoted by Adj. p.