Time for Tea: Measuring Discounting for Money and Consumption without the Utility Confound

Mohammed Abdellaoui, Emmanuel Kemel, Amma Panin*, Ferdinand M. Vieider

10 February 2023

Abstract

We present a novel method—called risk equivalents—that uses a single measure to elicit discount rates while avoiding concerns about the shape of the utility function. The method is valid under discounted expected utility (DEU), and also under several of its behavioral extensions including more general models that account for a biased perception of time and risk (such as time- or likelihood-insensitivity). We implement the method in a field experiment in India measuring inter-temporal discount rates for money and the consumption of tea.We empirically observe that discount rates elicited by traditional methods are related to utility curvature, whereas discount rates elicited by risk equivalents are not. Risk equivalents also mitigate differences in discount rates measured for money and for tea, suggesting that unaddressed utility curvature may play a role in results that demonstrate good-specific discounting. Risk equivalents are general, fast and tractable, three features that are particularly useful in field studies.

Keywords: time discounting; money vs consumption; utility confound

JEL-classification: D03; D81; D91

*Corresponding author: amma.panin@uclouvain.be; Voie du Roman Pays 34 1348, Louvain-la-Neuve, Belgium. Code and data to reproduce the analyses in this paper can be accessed at the following Open Science Foundation project: https://osf.io/ya7ke. The experiment reported in this paper was financed by the WZB Berlin Social Science Center. Ferdinand Vieider gratefully acknowledges financial support received from the German Science Foundation (DFG) under project VI 692/1-1. The authors declare no competing interests.

1 Introduction

Economists model intertemporal decisions as a sum of future utilities from consumption, discounted to the present by a discount function. Historically, most effort has gone into obtaining and refining measurements of the discount function, assuming linear utility. More recently, attention has focused on the biases introduced when neglecting the role of utility curvature, and on the possible ways to account for it. In this paper, we address this topic by presenting a simple task that can be used to measure discount factors without worrying about utility curvature. Participants trade off risk against time. They decide between a sure outcome x that is available at a later time ℓ (i.e. an intertemporal prospect x_ℓ) and the same outcome x that is available in the present with a probability p (i.e. a risky prospect (x, p)). Under the standard model of discounted expected utility (DEU), the discount factor is given by p^*, the probability that makes the decision maker indifferent between x_ℓ and $(x, p^\star)$. We further show that this probability remains a valid measure of the discount factor under various behavioral extensions of DEU. We call this probability the *risk equivalent.*

We use the task to elicit discount factors for different time delays with 583 respondents from several Indian villages in Karnataka state. We elicit discount factors for both money and tea, and compare how measured discounting differs between the two goods. In order to assess the validity of our method, we also elicit discount factors using the more traditional *present equivalent* task, which involves participants choosing between a variety of riskless smaller-sooner payments and a fixed larger-later payment to find a point of indifference.

This process results in four sets of discount factors: two methods (discount factors measured either by *risk equivalents* or by *present equivalents*), each applied to the consumption of two different goods (discount factors measured either for money or for tea). We first compare the two methods to each other to examine whether risk equivalents return discount rates that avoid the problem of utility curvature. We also explore differences in discounting for money and for tea. Finally, we use an expanded set of risk equivalents measured at different time periods to test for present bias and decreasing impatience. We also test for subadditivity, the tendency to discount more when a given time delay is split into shorter intervals (Read, 2001, Dohmen et al., 2017).

We observe that the discount factors measured by risk equivalents are not impacted by utility curvature (as measured from structural estimations based on present equivalents). They are also less heterogeneous, which is likely to be due to the fact that our method is not impacted by heterogeneity in utility, nor my measurement errors that may occur when utility has to be controlled for econonometrically. We interpret these results as evidence that risk equivalents really do avoid the utility confound. We also show that there is a degree of good-specific discounting that is mitigated when discount factors are measured with risk equivalents. Finally, we demonstrate that in our sample we do not find evidence of present bias although we do find evidence of subadditivity (the tendency to discount more when a long interval is split into multiple smaller intervals).

Our study thus addresses three aspects of preferences that have been found to impact the measurement of discount rates: utility curvature, good-specificity, and deviation from the DEU model of rational choice. This allows us to add to the existing literature, which we review next.

Several papers have studied how intertemporal utility curvature affects discounting. All these methods require multiple tasks to identify a single discount factor. In contrast, our approach requires only a single measure of a risk equivalent. Additionally, structural econometric specifications are needed to identify the utility function and account for its impact in the estimation of discount factors. One such approach has been to measure risk attitudes in addition to time discounting, and to correct discount factors from the curvature of the risky utility function (Chapman, 1996; Andersen et al., 2008). Another approach has been to jointly estimate utility curvature and discount factors from a series of intertemporal choices. Examples following this structural estimation approach include Andreoni and Sprenger (2012) who use a series of inter-temporal allocations called "convex budget tasks"; and Abdellaoui et al. (2013) who use the present equivalents of a series of streams of outcomes. The method that comes closest to risk equivalents is the "direct method" of Attema et al. (2016), which however requires choices between multi-period streams of outcomes. We include a detailed discussion of the merits of different methods farther below.

The question of whether discounting is good-specific has also received some attention, with mixed results. Reuben et al. (2010) compared discounting for money to discounting for chocolate, finding higher discount rates for chocolate. Their conclusions are, however, based on the assumption of linear utility (i.e., no desire for inter-temporal smoothing of consumption), which could be questioned in the case of a consumption good such as chocolate. Augenblick et al. (2015) investigated the stationarity of time preferences for money and effort in a dynamic experiment where subjects could revise their choices, and found decreasing impatience for effort, but not for money. While they did control for utility curvature in their setup, they could do so only at the aggregate level, assuming that everybody has the same utility. The test presented furthermore combined a gain frame (monetary payoffs) with a loss frame (effort provision). This constitutes a confound, given that time preferences have long been known to differ for gains and losses (Loewenstein and Prelec, 1992, Frederick et al., 2002, Abdellaoui et al., 2018). Ubfal (2016) compared discounting for 19 different goods in a field experiment in Uganda, finding that discounting differed between some of those goods. He could, however, not directly account for utility curvature and used mostly hypothetical choices.

Our analysis is based on a field study and considers incentivized choices with two types of outcomes that are relevant for to our study population in rural India: money and tea.The main advantage of our method is that it does not require the measurement of utility, or indeed any advanced structural econometrics. This is because the method delivers nonparametric discount factors while circumventing utility curvature without the need for any further adjustment. Besides the measurement of utility, several behavioral factors may further complicate attempts to measure impatience as a single "discount rate". For example, omitting the role of present bias (typically captured by a quasi-hyperbolic discount function; Laibson, 1997), or strongly decreasing impatience (typically captured by hyperbolic discount functions; e.g. Ebert and Prelec, 2007), would lead to overestimating the discount rate over short delays from the present, and to underestimating it over long delays with up-front waiting periods. Here again, these patterns are generally accounted for by using econometric estimations and fitting specific functional forms to the data. We show that, under plausible behavioral assumptions (Baucells and Heukamp, 2012; Abdellaoui et al., 2019), risk equivalents neutralize the impact of hyperbolic behavior and provide discount factors that are consistent with exponential discounting. Therefore, risk equivalents are not only robust to the impact of the utility, they are also robust to the impact of discounting biases.

In addition to its theoretical properties, an advantage of risk equivalents is that they involve a single consequence, and the measurement is made on the probability scale. Consequently, it can be used to measure discounting for intertemporal tradeoffs of qualitative consequences (e.g. health states, education level, social status). This is indeed a clear advantage of the risk equivalent method, since the difficulties of measuring utility curvature for such consequences makes the risky utility correction and convex time budget methods ill-suited in this sort of setting.[1]Individual discount

[1]Attema et al. (2016) have developed a method to directly measure discounting under discounted utility. The latter, however, requires choices between between future consumption **streams** instead of single outcomes. This makes them less easily applicable with the sort of subject pool we use in this paper, whose members are constrained

factors can be measured from a single task, avoiding the complications of structural econometric modeling. The method is therefore general, fast and tractable, three features that are particularly useful in field studies.

The next section introduces the method and describes the experiment. Section 3 presents the results, and section 4 concludes the paper.

2 Description of risk equivalents and the field experiment

2.1 Measuring inter-temporal discounting

We first introduce risk equivalents—the method used to measure inter-temporal discounting in this paper. Then, we present a more classical measurement method based on present values, used as a benchmark. The stimuli used for each method, ie the list of risk equivalents and present values, as well as descriptive statistics on their measured responses, are reported in Table 1.

2.1.1 Obtaining discount factors from *risk equivalent*s

To explain how our risk equivalent (RE) method allows to measure discounting, let us first fix an outcome x and two delays: a sooner delay s and a later delay ℓ. In the subsequent text, (x_t, p) denotes the risky option that gives x at time $t \in \{s, \ell\}$, with probability p and nothing with probability $1 - p$. The RE task consists in eliciting the probability $p^\star$ that makes the respondent indifferent between $(x_s, p^\star)$and $(x_\ell, 1)$, i.e. receiving x at a sooner date s with probability $p^\star$ or receiving x for sure at a later date ℓ. Under DEU, this indifference reduces to the equation $p^* D(s)u(x) = D(\ell)u(x)$, where $u(\cdot)$ stands for utility and $D(t) = e^{-rt}$ is the standard exponential discounting function. The discount function is strictly decreasing and such that $D(0) = 1$. The utility is strictly increasing and defined up a to positive linear transformation. We use normalization $u(0) = 0$. After simplifying and rearranging, we obtain

$$p^* = \frac{D(\ell)}{D(s)}. \tag{1}$$

When $s = 0$, we can directly access the discount factor $\delta_\ell = D(\ell)$ for a delay ℓ from the present: $\delta_\ell = p^*$. Under DEU, the discount function is exponential: $D(\ell) = e^{-r\ell}$, such that a single risk

in terms of time and familiarity with complex, abstract, experimental decision problems.

equivalent allows us to capture the discount rate that measures impatience, $r = -\frac{1}{\ell} ln(p^{*})$ (since $D(\ell) = \delta_{\ell}$ and $\delta_{\ell} = p^{*}$). Further adding tasks with $s > 0$ allows us to relax the assumption of DEU regarding the shape of D, and to assess whether the measured discount rate is indeed constant, or whether there is some form of decreasing impatience.

Non-exponential discounting is one of the two main deviations from DEU (Loewenstein and Prelec, 1992), the other one being nonlinear probability weighting (Tversky and Kahneman, 1992). Abdellaoui et al. (2019) propose and estimate a general model on choices involving both risk and time, called rank-dependent discounted utility. The model introduces a probability weighting function w and a non-exponential discount function $D(t)$ such that a prospect that gives x at time t with probability p or else y at the same time t with probability $1 - p$ is valued

$$D(t)[w(p)[u(x) - u(y)] + u(y)].$$

Under reasonable assumptions about the parametric specification of D and w, risk equivalents capture an exponential discount rate under this general model: If w follows the specification proposed by Prelec et al. (1998), according to which $w(p) = e^{-(-log(p))^{\gamma}}$, and D follows the constant sensitivity specification proposed by Ebert and Prelec (2007), whereby $D(t) = e^{-(rt)^{\beta}}$, then for $s = 0$ the risk equivalent can be modeled as follows:

$$e^{-(-log(p^{\star}))^{\gamma}} = e^{-(rt)^{\beta}}.$$

Assuming that $\gamma = \beta$, we obtain $p^{\star} = e^{-r\ell}$. Here again, a single risk equivalent provides an estimate of the discount rate that is independent from utility curvature, probability weighting, and the hyperbolicity parameter. The Prelec et al. (1998) and Ebert and Prelec (2007) specifications have been found to provide a good fit to the data (e.g. Ebert and Prelec 2007; Stott 2006) and the assumption that $\beta = \gamma$ is also reasonable given typical estimates in the empirical literature (e.g. Ebert and Prelec, 2007, Booij et al., 2010). Specifically, Abdellaoui et al. (2019) report means of individual estimates of 0.74 for γ and 0.84 for β. Both a t-test and a Wilcoxon test do not reject the null-hypothesis that these are equal.

The conclusions reached above also extend to models specifically devised to study the interaction between risk and time. Baucells and Heukamp (2012) proposed a model that accounts for these two behavioral components. A prospect (x_t, p) that gives x at t with probability p is valued

$e^{-f(-log(p)+rt)}u(x)$ where f is an increasing function. Under this model, the indifference given by a risk equivalent becomes

$$e^{-f(-log(p^{\star})-rs)}u(x) = e^{-f(-r\ell)}u(x),$$

which also simplifies to $p^{*} = e^{-r(\ell-s)}$. Under this model, risk equivalents are again capable of eschewing the impact of probability weighting and non-exponential discounting.

If present bias is due to the non-linear distortion of a contract survival probability by the certainty effect, as pointed out by Halevy (2008), it may arise from the contrast with a present payout that obtains with certainty. Since in the current setup the present is made uncertain as well, this contrast is eliminated. Risk equivalents are thus also resistant to present-bias induced by the contrast between a truly certain immediate payoff and a nominally certain future payoff, to which some subjective probability of contract fulfillment is attached.

To summarize, under DEU risk equivalents can overcome the potential utility confound, and under several plausible behavioral extensions of DEU, risk equivalents can neutralize the impact of both probability distortions and of non-exponential discounting. Therefore, we should observe that discount factors measured through risk equivalents do not feature the usual deviations from exponential discounting, such as present bias or strongly decreasing impatience.

In our experiment we elicited choices over two main sets of outcomes x: $x = Rs400$ of money or $x = 2kg$ of tea, which at the time of the elicitations cost about 400 Rs in local markets. This allows us to directly compare the discount functions for money and a consumption good like tea.[2] Indifferences were elicited using choicelists. Within a list, the probability of winning in the sooner period was varied in steps of 0.05 between 0 and 1. Steps of 0.05 were large enough to allow for easy physical representations of the probabilities, but also small enough to pick up on relevant differences in discounting: a participant who switched between 0.45 and 0.5 would have an implied monthly discount rate of 25%, whereas a participant who switched between 0.5 and 0.55 would have an implied monthly discount rate of 21%. All choice problems were represented physically, laying out the monetary sums or amounts of tea to be won, and using colored balls to represent probabilities. The choice lists differed in terms of the timing of the sooner period and the delay intervals measured in months. Altogether, the risk equivalent experiments contained the following

[2]Like most empirical studies on discounting, we will assume throughout that subjects bracket choices narrowly—see Cohen et al. (2020), section 4.2, for a detailed discussion of this assumption.

delay pairs: $(s,\ell) = \{(0,2),(2,4),(4,6),(0,3),(3,6),(0,6)\}$, with delays indicated in months.

2.1.2 Additional measures from *present equivalents*

We also presented participants with standard *smaller-sooner* versus *larger-later* tradeoffs. In these decisions, participants generally face a tradeoff between obtaining an outcome x for sure at time ℓ and an outcome y for sure at time s. We elicit the sooner amount y^* that makes a participant indifferent between the two options. This has been the typical way economists elicit discount factors (Cohen et al., 2020). Under discounted expected utility, and with the crucial—sometimes implicit—assumption that utility is linear, the discount factor $\delta_\ell = D(\ell)$ is obtained from the ratio y^*/x. We compare discount factors elicited with this measure and under these assumptions to the discount factors elicited with risk equivalents.

We include two present equivalent tasks. The main task offers a tradeoff between Rs 200 for sure in 3 months and amounts Rs y for sure in the present. The second task offers a tradeoff between an option that offered two payments— Rs 200 for sure in the present and Rs 400 in 3 months—and a second option that offers amount Rs y for sure in the present. This latter amount can be used to separate utility curvature from the discount function within this type of task (Abdellaoui et al., 2013). In order to elicit the present equivalent y^*, the amount Rs y varies in steps of Rs 20. Participants also make choices over tea. The tea choices have a similar structure to the money choices, with Rs 400, Rs 200 and Rs 20 replaced by 2kg, 1kg, and 100g respectively. Under the assumption of discounted expected utility (which corresponds to discounted utility in this case where there is no risk), we obtain:

$$PE_1 : u(y_1^*) = D(3)u(200) \tag{2}$$

$$PE_2 : u(y_2^*) = u(200) + D(3)u(400). \tag{3}$$

Assuming linear utility, we can obtain a "raw" measure for $\tilde{D}(3)$ from PE_1 with $\tilde{D}(3) = y_1^*/200$. We now address the more general case of non-linear utility, assuming a power utility function $u(x) = x^\rho$. The two present equivalent tasks allow us to calculate the best-fitting parameter $\hat{\rho}$ that can rationalize the choices. Abdellaoui et al. (2013) provide non-linear least square estimates of the intertemporal utility function using similar structural equations. Here, we derive estimates of a power utility function using structural equations (2) and (3). Accounting for the role of the

Table 1: PEs and REs: Stimuli and descriptive statistics of discount factors

Choicelist number	Description	Money			Tea		
		1st quartile	median	3rd quartile	1st quartile	median	3rd quartile
1	$pe^* \sim 200_3$	0.55	0.75	0.95	0.35	0.75	0.85
2	$pe^* \sim 200_0 + 400_3$	0.48	0.73	0.88	0.38	0.63	0.88
3	$(400_0, p*) \sim 400_3$	0.53	0.73	0.88	0.53	0.72	0.88
4	$(400_0, p^*) \sim 400_6$	0.48	0.68	0.83	0.48	0.68	0.83
5	$(400_3, p^*) \sim 400_6$	0.53	0.73	0.88	0.48	0.73	0.88
6	$(400_0, p^*) \sim 400_2$	0.55	0.73	0.88	0.48	0.73	0.88
7	$(400_2, p^*) \sim 400_4$	0.58	0.73	0.88	0.53	0.73	0.88
8	$(400_4, p^*) \sim 400_6$	0.58	0.73	0.88	0.48	0.73	0.88

Summary of the different decisions made by participants. The Description column presents the choicelists as an indifference between either a certain amount in the present and a later, certain amount (choicelists 1 and 2) or an indifference between a lottery in an earlier period and a later, certain amount (choicelists 3-8). The certain amount in the present that makes the decision maker indifferent is called the *present equivalent:* $pe^\star$ The probability that makes the decision maker indifferent is called the *risk equivalent:* $re^\star$. The next set of columns show the distributions of the discount factors elicited by the different choicelists.

power utility parameter, the raw discount factor and the true discount factor are related as follows: $\tilde{D}(3) = D(3)^{1/\rho}$. When $\rho < 1$, $\tilde{D}(3)$ under-estimates the true discount factor $D(3)$. When $\rho > 1$, $\tilde{D}(3)$ over-estimates the true discount factor.

2.1.3 Comparison between REs and alternative methods

According to the review of Cohen et al. (2020), the majority of time-preference investigations use tradeoffs between smaller-sooner and larger-later payments, and only a minority of studies explicitly control for the curvature of utility. To our knowledge, there are four methods that allow to adjust measured discounting for utility curvature. Here we present a comparison of our method of risk equivalents to these other methods.

Historically, the first method that has been used consists in estimating utility from risky choices, and using it for correcting discount factors derived from present equivalents (Chapman, 1996; Andersen et al., 2008). A minimum of two tasks is needed to obtain a discount factor under this method—one tradeoff under risk for the identification of utility, and one tradeoff over time for the measurement of a present equivalent (hence the name double price list, or *DPL*). The method requires parametric assumptions for the utility function, and the deployment of structural estimation techniques—including an additional assumption on the correct stochastic choice model—to recover the model parameters. One drawback of this approach is thus that the results may be sensitive to the functional assumptions made.

Another class of methods recover utility curvature directly from inter-temporal tradeoffs. Andreoni and Sprenger (2012) propose a convex time budget method (CTB), where subject allocate tokens to a sooner and a later date at different interest rates. The method requires several allocation tasks to zoom in on a discount factor for a given time delay. The method further results in the frequent presence of corner solutions, whereby all tokens are allocated to either the sooner or the later date. The non-degenerate intermediate allocations needed to identify utility curvature have furthermore been found to contain frequent violations of rationality principles, which may question the reliability of the estimates based on such choices (Chakraborty et al., 2017). Abdellaoui et al. (2013) propose an alternative method that jointly obtains utility and discount factors from a trade-off between a series of present equivalents and inter-temporal *streams* of outcomes. Just like CTBs, the method requires multiple measurements and entails a higher degree of complexity for subjects than standard PEs for a single future outcome. Both methods furthermore require functional assumptions and structural estimation techniques to jointly recover the parameters. This may result in distorted inferences in the presence of noisy measurements, or if the functional assumptions are not warranted.

The closest approach to ours is the "direct measurement" (*DM*) proposed by Attema et al. (2016). It involves choices between inter-temporal streams providing a same outcome x at several time periods. Like ours, the method neutralizes the role of utility (but still requires parametric assumptions about the discount function). Other than our REs, however, the method requires continuous time and captures the shape of the discount function over a given time *interval*. Estimation of a the discount factor for a specific time period requires parametrization. Given the multiple time periods involved, the method is also arguably more complex for subjects than the simple tradeoffs involved when using risk equivalents.

Risk equivalents have a number of advantages over these methods. The first is parsimony. REs allow the discount factor to be measured from a single tradeoff or task. In comparison, a minimum of two tasks are required when using the DPL design or the method of Abdellaoui et al. (2013). CTBs of Andreoni and Sprenger (2012) require even more allocations to zoom in on the discount factor. Second, REs do not require any parametric assumption about the utility function nor its estimation. They thus eschew the risk of distortions deriving from mis-specification of the functional forms, as well as distortions deriving from noise in estimated utlity parameters. This also makes the method simpler to deploy for the analyst. Beyond simplicity for the analyst, the method is arguably also easy to understand for sujects. REs involve only one outcome, two time periods and one probability. In contrast, the DPL method involves more than two outcomes and a probability (for risk) and two outcomes and two time periods (for PEs). The method of Abdellaoui et al. (2013) involves at least three outcomes and two time periods. The CTB involves two outcomes, two time periods and an allocation percentage. And the DM involves one outcome and many time periods.

We implement the RE method and investigate if it produces the expected results regarding the neutralization of the role of utility while comparing it to the PEs of Abdellaoui et al. (2013). Including the other methods in the comparison would have considerably extended the duration of the experiment, and is thus left for further research.

2.2 Participants and procedures

The data set studied in this paper is part of a large scale experiment conducted in Ramanagara district, Karnataka state, India. The main experiment focused on the discounting of money, and only a subset of respondents dealt with choices involving money and tea. We first randomly sampled 23 villages from all rural villages in the district (excluding the larger, urbanized, villages located close to the Bangalore-Mysore highway). Once a village had been selected, we contacted the village head to announce and explain the experiments. We then constructed a household roster based on lists available at the village level, and using identifying information from IDs and ration cards to obtain unique identifiers for each household. The initial roster contained a questionnaire, and subjects were paid for their participation. We used this occasion to announce the subsequent time choice experiments, and to obtain written consent for the use of data from each participant. The payment for this initial questionnaire was administered with a three day delay, using the procedure that was later to be used in the experiment (see Section 2.2.2). This was meant to build trust in the fact that future payments would truly be administered.

All households in the sampled villages were invited to participate for ethical reasons. About 80% of all listed households participated, yielding a total sample of 2983 households for the full set of experiments. The reasons some listed households did not participate vary. The include deaths, households merging, or eligible participants being away from the village at the time of the experiment. One member from each household was randomly chosen to participate in the experiment. The experiments were administered in individual interviews by 20 enumerators. The enumerators were extensively trained on the tasks by the authors, and had the opportunity to accumulate experience in pilot sessions conducted in different villages. They were supervised in the field by a postdoctoral researcher speaking the local language, who monitored randomly selected experimental sessions to ensure that the same procedures were followed by all the enumerators.

From this sample, a subset of 728 households in 9 randomly sampled villages completed choices with both money and tea, allowing for a within subject comparison of discounting of these two attributes. In our analysis, we focus on the 587 respondents who completed all the 8 tasks described in table 1. The difference emerge from other subjects leaving before finishing the experiment, mostly to attend to issues on their farms. We present summary statistics of all participants in Table 2. About 60% of participants were women and about 55% of participants managed household food, including tea purchases. The average household had about 3 members and they had about 200g of tea on stock at the time of the experiment. Thus in the experiment, households were making decisions about amounts of tea that were ten times greater than the average household stock. The decisions over money represent a little bit more than average daily household income. In general, there are few differences between the participants who are included in the analysis compared to those who are excluded from the analysis. (Participants are included if they completed all tasks.) Participants who are excluded are slightly more likely to be older, to be heads of households, and to be farmers. This is consistent with the fact that other pressing duties were the main reason given for exiting the experiments early.

Table 2: Participant demographics

	Excluded	Included	p-value
Proportion women	0.57	0.64	0.15
Proportion head of household	0.48	0.34	0.00
Proportion farmers	0.37	0.28	0.04
Proportion with some secondary education	0.23	0.29	0.19
Proportion who manage household food	0.53	0.56	0.63
Proportion could raise 400 Rs in emergency	0.90	0.94	0.17
Mean age	45.06	40.99	0.02
Mean number of people in household	3.67	4.3 5.48	0.21
Mean amount of tea (in g) at time of experiment	207.24	209.95	0.92
Mean monthly household income (in Rupees)	9252.45	11969.31	0.64
Mean household savings (in Rupees)	9170.34	11868.58	0.52

Notes: Summary statistics of study participants. "Excluded" refers to participants who did not complete the experiment. P-values represent results for tests that proportions (test of proportions) or means (t-test) are different between the two groups.

2.2.1 Instructions, payments and comprehension checks

Decisions in the choice tasks were recorded on laptop computers, although enumerators used physical devices to carefully explain the choices and outcomes, including envelopes containing cash and samples of tea in the appropriate weights. In addition to envelopes containing cash in increments of Rs20, each enumerator had 20 sachets of tea, each weighing 200g, as well as a 2kg bag of tea. Each enumerator also had a bag containing 20 red and 20 yellow ping-pong balls.Participants' first decisions were concrete choices over the physical representations of tea or money and the lotteries represented by the red and yellow balls. There was no abstraction. For example, to represent a tradeoff between 2kg of tea 3 months in the future and a 75% chance of 2kg of tea today, the enumerators would physically lay out 2kg of tea on the left of a mat and on the right side of the mat. They would further lay out a bag containing 15 red balls and 5 yellow balls. The bag would be next to a 2kg bag of tea. Participants were invited to count and inspect the balls in the bag. Participants were then told that they could choose either the option on the right or the left. If they chose the option on the right, they would pick a ball out of the bag and if the ball was red, then they would get the 2kg bag of tea today. If they chose the option on the left, then the enumerators would come back in 3 months with 2kg of tea.

During training, participants made at least 3 such physical choices and played out the lotteries if chosen. Enumerators also talked them through the extreme choices. For example, "If the bag on the right was full of 20 yellow balls, what would the chance be of getting the 2kg bag of tea today?" To represent changing probabilities, enumerators took out one red ball from the sack, added a yellow ball and then shuffled the sack. Participants would again be invited to count and inspect the balls. Training lasted about 15 minutes on average (after introducing the project and obtaining consent). The tradeoffs were quite intuitive to participants when presented in physical form. When moving on from training, participants could choose to make their payoff-relevant decisions with the physical lotteries or with the abstract representations on the laptops. Almost all participants switched to the laptop representation after a few physical rounds. Enumerators evaluated whether participants comprehended the experiment before moving on to the payoff-relevant sections, and nobody was excluded because of a lack of comprehension.

We can further evaluate comprehension by looking at consistency in responses, violations of stochastic dominance, and the prevalence of focal points such as 0, 0.5 or 1. In terms of consistency, we find sensible correlations between discount rates elicited for money and tea. For example, the correlation for our main measures elicited using risk equivalents—the 3-month discount factors for money and for tea—is 0.43 (p-value < 0.01). Differences in discount factors for money and tea elicited with risk equivalents are not correlated with having secondary education (difference in means = 0.05, p-value for t-test = 0.89) or with income (correlation coefficient = -0.02, p-value = 0.53), two plausible indicators of sophistication in tasks such as ours. The same differences elicited with present equivalents are positively correlated with income. Only 2 subjects made at least one dominated decision in a choicelist (i.e. choosing to play the lottery in the present when there was a 0% chance of getting the money or tea) and fewer than 1% of allchoicelists contain an extreme choice (i.e. always choosing the certain payment in the future, even when an equivalent amount is available in the present). As Figure 4 in the Appendix shows, participant responses did not cluster around focal points such as 0, 0.5 and 1.

The order of the choicelists was randomized within treatments, and the order of treatments was also randomized. At the end of the experiment, subjects randomly drew a number to select the payoff relevant question. After this, subjects randomly drew a number corresponding to one of the rows within the choicelist. Their decision on this row was recorded for payment. The whole procedure from instructions to signing the payment certificates took the median subject just under one hour.

2.2.2 Future payments

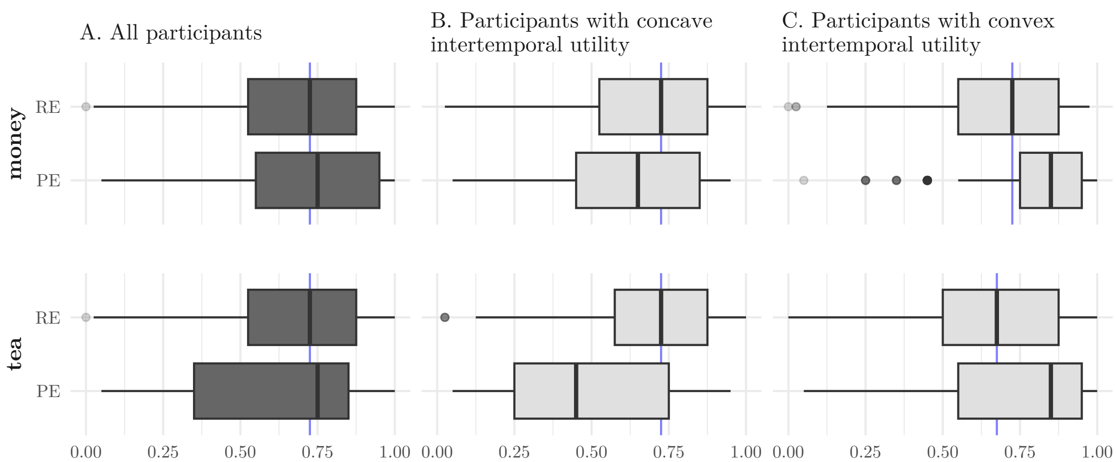
Trust in future payoffs is crucial when conducting intertempoal choice experiments. We took several measures to ensure such trust. For one, the soonest payoff was always administered with a three-day delay to avoid introducing differences in transaction costs. At the end of the experiment, the respondents received a certificate signed by the enumerator listing the amount to be paid, and the date at which the payment was due. The certificate contained the logo, the name, and the address of the office we maintained in the district (the WZB India Field Office). It also contained the address and logo of a local NGO facilitating our work, which was familiar to the participants, since they maintained operations supporting farming in the villages. The address and telephone numbers of both organizations were reported on the certificate, and participants were explicitly encouraged to get in touch in case they had any questions about the payments.

3 Results

In this section, we report a series of tests comparing the discount factors measured with risk equivalents to the discount factors measured with present equivalents. We first compare them to one-another, looking at differences across different profiles of estimated utility curvature; then we focus on differences between money and tea across the two methods; and finally, we examine deviations from exponential discounting across the two methods.

3.1 Discount factors from REs versus PEs

Figure 1: Distributions of 3-month discount factors measured by REs and PEs



Notes: Boxplots showing 3-month discount factors. Blue lines represent the median discount factor elicited with risk equivalents.

The boxplots in Figure 1 show summaries of the 3-month discount factors elicited by risk equivalents and present equivalents. These are presented separately for money and for tea. Panel A shows data for all participants while panels B and C show data for participants categorized as having either concave or convex intertemporal utility. These categorizations were made according to the best-fit parameters of a power utility function, using responses to the present equivalent tasks. (See section 2.1.2 for a description of this process. Descriptive statistics on the estimated utility parameters are reported in Appendix A).

The plots reveal several key features related to our hypotheses. First, focusing on Panel A, we observe that discount factors measured by risk equivalents are generally lower than those measured by present equivalents. This is confirmed by two-sided Kolmogorov-Smirnov tests (statistics and p-values: $0.19, p \leq 0.01$ for money and $0.19, p \leq 0.01$ for tea). The median three months discount factor for money is 0.73 measured by risk equivalents and 0.75 measured by present equivalents. For tea, the median discount factor is 0.73 measured by risk equivalents and 0.75 measured by present equivalents. The variances of the distributions of discount factors measured by risk equivalents are also significantly smaller (ratio of variances and p-values of an F-test for equality of variances: $0.79, p \leq 0.01$ for money; $0.61, p \leq 0.01$ for tea). One possible reason for this is that discount factors measured from PEs contain information about utility and discounting whereas those measured from REs contain information about discounting only.

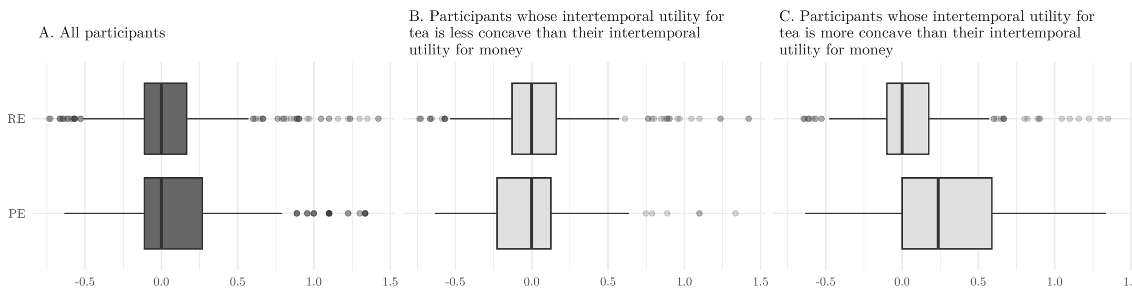
Discount factors measured from REs are slightly smaller than those measured from PEs. However, this aggregate result masks important differences when subjects are categorized according to their inter-temporal utility function. We observe that a slight majority of subjects exhibit convex inter-temporal utility—approximately 52% for money and 54% for tea (see Appendix). This echoes previous work (e.g l'Haridon and Vieider, 2019) that observes more convex risky utility in developing countries than in WEIRD (Western, Educated, Industrialized, Rich and Democratic) countries.

In Panels B and C, the distributions of discount factors are shown separately for participants classified as having concave or convex intertemporal utility as estimated according to Section 2.1.2. For money, the median discount factors measured by PEs are radically different depending on the curvature of utility: 0.65 among respondents with concave utility versus 0.85 among respondents with convex utility ($p < 0.01$). A yet larger gap is observed for tea: 0.45 vs 0.85 ($p < 0.01$). In contrast, much smaller gaps are observed from REs: we observe no gap for money (0.73 for both, $p = 0.85$) and a small gap (0.73 vs 0.68, $p = 0.01$) for tea.

Overall, we observe that non-linear utility clearly biases discount factors measured form PEs, whereas discount factors measured from REs are not impacted. The results in Figure 1 also suggest that tea is discounted more steeply than money, and that this difference may be greater when discounting is measured by present equivalents. We discuss good-specific discounting in the next section.

3.2 Discount factors for money and tea are more similar when measured with risk equivalents

Figure 2: Distributions of logged ratios of 3-month discount factors for money over tea



Notes: Logged ratios of the discount factor for money divided by the discount factor for tea. Ratios calculated at the individual level. Lowest 5% and highest 5% of ratios are excluded for plotting purposes, but all values included in statistical tests.

We can explore the extent of similarity between the two discount factors by examining the log of the ratio of discount factors measured for money over discount factors measured for tea. This value should be close to 0 if discounting is not good-specific, and deviations from 0 indicate differences in discount rates. A value greater than 0 would suggest that money is discounted less than tea. Panel A of Figure 2 shows boxplots of these log ratios for discount factors measured with risk equivalents and discount factors measured with present equivalents. We see that the distributions have similar means and medians. Despite these similar patterns at the aggregated level, accounting for heterogeneity reveals that the distribution of log ratios is much more dispersed for PEs than for REs. A variance test confirms that the variance is larger for the former than for the latter ($p < 0.01$). This means that there are larger differences between the discount rate for tea and the discount rate for money when measured with PEs. This also highlights the advantage of individual-level analysis, which captures differences that may be hidden at the aggregate level.

One theoretical explanation for this is that the discount rates computed from PEs do not control for utility curvature. Therefore, differences between utility for tea and utility for money translate into differences in discount rates when using PEs. This is not the case with discount rates measured from REs. In other words, some of the differences between discount rates for tea and money measured by PEs can be explained by differences in utility curvature between the two commodities. We investigate this relationship further in panels B and C of Figure 2.

We split the sample according to whether utility of tea is less (48% of the sample) or more (52% of the sample) concave than utility for money. Less concavity in the utility for tea should produce relatively larger discount factors for tea. This would imply that the log of the ratio of discount factors of money over discount factors of tea should be smaller than 0. Inversely, more concavity in the utility for tea should produce smaller discount factors for tea, i.e log ratios that are larger than 0. This is the pattern observed in Panels B and C of Figure 2. For individuals with less concave utility for tea (Panel B), the distribution of log ratios measured by PEs is shifted to the left of 0. Similarly, for individuals with more concave utility for tea (Panel C), the distribution of log ratios measured by PEs is shifted to the right. These shifts in distributions are statistically significant (Wilcoxon test, $p < 0.01$). Put differently, the log ratios measured from PEs are impacted by the difference of utility between tea and money. In contrast, the log ratios measured from REs for the two subgroups are stable to differences in utility (i.e. there is no difference in the distributions presented in Panels B and C for REs, Wilcoxon test, $p = 0.52$).

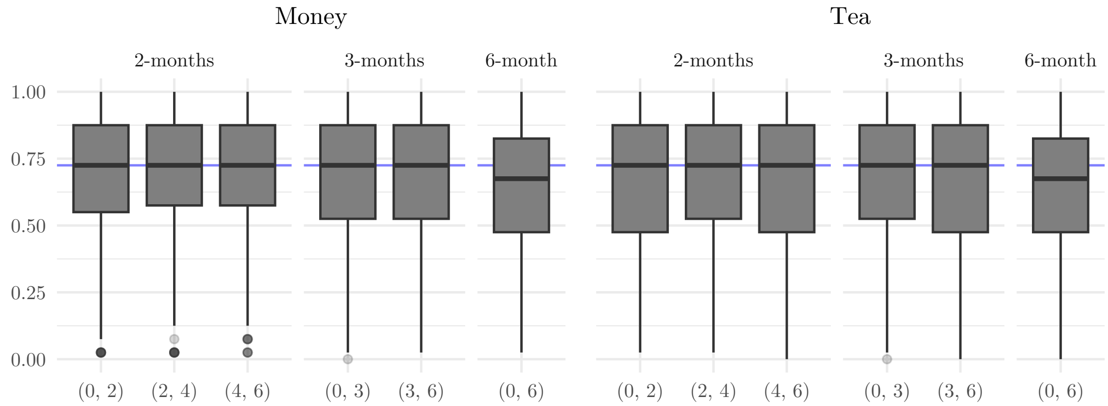
Overall, the analysis shows that REs produce discount rates that are more robust to differences in consumption goods. The analysis also accords with the hypothesis that one reason for this is the robustness to differences in utility curvatures across consumption goods.

3.3 Risk equivalents provide no evidence of present-bias, some evidence of subadditivity

We have established that compared to present equivalents, discount factors measured by risk equivalents are on average lower and more likely to be similar across commodities. Both these results are consistent with the hypothesis that risk equivalents neatly sidestep the problem of utility curvature. Given that they produce these cleaner measures, we now use risk equivalents to examine some broader patterns of discounting behaviour.

Figure 3 shows distributions of all discount factors measured by risk equivalents. The six different discount factors reflect measurements taken over different interval lengths and with different delays until the earliest period. The measures include two months discount factors at three different delays where i) the earlier period is in the present and the later period is two months away, ii) the earlier period is two months away and the later period is four months away, and finally iii) the earlier period is four months away and the later period is six months away. The measures also include

Figure 3: Boxplots of discount factors at different time delays and over different intervals



Notes: Boxplots of discount factors at different time delays and over different intervals. Intervals are represented on the x-axis as (t_s = sooner period, t_l = later period) where t is measured in months. Blue lines represent the median of the 3-month discount factors measured from ($t_s = 0, t_l = 3$).

three months discount factors at two different delays: iv) the earlier period is in the present and the later period is in three months and v) the earlier is three months away and the later period is six months away. And finally, the measures include vi) a six months discount factor measured with the earlier period at present and the later period six months away.

We consider the discount factors measured over the same time interval, but with dates where the early consumption period obtains in the future. These allow us to look for evidence of present-bias or subadditivity, the tendency to discount more when a long interval is split into multiple smaller intervals. Firstly, discount rates measured for the 0-3 period are not statistically different from discount rates measured for the 3-6 period. Discount rates measured for the 0-2 period are also not statistically different from discount rates measured for the 2-4 and 4-6 periods. Thus we find no evidence of any sort of decreasing impatience. This may not be surprising since our earliest consumption period was not immediately in the present (Imai et al., 2021). The absence of evidence about decreasing impatience is the same for money and tea. This is also consistent with the theoretical feature of risk equivalents of neutralizing time insensitivity—the source of decreasing impatience in the non-exponential discount function of Ebert and Prelec (2007). However, we do find strong evidence of subadditivity. If there is no subadditivity, the product of the 0-3 and 3-6 discount factors should be equal to the six months discount factor. For both money and tea, we can reject the null hypothesis that this is true (paired one-sided t-tests: $p < 0.01$ for both money and tea). A similar analysis and results apply to the two months discount factors.

4 Conclusion

We introduced the *risk equivalent* task to elicit discounting in a field setting and evaluated whether it achieves its main goal of providing a simple way to measure discounting while side-stepping the of utility curvature confound. Compared to the more traditional *present equivalent* task, discount factors measured by risk equivalents are not impacted by utility curvature. We also compared discounting for money and tea and found significant differences between the two. Importantly, these differences were mitigated in the measures made with risk equivalents, suggesting that differences in good-specifi utility curvature have a role to play in results that demonstrate differences in good-specific discounting. Finally, we used the risk equivalents to explore some general patterns of discounting behaviour, finding little evidence of decreasing impatience but significant evidence of subadditivity. Subadditivity is a decision pattern that is less explored in the literature but may have important implications for the design of contracts or payment plans when the number of sub-intervals between consumption periods might be varied.

Overall the paper demonstrates that risk equivalents are theoretically appealing, easy to implement, and flexible enough to measure complex patterns of intertemporal decision making. They could be a valuable addition to the toolkit of field studies.

References

Abdellaoui, Mohammed, Cédric Gutierrez, and Emmanuel Kemel, "Temporal discounting of gains and losses of time: An experimental investigation," *Journal of Risk and Uncertainty*, 2018, *57* (1), 1–28.

_ **, Emmanuel Kemel, Amma Panin, and Ferdinand M Vieider**, "Measuring time and risk preferences in an integrated framework," *Games and Economic Behavior*, 2019, *115*, 459–469.

_ **, Han Bleichrodt, Olivier l'Haridon, and Corina Paraschiv**, "Is there one unifying concept of utility? An experimental comparison of utility under risk and utility over time," *Management Science*, 2013, *59* (9), 2153–2169.

Andersen, Steffen, Glenn W Harrison, Morten Igel Lau, and Elisabeth Rustrom, "Eliciting Risk and Time Preferences," *Econometrica*, 2008, *76* (3), 583–618.

Andreoni, James and Charles Sprenger, "Estimating Time Preferences from Convex Budgets," *American Economic Review*, 2012, *102* (7), 3333–3356.

Attema, Arthur E, Han Bleichrodt, Yu Gao, Zhenxing Huang, and Peter P Wakker, "Measuring discounting without measuring utility," *American Economic Review*, 2016, *106* (6), 1476–94.

Augenblick, Ned, Muriel Niederle, and Charles Sprenger, "Working over time: Dynamic inconsistency in real effort tasks," *The Quarterly Journal of Economics*, 2015, *130* (3), 1067–1115.

Baucells, Manel and Franz H Heukamp, "Probability and time trade-off," *Management Science*, 2012, *58* (4), 831–842.

Booij, Adam S, Bernard Van Praag, and Gijs Van De Kuilen, "A parametric analysis of prospect theory?s functionals for the general population," *Theory and Decision*, 2010, *68* (1), 115–148.

Chakraborty, Anujit, Evan M Calford, Guidon Fenig, and Yoram Halevy, "External and internal consistency of choices made in convex time budgets," *Experimental economics*, 2017, *20*, 687–706.

Chapman, Gretchen B, "Temporal discounting and utility for health and money.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1996, *22* (3), 771.

Cohen, Jonathan, Keith Marzilli Ericson, David Laibson, and John Myles White, "Measuring time preferences," *Journal of Economic Literature*, 2020, *58* (2), 299–347.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde, "The robustness and pervasiveness of sub-additivity in intertemporal choice," in "Working Paper" 2017.

Ebert, J. E. J. and D. Prelec, "The Fragility of Time: Time-Insensitivity and Valuation of the Near and Far Future," *Management Science*, September 2007, *53* (9), 1423–1438.

Frederick, Shane, George Loewenstein, and Ted O'donoghue, "Time discounting and time preference: A critical review," *Journal of economic literature*, 2002, *40* (2), 351–401.

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

Imai, Taisuke, Tom A Rutter, and Colin F Camerer, "Meta-Analysis of Present-Bias Estimation Using Convex Time Budgets," *The Economic Journal*, 2021, *131* (636), 1788–1814.

Laibson, David, "Golden eggs and hyperbolic discounting," *The Quarterly Journal of Economics*, 1997, *112* (2), 443–478.

l'Haridon, Olivier and Ferdinand M Vieider, "All over the map: A worldwide comparison of risk preferences," *Quantitative Economics*, 2019, *10* (1), 185–215.

Loewenstein, George and Drazen Prelec, "Anomalies in intertemporal choice: Evidence and an interpretation," *The Quarterly Journal of Economics*, 1992, *107* (2), 573–597.

Prelec, Drazen et al., "The probability weighting function," *Econometrica*, 1998, *66*, 497–528.

Read, Daniel, "Is Time-Discounting Hyperbolic or Subadditive?," *Journal of Risk and Uncertainty*, 2001, *23* (1), 5–32.

Reuben, Ernesto, Paola Sapienza, and Luigi Zingales, "Time discounting for primary and monetary rewards," *Economics Letters*, 2010, *106* (2), 125–127.

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

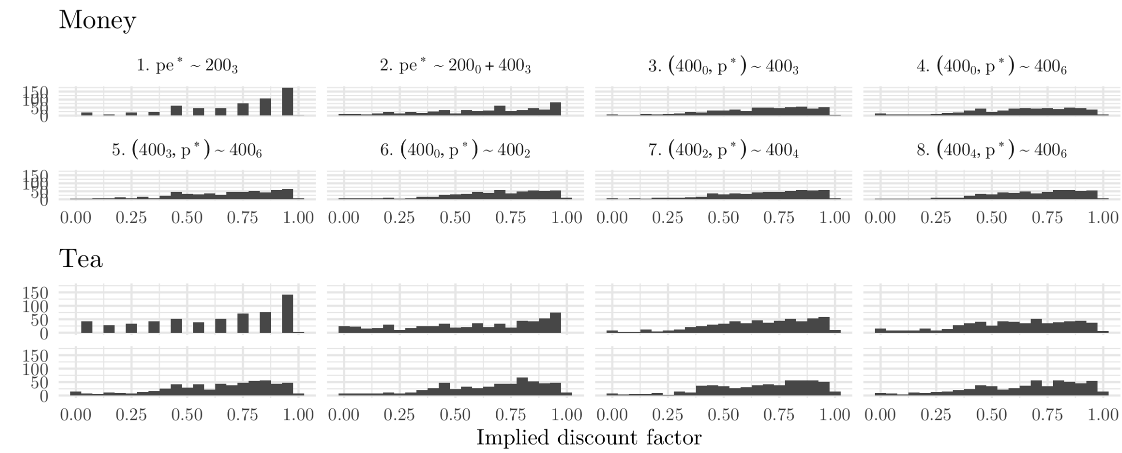
Tversky, Amos and Daniel Kahneman, "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and uncertainty*, 1992, *5* (4), 297–323.

Ubfal, Diego, "How general are time preferences? Eliciting good-specific discount rates," *Journal of Development Economics*, 2016, *118*, 150–170.

Appendices

Appendix A: Histogram of discount factors for main choicelists

Figure 4: Counts of participants at different discount factors for the main choicelists



Appendix B: Descriptive statistics on utility parameters estimated from present values

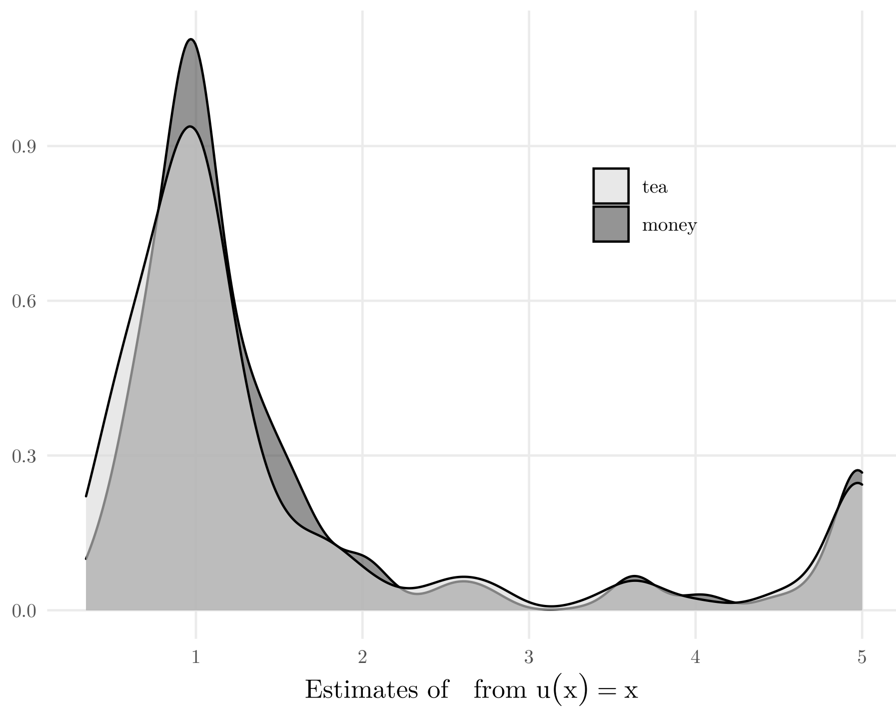
Notes: Histograms showing the counts of participants at the discount factors implied by their main *present equivalents* (choicelists 1 and 2) or r*isk equivalent* (choicelist 3). The width of each histogram bin is 0.05, reflecting that present equivalents were elicited in steps of Rs 20 (5\% of Rs 400) to be comparable to risk equivalents which were elicited in steps of 0.05

This appendix reports and illustrates statistics about the utility parameters for money and tea, and their relationship.

Table 3: Summary statistics of utility curvature parameter

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Money	0.36	0.88	1.04	1.62	1.56	5.00
Tea	0.34	0.80	1.02	1.61	1.64	5.00

Figure 5: Distribution of curvature parameter



Notes: the utility parameters for money and tea are weakly but significantly correlated. The Pearson correlation is 0.14 ($p < 0.01$).

Table 4: Counts of participants by convex or concave utility for money and tea

	Tea is convex	Tea is concave
Money is convex	192	138
Money is concave	124	129

Notes: Chi-square test rejects null that curvatures across money and tea are independently distributed (test statistic $= 4.48$; $p = 0.034$).