The Pervasiveness and Robustness of Nonadditivity in Intertemporal Choice

Thomas Dohmen[1], Armin Falk[2], David Huffman[3], Uwe Sunde[4]

June 4, 2022

Abstract

Present bias has been the most studied deviation from the canonical assumption of constant discounting, and has typically been modeled within the discounted utility (DU) model using a declining discount function. In our intertemporal choice experiments with representative samples, however, we find that the first-order deviation from the DU model is nonadditivity: discounting depends on how time horizons are subdivided. About 70 percent of individuals violate additivity, and the majority are subadditive – impatience is higher when time is subdivided. Nonadditivity cannot be rationalized by the DU model for any shape of discount function, including declining discounting. Robustness checks suggest that various confounds that have been raised about intertemporal choice experiments do not drive the results. The bias does not appear to be cognitive, as it is unrelated to cognitive ability or education. It is negatively correlated with wealth, suggesting an adverse impact on financial decisions. Among individuals who violate additivity, a non-trivial number still exhibit signs of present bias, even though this cannot be rationalized by the DU model. Thus, the findings suggest a value of models that can accommodate both nonadditivity and present bias. The results also have several methodological implications: (1) Commonly used designs to measure present bias may in fact be capturing subadditivity; (2) traditional experimental designs should be modified to allow testing for additivity; (3) different design choices for measures of time preference, which would be innocuous given additivity, can be expected to generate systematically different results if individuals are nonadditive.

Keywords: Time preference, hyperbolic discounting, self-control, dynamic inconsistency, intransitivity, sub-additivity.
JEL codes: D01, D90, D03, E21

[1] University of Bonn, ROA, IZA, and DIW; e-mail: tdohmen@uni-bonn.de;
[2] University of Bonn, briq, IZA, and CEPR; e-mail: Armin.Falk@briq-institute.org;
[3] University of Pittsburgh and IZA; e-mail: huffmand@pitt.edu;
[4] University of Munich, CEPR, and IZA; e-mail: uwe.sunde@econ.lmu.de.

1 Introduction

In recent decades, there has been a large amount of research in economics on the theoretical implications of present-biased preferences. In contrast to the canonical discounted utility (DU) model, which assumes a constant rate of impatience between any two adjacent periods, the idea of present biased preferences is that individuals may be more impatient when trade-offs involve consumption in the present period.[1] The main approach to modeling present bias has been to allow for some form of non-constant, declining discount function within the framework of the DU model, e.g., a hyperbolic or quasi-hyperbolic discount function (e.g., Phelps and Pollak, 1968; Ainslie, 1992; Loewenstein and Prelec, 1992; Laibson, 1997; O'Donoghue and Rabin, 1999).[2]

The approach of modeling present bias using the DU framework has in turn informed empirical methodology, with commonly used measures of present bias taking assumptions of the DU model as given (for surveys see, e.g., Frederick et al., 2012; Cohen et al., 2020). The traditional approach has been to use choice experiments with rewards (money or consumption), available either sooner or later, and contrast an individual's intertemporal trade-offs over two different time horizons. The by far most commonly used design, which we denote an overlapping horizon (OH) design, compares trade-offs over short and long time horizons that overlap and both start in the present, e.g., today to 6 months compared to today to 12 months. Such designs have typically found greater impatience for the shorter horizon, which has been taken as evidence of declining discounting and present bias.[3]

An alternative approach, which has been developed in psychology but which received very little attention in economics, is related to non-additivity of time perception. The usual way of modeling intertemporal choice the DU framework, including the conventional approaches of modeling present bias, implicitly assumes that discounting is unaffected by how a time horizon is subdivided. If, instead of being additive, discounting is subadditive, such that discounting is more extreme when time is measured in subintervals, this could be an alternative reason to find different discounting for the shorter time horizon in OH

[1] The canonical DU model is due to Samuelson (1936). O'Donoghue and Rabin (1999) coined the term present bias to capture this deviation from the canonical assumption.

[2] The earliest theoretical treatment of non-constant discounting in economics was by Strotz (1956).

[3] With a declining discount function, the average between-period discount rate will be higher for the shorter horizon, because the average does not include as many periods from far in the future.

designs. In fact, there are plausible psycho-physical mechanisms that can lead to subadditivity, for example subjective perceptions of time duration that are a concave function of objective duration.[4] Distinguishing declining discounting from subadditivity requires the addition of a third time horizon to an OH design, e.g., 6 months to 12 months in our example.

This paper argues for increased attention to nonadditivity of time discounting as an important deviation from the DU model. As we discuss below, nonadditivity has been found previously in the lab by a set of studies in psychology (e.g., Baron, 2000; Read, 2001; Read and Roelofsma, 2003; Read et al., 2005; Zaubermann et al., 2009), but this has had relatively little impact on the economics literature on intertemporal choice. We provide evidence that accounting for subadditivity is important for application in economics, because: (1) the phenomenon is very prevalent in representative samples – in our data even more prevalent than patterns that could be interpreted as present bias; (2) the distortion is large in magnitude; (3) the phenomenon is robust to a range of confounds that have been raised in the economics literature about intertemporal choice experiments with monetary rewards; (4) pervasive subadditivity has important methodological implications, because of how it changes interpretation of the measures economists typically use to capture time preference parameters, and for ways in which such designs can be improved; (5) subadditivity, like present bias, has important implications for economic behavior, but through a mechanism of how time horizons are perceived or framed rather than when they occur.

Our results are based on intertemporal choice experiments that nest traditional approaches to testing for present bias based on two time horizons, but also allow testing additivity, because they include appropriately chosen third time horizons. We use two representative samples of the German population (N=500; N=1,500) in order to assess whether nonadditivity is important in the sense of being prevalent.[5] The experiments involve real monetary stakes, and check robustness to different levels of stakes and different

[4] For discussions see, e.g., Read (2001); Ebert and Prelec, (2007); Zauberman et al. (2009). Intuitively, this may lead people to be more impatient for shorter time horizons (subintervals), because subjective duration seems relatively long, and less impatient for longer time horizons, because subjective duration increases less than objective duration.

[5] In light of evidence that preferences can vary across cultures and countries (e.g., Henrich et al., 2010; Falk et al., 2018), we do not claim that the precise frequencies we observe will generalize to non-German populations, or even to other random samples of Germans. Broadly speaking, however, we think our findings imply that one should expect a substantial prevalence of nonadditivity in most populations and samples.

methods of payment delivery.

We first present aggregate results that are consistent with subadditivity. In both of our datasets, median discount rates are similar for short time horizons, regardless of when they occur, contrary to present bias and more consistent with constant discounting. At the same time, median discount rates are substantially higher for short horizons than for long horizons, contrary to constant discounting, but consistent with subadditivity. The magnitude of the subadditivity effect is also large. For example, in one of our datasets, the median choices made for horizons today to 6 months and 6 months to 12 months imply 30 percent greater discounting (when computed over an annual time horizon) than do the choices observed for the horizon today to 12 months. The degree of subadditivity appears to be even larger as the subinterval gets shorter, e.g., the annual discount rate for today to 1 month is about 50 percentage points higher than for today to 12 months (but very similar for today to 1 month compared to 12 months to 12 months).[6] The results are consistent with a very simple (descriptive) model of subadditivity, in which the mechanism is subjective perceptions of time duration being a concave function of objective duration.

We discuss arguments, and provide empirical evidence, that suggest the results are not driven by a set of confounds that have been discussed in the economics literature about intertemporal choice experiments with monetary rewards: Curvature of utility; arbitrage between lab and field interest rates; savings motives; imperfect credibility of payments; or anticipated taste shocks. Another concern could be that the results reflect some other bias or arithmetic heuristic specific to intertemporal choice experiments using monetary rewards, but we argue and provide evidence that the subadditivity we find cannot be explained by two such candidate explanations, exponential growth bias, or a rule of thumb heuristic involving demanding double the money to wait twice as long.[7]

[6] In our data the effect of subadditivity is substantially larger than another deviation from the DU model that has been studied in the previous literature, the so-called "magnitude effect" (see, e.g., Thaler, 1981). While discount rates do decrease with higher stakes in our experiments, consistent with the magnitude effect, this is a relatively small compared to the impact of subdividing time horizons on discounting.

[7] Some recent studies have used consumption rather than money to measure time discounting, which relaxes some of the assumptions made with monetary experiments, but requires other assumptions (for a discussion see Cohen et al., 2020). Since our robustness checks suggest that the results are not explained by features specific to monetary-base experiments, we do not see a reason to expect that additivity violations will be less prevalent with consumption based experiments, although this is an interesting direction for future research. Regardless, monetary-based experiments have been used extensively in the existing literature, and continue to be used in the present, so it is important to understand how to interpret behavior in such experiments.

We next turn to an individual-level analysis, and show that the same conclusions hold. About 70 percent of the general population violates additivity, with most being subadditive, even with a conservative definition that treats small deviations as additive. By contrast, there are very few individuals, about 5 percent, who can be classified as declining discounters, i.e., exhibiting present bias and satisfying additivity.[8] Notably, however, some individuals who violate additivity still have choice patterns that correspond to the idea of present bias. In our data about 16 percent of the sample falls into this category. Other individuals who are nonadditive show signs of future bias, meaning they are more patient when choices involve the present, and still others have no bias related to choices involving the present. Thus, the fact that subadditivity is pervasive does not mean that these other biases are not present as well. It just means that they cannot be rationalized by the DU model.

Our individual analysis also explores correlates of nonadditivity, to shed light on potential mechanisms, and also the relationship of nonadditivity to economic outcomes. We find that subadditivity is largely the same across different demographic groups, underlining that this is a pervasive feature of decision making. There is no significant relationship to cognitive ability or education, suggesting that the mechanism underlying nonadditivity is not just cognitive mistakes. The one measure that does predict fewer additivity violations is a proxy for financial sophistication, which could indicate that financial market experience or sophistication may help de-bias individuals, but subadditivity is still the predominant choice pattern even among those who are sophisticated. Additivity violations are associated with worse financial outcomes in terms of income and wealth, consistent with nonadditivity leading to distortions away from optimal financial decision making.

Our analysis concludes with a discussion of three sets of methodological implications. A first implication is that different, commonly used approaches in the economics literature for testing between different discount functions can be expected to find systematically different results, if individuals are nonadditive. For example, if individuals are subadditive, OH designs will tend to find declining discounting, because this is confounded with present bias. A less-commonly used approach – which we denote a shifted horizon (SH) design – compares discounting over two time horizons over the same length, with one shifted in

[8] Roughly similar proportions of the population could be classified as increasing or constant discounters satisfying additivity.

time, e.g., today to 6 months versus 6 to 12 months. In such designs, subadditivity will not give the appearance of declining discounting, because time horizon length is constant, and thus one would expect such designs to be less likely to find evidence interpreted as declining discounting. A third type of approach, the convex time budget (CTB) method, has been adopted by a number of recent studies, and typically uses four or more time horizons of various lengths and timings, which typically nest both OH and SH comparisons. Results from CTB studies are based on pooling across these multiple horizons. We discuss how results in CTB designs can be influenced by subadditivity depending on the particular mix of horizons chosen. The results of our experiments are strongly consistent with the predicted variation in results across OH and SH measures, and our survey of the previous literature also finds that OH studies almost always find declining discounting, whereas results for SH and CTB designs are more mixed. This explanation for varying results across studies has not been discussed in the literature.

A second methodological implication is that it is important to implement designs with appropriately chosen time horizons to allow testing for nonadditivity. Otherwise, interpreting behavior in such designs as consistent with the DU model may be incorrect. For example, our data show that among those who are present biased according to an SH comparison, accounting for the third time horizon reveals that relatively few individuals satisfy additivity. The behavior of the majority of individuals who are present biased according to the SH design can therefore not be rationalized by the DU model with declining discounting. Testing for additivity is also important for whether results of OH or CTB designs can be rationalized with the DU model. Our findings on the pervasiveness and robustness of nonadditivity thus offer a new lens through which to interpret the conclusions of previous studies, about the ability of the DU model with a particular shape discount function to explain intertemporal choice.

A third methodological implication is for approaches used by economists to estimate quantitative values of time preference parameters for economic models: If individuals are nonadditive, the specific lengths of time horizons chosen can influence parameter estimates. We show theoretically that this is the case for studies using a single time horizon to calibrate the exponential discount rate, with subadditivity implying that longer time horizons will generate lower estimates of the discount rate. We also show implications of subadditivity for studies using OH or SH designs to calibrate parameters of the quasi-hyperbolic model. The particular time horizons chosen can matter for the estimate of β and δ in OH designs, and for the estimate of δ in SH designs. We also discuss implications for Convex Time Budget (CTB) designs to estimate preference parameters in the quasi-hyperbolic model, while accounting for curvature of utility. We discuss how the particular set of time horizons chosen can matter for how nonadditivity will influence preference parameter estimates in this type of design.

Our findings complement several previous literatures. One is the theoretical literature on intertemporal choice. While much progress has been made on working out the theoretical implications of present bias in the context of the DU model (seminal papers include Phelps and Pollak, 1968; Ainslie, 1992; Loewenstein and Prelec, 1992; Laibson, 1997; O'Donoghue and Rabin, 1999), our results highlight the value of alternative frameworks that can capture nonadditivity (see, e.g., Rubinstein, 2003; Scholten and Read, 2006; Ok and Masatlioglu, 2007, Vieider, 2021), and point to the value of future theoretical work on the implications of nonadditivity for economic behavior. Notably, nonadditivity has more radical theoretical implications than present bias, as it violates assumptions of the DU model and cannot be accommodated by any shape of the discount function.[9] Given that we still find evidence consistent with present bias, despite violations of additivity, our findings also point to the value of non-DU models that can accommodate both pheonomena.

Our paper also complements a previous literature on nonadditivity of time discounting. While a number of studies in psychology have found evidence in the lab that time discounting may be subadditive (Baron, 2000; Read, 2001; Read and Roelofsma, 2003; Read et al., 2005; Zaubermann et al., 2009), studying nonadditivity has not gained much traction in economics, and the phenomenon has not informed interpretation or design of preference measures in economics. Indeed, our survey of 83 experimental studies on intertemporal choice published in economics in the last two decades (see Appendix A) finds that, aside from the literature referenced above, only two studies test for additivity violations.[10] The possibility of nonadditivity is mentioned in some important survey papers

[9] Indeed, non-additivity implies a violation of transitivity over time, under the standard workhorse assumptions of time-separable and stationary utility.

[10] McAlvanah (2010) finds subadditivity in two laboratory experiments with student subjects (N=55; N=41), and Vieider (2021) finds evidence of subadditivity in an experiment with student subjects (N=175). The 83 papers are from the economics literature in the sense that they are either published in economics journals or involve one or more economists as authors.

that are heavily cited (Frederick et al., 2002; Cohen et al., 2020), so nonadditivity is not unknown in economics. Rather, the fact that the empirical literature gives the issue almost no consideration seems to reflect an implicit assumption nonadditivity is not a major issue, either because it is not prevalent and affects only a few people, or because it is small in magnitude. Our paper is complementary because it demonstrates the pervasiveness and substantial magnitude of subadditivity in representative samples, the robustness of additivity violations to the methodology of experimental economics (e.g., real incentives) and to various critiques of intertemporal choice experiments, and provides the first evidence on whether subadditivity is related to cognitive ability and economic outcomes. Another difference is our investigation of whether patterns such as present bias are found among individuals who are nonadditive; while the previous literature on nonadditivity mainly focused on the conclusion that choices could not be rationalized by declining discounting in the DU model, our results call for retaining the idea of present bias, albeit modeled in some other way.

The paper also contributes to the empirical literature on intertemporal choice experiments in economics (see our literature survey in Table A.1, as well as surveys by Frederick et al., 2012, and Cohen et al., 2020). Our findings show the pervasive violation of a key identifying assumption of standard approaches to measuring preferences in the context of the DU model. As such, the research suggests a new lens through which to interpret previous findings. Subadditivity also offers an explanation for heterogeneity in results across studies, according to different types of design choices, and suggests specific changes to experimental designs used in the literature that can be valuable for clarifying what types of models can rationalize the observed behavior.

The rest of the paper is organized as follows. Section 2 describes the data sets, treatments, and behavioral predictions of standard discounting models. Section 3 presents the aggregate results, and Section 4 provides an analysis of individual heterogeneity. Section 5 discusses methodological implications, and Section 6 concludes.

2 Design of the Experiment

2.1 Data Collection and Experimental Procedures

Our analysis uses two data sets. One data set, which we call the SOEP data involves a sub-sample of participants in the German Socio-Economic Panel (SOEP), a large panel data set for Germany (for a detailed description of the SOEP see Goebel et al., 2019). The second data set is the SOEP Cross Sectional Study, which we call CSS data for short. This data set involves a separate sample of individuals that was collected by the SOEP administration as part of the process of “pretesting” questions for potential use in the SOEP core survey.

Data collection for both data sets was done by the same professional surveying company that administered the SOEP in these years, using the same sampling procedure as for the SOEP.[11] Both the CSS and SOEP samples were constructed so as to be representative of the adult population, age 17 and older, living in Germany.[12] Subjects were visited by interviewers in their own homes. In total the CSS data include 500 subjects who participated in the intertemporal choice experiments, and the SOEP data include 1,503 such subjects.

Participants in our studies went through a computer-assisted personal interview (CAPI) conducted with a laptop. First, subjects answered a detailed questionnaire. Topics included demographic characteristics, financial situation, health, and attitudes. The full questionnaire, in German and translated into English, is available upon request. Second, at the end of the questionnaire, subjects were invited to participate in a paid experiment.

The first step in the experimental procedure involved the experimenter presenting subjects with some example choices, explaining the types of choices the subject would face, and how payment would work. Once there were no more questions, the experiment began. An example of the script and instructions used in the experiments is presented in Appendix G below, translated from German into English.

Our experiments were designed to give a measure of the annual Internal Rate of Re-

[11] For each of 179 randomly chosen primary sampling units (voting districts), an interviewer was given a randomly chosen starting address. Starting at that specific local address, the interviewer contacted every third household and had to motivate one adult person aged 17 or older to participate. For a detailed discussion of the random walk method of sampling see Thompson (2006).

[12] Respondents had to turn 18 during the year of the interview to be eligible.

turn (IRR) needed to induce an individual to wait for a given time horizon. As we discuss below, in the context of the canonical DU model, the IRR is under certain assumptions informative about the discount factor. A time horizon $T_{s,t}$ is defined by a starting date s, and an ending date t. For a given horizon an individual made a choice between an early payment, X, available at the start of the horizon, and series of larger, later payments, Z, available at the end of the horizon (in a slight abuse of notation, in this section we suppress superscripts for early and late payments indicating time horizon). For example, in horizon $T_{0,12}$, the early payment is at time 0, and the later payment is 12 months in the future. In all choices for a given horizon, the amount of the early payment, X, was held constant, but the later payment, Z, was larger in each subsequent choice.

For most time horizons, the value of the delayed payment in the first choice was calibrated to be consistent with an annual IRR of 2.5 percent, assuming semi-annual compounding, and each subsequent value of Z implied an additional 2.5 percentage point increase in the annual rate of return, up to a maximum of 50 percent.[13] To achieve the same menus of annual IRRs for horizons of different lengths, we varied the menus of delayed payment amounts appropriately. Having the same range of possible IRRs for different horizons was attractive in order to reduce problems of censoring that might interfere with comparing IRRs across time horizons.[14] In one treatment, we implemented choices with a coarser measurement of the IRR, in steps of 5 percentage points, but in this case measured annual IRRs as high as 100 percent.

[13] We chose semi-annual compounding of the annual interest rate as a compromise between the two types of compounding German subjects are most familiar with: quarterly compounding on typical bank accounts, and annual reports on the rate of return from savings accounts, pension funds, or stock holdings. Using semi-annual compounding also helps avoid prominent round numbers in the choices, which could potentially influence switching choices.

[14] An alternative design choice would have been to hold constant the menus of delayed payment amounts across horizons of different lengths, which would have necessarily entailed varying the menus of annual IRRs offered for short versus long horizons (and also measuring the IRRs in coarser intervals for the short horizons). Like our design, this approach involves varying two parameters across horizons of different lengths, albeit different parameters. This alternative approach would necessarily involve having upper bound IRRs that are substantially smaller for long horizons that short horizons (for a horizon twice as long the upper bound would be half as large). We did not use such a design because we were concerned about the possibility of censoring at the upper bound for the longer horizons, which would have made it difficult to know how the censored IRR magnitude compared to potentially higher magnitudes elicited for shorter horizons. Our survey of the literature indicates that both types of designs have been used in studies measuring discounting and present bias, so both have precedent in studies on intertemporal choice. Among the few studies that test additivity, some have used the design we use (e.g., Scholten and Read, 2006), others the alternative (e.g., Read, 2000), and both have found evidence of subadditivity. This suggests that this particular design choice is not crucial for finding violations of additivity. It would be interesting for future research on nonadditivity to compare both types of designs in the same framework.

We obtain a measure of the IRR needed to induce an individual to wait for a given time horizon by observing the smallest value of Z that induces them to wait. More precisely, we obtain upper and lower bounds for the annual IRR, separated by 2.5 percentage points, due to the discrete variation in late payment amounts; in our analysis, we focus on lower bounds. Across treatments, we varied s and t as well as the amounts X and Z.

IRRs were elicited in an incentive compatible manner: Subjects were presented with the different choices, one at a time, on the computer screen, and subjects knew that one choice would be randomly selected at the end of the experiment and implemented. Furthermore, subjects knew that at the end of the experiment a random device would determine whether they were actually paid, with the probability of being paid equal to 1/7 in the CSS and 1/9 in the SOEP data. This procedure gave subjects an incentive to choose according to their true preferences in each choice situation.

A key feature of our design was the high credibility of payments. One aspect that contributed to credibility of payments for both data sets was the fact that the agency conducting the experiments is well-known and trusted by the German public.[15] Interviewers also left their contact details at the end of the experiment, making it easy for subjects to contact the interviewing agency, but there were no reports, from any of the interviewers, about subjects expressing concerns regarding credibility of payments. There is an even stronger argument for credibility in the case of our SOEP data, however, because all participants were members of the SOEP panel itself. Unlike participants in most intertemporal choice experiments, these individuals were in a long-term relationship with the individual surveyor conducting the experiment. This implies that payments were highly credible regardless of timing. We did also use a "front-end delay" approach to achieving equal credibility of early and late payments that has often been used in the literature (see Coller and Williams, 1999). Specifically, all payments arrived by mail after the experiment. Thus, early payments did not have any special credibility arising from being paid during the experiment session itself.

A further robustness check regarding credibility is possible because we varied the payment procedure across data sets. Specifically, for the CSS data set, all payments were mailed on the day after the interview and thus would arrive within at most two days due

[15] The agency employed at the time was known as a reliable and reputable polling company by German public television stations; television news programs regularly feature results from the agency's opinion polls on social and political issues.

to the well-known two-day guarantee for delivery by the German postal service. Checks for immediate payments could be cashed immediately, once they arrived, while checks for payments in the future were post-dated through a special arrangement with the issuing bank and could only be cashed at the specified time. For the SOEP data, by contrast, payments were also sent by mail, but the timing of the mailing reflected the timing of the payments, i.e., checks for immediate payments were mailed immediately after the experiment and arrived within two days, while checks for later payments were mailed punctually on the corresponding later date. In this case neither check was post-dated. If credibility concerns are an important issue for subjects, one would expect to see that changes in the payment procedure affect results. In our analysis we compare results across the two data sets.

A final note on the design concerns the length of the front-end delay. Many previous studies that tested discounting assumptions have featured front-end-delays ranging from one day to as long as one month (e.g., Meier and Sprenger, 2010; Harrison et al., 2002). While useful for equalizing credibility, such a delay might tend to reduce "immediacy" of early payments. This could matter for the predictions of some non-constant discounting models that assume a discrete drop in discount rates between the present and the future, such as the quasi-hyperbolic discounting model. The verdict is arguably still out on exactly how quickly such a drop might occur, but if it is assumed to occur between the present and the next day, then both early and late payments in our experiments are beyond the time frame in which present bias exerts its influence.[16] In such a case, the model would make still make testable predictions, but these would be the same as those of the constant discounting model, as discussed below. Making a trade-off between credibility and immediacy, we chose the shortest possible front-end delay compatible with avoiding a same-day credibility problem.[17]

[16] Augenblick and Rabin (2019), for example, provide evidence that is more consistent with quasi-hyperbolic discounting than hyperbolic discounting, but because all future dates in their study are at least four days in the future, it is not clear exactly where the drop occurs between today and four days into the future.

[17] Mainly for the purposes of concreteness, so as to avoid having to say 1 or 2 days repeatedly, the experimental instructions told subjects that immediate rewards would be referred to as being received "Today"[quotes included]. At the same time, the instructions were very clear that all rewards would arrive after the experiment by post, and that: "Today means you can cash the check you receive by post immediately." Ultimately, we see little evidence that this wording led to differential behavior in early versus later time horizons. For example, as shown below, the observed impatience is similar regardless of whether the time horizon involves early payments "Today" or involves early payments in 12 months.

Table 1: Summary of Treatments

Measure	Data set	Sub-sample	Early payment (in Euro)	Upper-bound IRR	Obs.
$T_{0,12}$	CSS	n.a.	100	52.5%	500
$T_{0,6}$	CSS	n.a.	100	52.5%	500
$T_{6,12}$	CSS	n.a.	100	52.5%	500
$T_{0,12}$	SOEP	1 & 2	200	52.5%	977
$T_{0,6}$	SOEP	1	200	52.5%	490
$T_{0,1}$	SOEP	2	200	52.5%	487
$T_{0,1b}$	SOEP	3	200	105%	526
$T_{12,13}$	SOEP	3	200	105%	526

2.2 Treatments

Table 1 summarizes the various treatments (the full sets of choices for each time horizon are provided in Appendix H). As shown in the table, the CSS data involved three different measures of annual IRR for each subject: 0 to 6 months ($T_{0,6}$), 0 to 12 months ($T_{0,12}$), and 6 months to 12 months ($T_{6,12}$). This design was chosen in order to nest the traditional approaches of OH comparisons ($T_{0,6}$ vs. $T_{0,12}$) and SH comparisons ($T_{0,6}$ vs. $T_{6,12}$) for each individual, and to allow testing for additivity by comparing total discounting implied by the two six-month subintervals to discounting implied by $T_{0,12}$. The order of the treatments in the CSS data was randomized across individuals. The early payment was always 100 euros, and the largest delayed payment always implied an annual IRR of 50 percent (compounded semi-annually) for waiting the specified length of time. If individuals never chose the later payment, their IRR was right-censored, and coded as having a (lower-bound) value of 52.5 percent.

In the SOEP data we used higher stakes, different treatments, and different payment procedures, in order to assess whether the results from CSS replicate with different parameter values and design features. The early payment was always 200 euros in the SOEP and thus stakes were higher than in the CSS, even accounting for the lower payment probability of 1/9 rather than 1/7.[18] The SOEP also differed from the CSS in having only two treatments for each individual (see Table 1). This was due to time constraints in

[18] The relatively large nominal values involved in the experiment help mitigate distortions due to subjects rounding delayed payment amounts up to the nearest dollar. See Andersen et al. (2011) for a discussion of this issue.

the survey implementation. We varied these treatments across sub-samples, however, in a way that complements the CSS study design and led to a total of five different treatments.

Specifically, in the SOEP data, the first sub-sample of 490 individuals had measures for 0 to 6 months ($T_{0,6}$) and 0 to 12 months ($T_{0,12}$). This allows an OH comparison with time horizon lengths that are directly comparable to the CSS data. The second sub-sample of 487 were asked about 0 to 1 month ($T_{0,1}$) and 0 to 12 months ($T_{0,12}$), shedding light on how OH comparisons change as the discrepancy in horizon length increases. For the third sub-sample of 526 we measured discounting for 0 to 1 month ($T_{0,1b}$) and 12 to 13 months ($T_{12,13}$), to allow an SH comparison, with a relatively large time delay between horizons. The measures for the third sub-sample were also different because IRRs were measured in steps of 5 percent rather than 2.5 percent, and the upper-bound IRR in each horizon was 105 percent rather than 52.5 percent. This was designed to help explore the nature of discounting without potential censoring at 52.5 percent. We denote the one-month measure in this sub-sample $T_{0,1b}$, to distinguish it from $T_{0,1}$ in the second sub-sample.

Order was predetermined in the SOEP data: for the first two sub-samples, the $T_{0,12}$ measure was always elicited first; for the third sub-sample, $T_{0,1b}$ was elicited first. A random device on the computer selected whether an individual was assigned to the first, second, or third sub-sample experiments.

2.3 Behavioral predictions of the DU model

In this section we present predictions based on the traditional DU model. To derive the predictions we employ assumptions that are often maintained, explicitly or implicitly, in the literature using intertemporal choice experiments. These include workhorse assumptions like time separability, but also the following: (1) equal credibility of payments regardless of timing or amount; (2) utility is locally linear; (3) people treat monetary payments like consumption opportunities and do not engage in arbitrage between the interest rates offered in the experiment and market interest rates; (4) a time-stationary period utility function that does not vary with calendar date. The recent literature has raised concerns about assumptions (1) to (4), so later in the analysis we discuss whether our results might be driven by the failure of one of these assumptions.

We illustrate the predictions by considering, without loss of generality, an example with three time horizons, $T_{0,6}$, $T_{0,12}$, and $T_{6,12}$, corresponding to the structure of the CSS data. The early payment is always 100 for each horizon, compounding occurs once every 6 months, and for simplicity a period is assumed to be 6-months long. We assume for now that early payments at time 0 are literally available on the day of the experiment, when preferences are measured.

When making decisions in $T_{0,6}$, $T_{0,12}$, and $T_{6,12}$, subjects decide for the early or late payment depending on whether or not the offered *annual* rate of return r in a given choice is sufficiently attractive to induce waiting. Thus, decisions involve the following comparisons:

$$T_{0,6} : \left(1+\frac{r}{2}\right) 100 \lesseqgtr \mathbf{Z}^{\mathbf{T_{0,6}}}; \quad T_{0,12} : \left(1+\frac{r}{2}\right)^2 100 \lesseqgtr \mathbf{Z}^{\mathbf{T_{0,12}}}; \quad T_{6,12} : \left(1+\frac{r}{2}\right) 100 \lesseqgtr \mathbf{Z}^{\mathbf{T_{6,12}}}$$

where $\mathbf{Z}^{\mathbf{T_{0,6}}}$, $\mathbf{Z}^{\mathbf{T_{0,12}}}$, and $\mathbf{Z}^{\mathbf{T_{6,12}}}$ denote sets of later payments available for the corresponding time horizons. We denote by $Z^{T_{0,6}}$, $Z^{T_{0,12}}$, and $Z^{T_{6,12}}$ the smallest element of these sets of later payments, which make the individual indifferent between the earlier or the later payments for the corresponding time horizon. Observations on $Z^{T_{0,6}}$, $Z^{T_{0,12}}$, and $Z^{T_{6,12}}$ obtained from the experiments establish the points of indifference for each horizon, and define the internal rates of return:

$$\left(1+\frac{IRR^{T_{0,6}}}{2}\right) = \frac{Z^{T_{0,6}}}{100}; \quad \left(1+\frac{IRR^{T_{0,12}}}{2}\right)^2 = \frac{Z^{T_{0,12}}}{100}; \quad \left(1+\frac{IRR^{T_{6,12}}}{2}\right) = \frac{Z^{T_{6,12}}}{100} \quad (1)$$

We neglect the fact that the delayed payment is actually a variable measured on a discrete grid in the experiment, and thus that we can infer only a range for the IRR; this has no consequences for the qualitative predictions, and eases exposition.[19]

2.3.1 Additivity

The DU model makes predictions about what we should observe for the IRRs measured in our experiments. Some of these predictions do not depend on the shape of the discount function, and some do. An example of the former type of prediction is additivity, which we discuss in this subsection.

[19] The lowest delayed payment that is preferred establishes an upper bound for the IRR, while the largest delayed payment that is not preferred establishes the lower bound. One can think of the predictions as being derived based on lower bounds everywhere (or equivalently in terms of upper bounds).

The DU model assumes a rate of time preference for each time period, which we denote by ρ_t, with $t \in 1, 2$ for periods 1 and 2 in our example. The rate of time preference may or may not be constant across periods. In this case an individual's points of indifference in $T_{0,6}$, $T_{0,12}$, and $T_{6,12}$ are determined by:

$$(1+\rho_1) = \frac{Z^{T_{0,6}}}{100}; \qquad (1+\rho_1)(1+\rho_2) = \frac{Z^{T_{0,12}}}{100}; \qquad (1+\rho_2) = \frac{Z^{T_{6,12}}}{100}. \tag{2}$$

This shows that in the DU model, regardless of whether discounting is constant ($\rho_1 = \rho_2$), increasing ($\rho_1 < \rho_2$), or declining ($\rho_1 > \rho_2$), the following relationship must hold between the monetary amounts the individual needs to induce waiting over the different time periods:

$$\frac{Z^{T_{0,12}}}{100} = \frac{Z^{T_{0,6}}}{100} \cdot \frac{Z^{T_{6,12}}}{100}. \tag{3}$$

From (1) this condition can be expressed in terms of annual IRRs as:

$$\left(1 + \frac{IRR^{T_{0,12}}}{2}\right)^2 = \left(1 + \frac{IRR^{T_{0,6}}}{2}\right) \cdot \left(1 + \frac{IRR^{T_{6,12}}}{2}\right). \tag{4}$$

This condition shows that the DU model requires discounting to be additive, in that the total discounting over one year inferred from the twelve-month choice must equal the total discounting implied by the two six-month choices. If additivity is violated, we denote by subadditivity the case in when discounting is greater when measured using subintervals, i.e., when the left-hand side of (4) is smaller than the right-hand side, and by superadditivity the opposite case in which the left-hand side is greater than the right-hand side. For later use, it is worth noting that a sufficient condition for subadditivity is that $(IRR^{T_{0,12}})^2 < IRR^{T_{0,6}} \cdot IRR^{T_{6,12}}$, and a sufficient condition for superadditivity is the opposite case, $(IRR^{T_{0,12}})^2 > IRR^{T_{0,6}} \cdot IRR^{T_{6,12}}$.

Individuals who violate additivity are not consistent with any discount function in the standard version of the DU model.[20] Indeed, nonadditivity implies a violation of

[20] Additivity is an implication of the DU model regardless of the number of time periods or the shape of the discount function, i.e., how the discount rate changes over time. To see this, note that a discount function assigns each period a discount rate, $\rho_t > 0$, or equivalently, a discount factor $\Delta_t = \frac{1}{1+\rho_t}$. The present value of a reward V available at time T is then worth $V \prod_{t=0}^{T} \Delta_t$. Dividing the time horizon into two subintervals, 0 to T' and T' to T, the reward V is worth $V \prod_{t=0}^{T'} \Delta_t$ at date T', and is further

transitivity over time, under standard workhorse assumptions of the DU model such as time separable and time stationary utility (see Appendix D for a demonstration that under such conditions nonadditivity implies intransitivity).

2.3.2 Constant, declining, increasing, and quasi-hyperbolic discounting:

Depending on the shape of the discount function the DU model makes additional predictions about the annual IRRs for the different time horizons. Using (4) the annual IRRs in the DU model can be written as:

$$IRR^{T_{0,6}} = 2[(1+\rho_1)-1]; \quad IRR^{T_{0,12}} = 2[((1+\rho_1)(1+\rho_2))^{\frac{1}{2}}-1]; \quad IRR^{T_{6,12}} = 2[(1+\rho_2)-1]. \tag{5}$$

Constant discounting: In the case of constant discounting, $\rho_1 = \rho_2 = \rho$. From (5) we can see that this implies constant annual IRRs (and hence IRRs) across the three time horizons: $IRR^{T_{0,6}} = IRR^{T_{0,12}} = IRR^{T_{6,12}}$.

Declining discounting: Declining discounting entails $\rho_1 > \rho_2$. According to (5) this implies $IRR^{T_{0,6}} > IRR^{T_{0,12}} > IRR^{T_{6,12}}$. Intuitively, impatience should be greatest in $T_{0,6}$ with declining discounting, because it includes the present and extends the least far into the future. Behavior in $T_{6,12}$ should be the most patient, because it excludes the present and only concerns payments relatively far into the future. Behavior in $T_{0,12}$ should be in-between.

Increasing discounting: Increasing discounting entails $\rho_1 > \rho_2$. Using (5) this implies the following relationship for annual IRRs elicited over the three time horizons, $IRR^{T_{0,6}} < IRR^{T_{0,12}} < IRR^{T_{6,12}}$.

Quasi-hyperbolic discounting: A special case of declining discounting is the specific functional form commonly used to approximate the hyperbolic discount function, namely the quasi-hyperbolic or β - δ discount function (e.g., Phelps and Pollak, 1968; Laibson, 1997;

---

discounted at time T to be worth $V \prod_{t=0}^{T'} \Delta t V \prod_{t=T'}^{T} \Delta_t$ at time T. This can be re-written as $V \prod_{t=0}^{T} \Delta_t$, which is the same as if the time horizon were not sub-divided. Thus, discounting must be the same whether it is elicited in subintervals or over the interval as a whole.

O'Donoghue and Rabin, 1999). This function involves a discrete drop in the discount rate between the present and the near future, and then constant discount rates between all more-distant future periods. In the literature studies differ in their assumptions about how quickly after the present the drop in discount rates occurs. If the present bias is assumed to extend more than 2 days but less than 6 months into the future, then the quasi-hyperbolic model makes the same predictions as models with hyperbolic, or other forms, of continuously declining discounting, $IRR^{T_{0,6}} > IRR^{T_{0,12}} > IRR^{T_{6,12}}$. If the drop in discount rates is assumed to occur within the two-day window between the date of the experiment and the arrival of the early payments, then only the constant discounting part of the quasi-hyperbolic discount function is relevant for the choices in our experiment, and the model predicts $IRR^{T_{0,6}} = IRR^{T_{0,12}} = IRR^{T_{6,12}}$, the same as with constant discounting. Another, more recent version of the quasi-hyperbolic model involves a fixed, rather than variable, cost of receiving payments in the future (Benhabib et al., 2010). We show in the Appendix that while this model can predict a "magnitude effect,", i.e., higher measured patience as stake sizes increase, for a given level of stakes it generates the same qualitative predictions for IRRs across time horizons as the declining discounting or quasi-hyperbolic models.

3 Aggregate Results on IRR and Time Horizon

In this section we present results on the aggregate intertemporal choice patterns, and then discuss these in light of the behavioral predictions.

3.1 Comparisons of Cumulative Distributions

Figure 1 presents the cumulative distribution functions of the annual IRR for each of the different time horizons. The top panel shows results from the CSS data, and the bottom panel shows results from the SOEP data. For the SOEP data, we pool the $T_{0,12}$ measures across the two sub-samples that have this treatment, as the order and stake sizes are identical.[21]

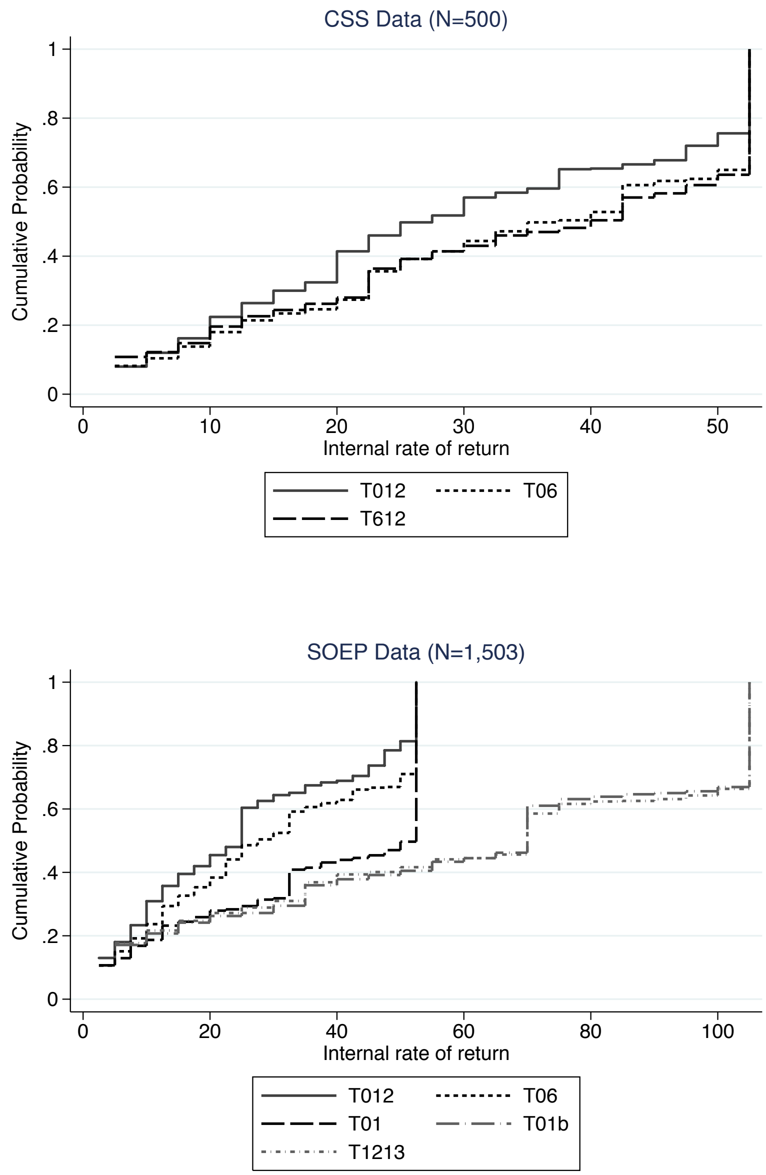
In the top panel of Figure 1, we see that people in the CSS data exhibit a different

[21] Empirically, the cumulatives for the two measures considered separately are very similar, and are not significantly different, using either parametric ($p < 0.849$; Kolmogorov Smirnov) or non-parametric tests ($p < 0.539$; Mann-Whitney). Medians are also identical for the two sub-samples.

distribution of IRRs for $T_{0,12}$ than for $T_{0,6}$, or $T_{6,12}$ ($p < 0.001$; $p < 0.001$; Kolmogorov-Smirnov), in the direction of lower IRRs (and hence greater patience) for the 12-month horizon. People are quite similar in impatience, however, comparing the distributions of IRRs for the $T_{0,6}$ and $T_{6,12}$ measures, and there is no significant difference ($p < 0.90$; Kolmogorov-Smirnov). Thus, people tend to be more patient for long horizons than short horizons, but similarly impatient over the short horizons regardless of the starting date. The quantitative magnitudes of the difference in IRRs is also substantial: The median IRRs for $T_{0,6}$ and $T_{6,12}$ are both at least 10 percentage points higher than the median IRR for $T_{0,12}$, which means a 30 percent higher median annual IRR.[22]

[22] Medians are 37.5 and 40 for $T_{0,6}$ and $T_{6,12}$, compared to 27.5 for $T_{0,12}$.

Figure 1: Cumulative Distributions of IRR



The bottom panel shows the same patterns in the SOEP data. IRRs increase monotonically as time horizon length decreases, with greater impatience for $T_{0,6}$ than $T_{0,12}$ ($p < 0.001$; Kolmogorov-Smirnov), and even greater impatience for $T_{0,1}$ than $T_{0,6}$ ($p < 0.001$; Kolmogorov-Smirnov). The difference going from the longest to shortest horizon is very large: The median IRR for $T_{0,1}$, censored above at 52.5, is more than double the median IRR of 25 observed for $T_{0,12}$.[23] At the same time, IRRs are relatively insensitive to the starting date of the time horizon, in that people are similarly impatient for T01b and T1213 ($p < 0.997$; Kolmogorov-Smirnov). The distribution for $T_{0,1}$ is almost identical to $T_{0,1b}$ and $T_{12,13}$, except for a deviation in the direction of greater patience starting around the middle of the range for $T_{0,1}$. This is potentially due to either an order effect, or a framing effect of the different upper bound for the IRR, or both.[24] Regardless, the cumulative distributions for the one-month horizons are more similar to each other than to the cumulative distributions for longer horizons.

3.2 Interval Regression Analysis

Tables 2 and Table 3 provide another way to look at the results, using interval regressions that correct for right- and left-censoring of the dependent variable. The dependent variable is the measured IRR, and independent variables are dummy variables for time horizon length. Standard errors are robust, clustering on individual.

Column (1) of Table 2 shows that IRRs in the CSS data are significantly lower for the $T_{0,12}$ measure compared to $T_{0,6}$, by more than 6 percentage points, while there is no significant difference between $T_{6,12}$ and $T_{0,6}$.[25] Table 3 presents similar regression analysis based on the SOEP data. Results are reported separately for the three sub-samples, in Panel A, B, and C, respectively. Looking at Column (1) we again see a pattern of lower IRRs for longer horizons, but similar IRRs across horizons of the same length regardless of the starting date: $T_{0,12}$ is significantly lower than $T_{0,6}$ by about 5 percentage points, and

[23] The median IRR for $T_{0,6}$ is 27.5 in the SOEP data, 2.5 percentage points or 10 percent larger than for $T_{0,12}$. The difference in IRRs tends to be larger at some other points in the distribution.

[24] Over the common support, the distributions for $T_{0,1b}$ and $T_{12,13}$ are each significantly different from the distribution for $T_{0,1}$ ($p < 0.001$; $p < 0.001$; Kolmogorov-Smirnov).

[25] The regression analysis also allows checking robustness with respect to order effects, in the CSS data where order was randomized. We add dummy variables for the different possible treatment orders, and interactions of $T_{0,12}$ and $T_{6,12}$ with all of the different orders, to the specification used in Column (1) of Table 2. We again find a significant difference between $T_{0,12}$ and $T_{0,6}$, but not between $T_{0,6}$ and $T_{6,12}$. Furthermore, all interaction terms are not statistically significant, individually or jointly.

lower than $T_{0,1}$ by about 20 percentage points, while $T_{12,13}$ is not significantly different from $T_{0,1b}$ and the point estimate indicates only a 1 percentage point difference.

In columns (2) through (9), Table 2 and Table 3 explore whether similar results are observed for different demographic groups in the population. We consider sub-populations defined by gender, age, education level, and cognitive ability.[26] Looking at the regression estimates, it is apparent that the qualitative results on time horizon effects are similar across all of the sub-groups, in both data sets. Thus, the patterns we find are not isolated to specific populations, but rather are a seemingly general feature of intertemporal choice.[27] We discuss results in the subsequent columns of both tables in later robustness checks.

3.3 Discussion: Aggregate Results and Model Predictions

To summarize, the aggregate results reveal a consistent choice pattern with two key features: (1) People are more impatient for short than long time horizons; (2) people are relatively insensitive to when a given short time horizon starts. In the CSS data this is shown by $IRR^{T_{0,6}} = IRR^{T_{6,12}} > IRR^{T_{0,12}}$, and in the SOEP data, by $IRR^{T_{0,1b}} = IRR^{T_{12,13}} \approx IRR^{T_{0,1}} > IRR^{T_{0,6}} > IRR^{T_{0,12}}$.[28]

These findings are not well explained by either constant, declining, or increasing discounting, maintaining the usual identifying assumptions set out in Section 2.3. The sensitivity of IRRs to time horizon length is inconsistent with constant discounting, while the insensitivity of IRRs to starting date of a given horizon is contrary to the key prediction of declining or increasing discounting (and also to two-system models that predict declining

[26] Cognitive ability is measured by two tests. The scores on the tests are standardized and then averaged to form a single measure of cognitive ability. One test involved matching numbers and unfamiliar symbols for 90 seconds, capturing speed of processing, and the other involved naming as many animals as possible in 90 seconds, providing a measure of crystalized intelligence. Both tests correlate with corresponding sub-modules of widely used tests of cognitive ability. For a more detailed description of these tests see Dohmen et al. (2010).

[27] In the CSS data we see a gender difference (Columns (2) vs. (3) in Table 2), in that the difference in IRRs for $T_{0,12}$ versus $T_{0,6}$ is about two times larger for men than women. In the SOEP data, however, we see little evidence of a gender difference (Columns (2) vs. (3) in Table 3).

[28] It is noteworthy that IRRs tend to be lower in the SOEP data than the CSS data, for comparable time horizons (e.g., compare the constant term in Column (1) of Table 2 to the constant term in Column (1), Panel A, of Table 3). One interpretation is a "magnitude effect", such that the level of impatience decreases with stake size for a given pair of horizons. Such an effect would be an anomaly for standard discounting models, where the IRR is assumed to be independent of stake size. Many other studies also report finding lower IRRs as stake sizes increase (see Frederick et al., 2002). One explanation is provided by the model of fixed-cost declining discounting, proposed by Benhabib et al., 2010. An alternative explanation for this stylized fact could be that the unobserved utility function is more linear for higher stakes (see Andersen et al., 2011, for a discussion).

Table 2: IRR as a Function of Time Horizon, by Demographic Groups and Characteristics: CSS Data

	All	Males	Females	Age ≤med.	Age >med.	Less educated	More educated	IQ ≤med.	IQ >med.	More willing to take risks	Less willing to take risks	Did not Consid. market int.	Considered market int.	Plans to spend	Plans to save	Credit constrained	Not credit constrained
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
$T_{0,12}$	-6.42***	-8.20***	-4.82***	-6.56***	-6.34***	-6.63***	-5.83***	-6.09***	-6.66***	-6.07***	-6.66***	-8.55***	-3.93***	-6.61***	-6.15***	-5.03**	-6.64***
	(0.85)	(1.20)	(1.19)	(1.05)	(1.39)	(1.04)	(1.42)	(1.53)	(0.96)	(1.12)	(1.19)	(1.20)	(1.22)	(1.03)	(1.50)	(2.17)	(0.92)
$T_{6,12}$	0.14	0.25	0.07	-1.26	1.75	0.71	-1.30	2.22	-1.38	1.44	-0.66	-1.80	2.82*	-0.02	0.42	3.80	-0.77
	(1.11)	(1.50)	(1.63)	(1.47)	(1.70)	(1.32)	(2.08)	(1.68)	(1.48)	(1.40)	(1.59)	(1.63)	(1.46)	(1.36)	(1.94)	(2.91)	(1.21)
Observations	1500	693	807	768	732	1113	387	657	843	564	936	936	558	1002	498	330	1131

Notes: Interval regression estimates. Dependent variable is the IRR for a given time horizon. The reference time horizon is $T_{0,6}$. Demographic groups are defined by: More educated indicates that an individual completed the Abitur, a college entrance exam in Germany; IQ is measured by the average of an individual's score on two different IQ tests; self-reported willingness to take risks on an 11-point scale from completely unwilling to completely willing; thinking about market interest rates during the choice experiments is self-reported; self-reported plans to save most or all of the immediate payment in the experiment; credit constraints are measured by a question asking about ability to borrow money in the event of an unexpected expense. Robust standard errors are in parentheses, adjusted for clustering on individual. *, ** indicates significance at 10 and 5 percent level..

Table 3: IRR as a Function of Time Horizon, by Demographic Groups and Characteristics: SOEP Data

	All	Males	Females	Age ≤med.	Age >med.	Less educated	More educated	IQ ≤med.	IQ >med.	More willing to take risks	Less willing to take risks	Did not Consid. market int.	Considered market int.	Plans to spend	Plans to save	Credit constrained	Not credit constrained
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
Panel A																	
Sample 1:																	
$T_{0,12}$	-5.26***	-4.86***	-5.66***	-5.90***	-4.58***	-5.33***	-4.97***	-3.08***	-6.55***	-5.03***	-5.40***	-6.61***	-3.58***	-6.69***	-3.69***	-5.07***	-5.33***
	(0.68)	(0.89)	(1.03)	(0.94)	(0.98)	(0.77)	(1.44)	(0.93)	(0.93)	(0.91)	(0.97)	(0.95)	(0.97)	(1.01)	(0.89)	(1.87)	(0.72)
Constant	31.17***	28.31***	34.15***	30.48***	31.85***	34.11***	21.46***	31.87***	30.74***	34.83***	28.31***	37.99***	21.45***	34.44***	27.50***	43.85***	27.63***
	(1.34)	(1.86)	(1.91)	(1.87)	(1.91)	(1.62)	(2.17)	(2.13)	(1.72)	(1.97)	(1.79)	(1.94)	(1.57)	(1.82)	(1.95)	(3.22)	(1.41)
Observations	980	504	476	486	494	730	218	358	622	418	562	622	358	506	474	230	744
Panel B																	
Sample 2:																	
$T_{0,12}$	-19.67***	-21.05***	-18.46***	-19.10***	-20.28***	-20.15***	-16.64***	-18.63***	-20.26***	-19.58***	-19.80***	-22.93***	-15.45***	-22.73***	-16.76***	-27.41***	-17.87***
	(1.35)	(2.04)	(1.80)	(1.97)	(1.86)	(1.62)	(2.39)	(2.12)	(1.75)	(2.10)	(1.77)	(1.89)	(1.86)	(1.74)	(2.03)	(3.70)	(1.44)
Constant	44.96***	46.61***	43.51***	41.08***	48.84***	47.50***	32.35***	46.24***	44.17***	49.39***	41.50***	53.12***	32.28***	50.16***	40.03***	60.49***	41.10***
	(1.80)	(2.64)	(2.47)	(2.42)	(2.70)	(2.11)	(3.51)	(2.83)	(2.34)	(2.88)	(2.29)	(2.52)	(2.33)	(2.56)	(2.51)	(4.49)	(1.94)
Observations	974	444	530	468	506	766	180	352	622	440	534	642	328	470	504	212	762
Panel C																	
Sample 3:																	
$T_{12,13}$	-1.45	-0.22	-2.42	-0.10	-3.14	1.78	-7.48	2.32	-3.15	1.77	-4.50	-7.05	5.04	-4.15	0.11	-3.60	-0.93
	(3.36)	(5.03)	(4.52)	(4.33)	(5.27)	(4.15)	(6.03)	(6.28)	(3.95)	(4.81)	(4.70)	(4.95)	(4.22)	(6.32)	(3.80)	(7.12)	(3.70)
Constant	97.25***	96.53***	97.83***	91.92***	103.49***	102.48***	82.34***	118.21***	85.41***	102.17***	92.05***	122.69***	57.85***	123.53***	78.92***	139.69***	89.92***
	(5.16)	(7.72)	(6.94)	(6.79)	(7.89)	(6.36)	(8.85)	(9.91)	(5.86)	(7.46)	(7.11)	(7.77)	(5.52)	(9.45)	(5.83)	(18.21)	(5.25)
Observations	1052	464	588	552	500	784	240	424	628	544	508	722	330	496	556	188	860

Notes: Interval regression estimates, separately by sub-sample of the SOEP data. Dependent variable is the IRR for a given time horizon, with two time horizons (observations) per individual. Reference category is $T_{0,6}$ in Panel A, $T_{0,1}$ in Panel B, and $T_{0,1b}$ in Panel C. Demographic groups are defined by: More educated indicates that an individual completed the Abitur, a college entrance exam in Germany; IQ is measured by the average of an individual's score on two different IQ tests; self-reported willingness to take risks on an 11-point scale from completely unwilling to completely willing; thinking about market interest rates during the choice experiments is self-reported; self-reported plans to save most or all of the immediate payment in the experiment; credit constraints are measured by a question asking about ability to borrow money in the event of an unexpected expense. Robust standard errors are in parentheses, adjusted for clustering on individual. *, ** indicates significance at 10 and 5 percent level..

discounting; see, e.g., Fudenberg and Levine, 2012). The data are also not consistent with alternative versions of the declining discounting model, quasi-hyperbolic or fixed cost, regardless of assumptions about the length of the present.

The aggregate results are consistent with subadditivity. The patterns we observe in the CSS data imply that, on average, $(IRR^{T_{0,12}})^2 < IRR^{T_{0,6}} \cdot IRR^{T_{6,12}}$. Thus, impatience over 12 months is greater when it is elicited in two sub-intervals, than when it is elicited directly for the entire 12 month horizon. While we do not have three time horizons per individual in the SOEP data, and thus cannot test for additivity directly, we see the same pattern, that internal rates of return are higher for short horizons and lower for longer horizons, which is consistent with subadditivity. A simple (descriptive) model proposed by Read (2001) matches these qualitative features of the data. Specifically, the model specifies an indifference relation for the early and delayed payment for a given time horizon as:

$$(1+\rho)^{t^{\sigma}} 100 = Z^t \tag{6}$$

Where $0 < \sigma < 1$. The parameter σ can be thought of as a reduced form way to capture a mechanism of biased perception of time, such that subjective perception of time duration is a concave function of objective time duration. Applied to the three time horizons in the CSS data, e.g., we have:

$$(1+\rho) = \frac{Z^{T_{0,6}}}{100}; \qquad (1+\rho)^{2^{\sigma}} = \frac{Z'^{T_{0,12}}}{100}; \qquad (1+\rho) = \frac{Z^{T_{6,12}}}{100}. \tag{7}$$

The model implies:

$$\frac{Z'^{T_{0,12}}}{100} < \frac{Z^{T_{0,6}}}{100} \cdot \frac{Z^{T_{6,12}}}{100}. \tag{8}$$

In other words, discounting will be subadditive.

3.4 Investigating Sensitivity to Other Maintained Assumptions

The theoretical predictions were based on maintained Assumptions (1) to (4), which have been questioned in the recent literature in economics. In this section we discuss whether it seems plausible conceptually, and empirically, that failures of one or more of these assumptions could be driving the pattern of aggregate results. Columns (10) to (17) of Table 2 and Table 3 present some results that are useful as robustness checks.

Assumption (1) is that the payments in the experiment were not differentially credible across different time horizons. If payments are not perfectly credible, and credibility varies with both timing of the payment and payment amount, then particular patterns of credibility concerns could conceivably generate the sub-additive choice pattern we observe.[29] As discussed above, however, there are several features of our setting that help ensure that credibility of all payments is high regardless of timing, for example subjects being in a panel survey and thus a long-term relationship with the experimenter. Furthermore, we varied the payment procedure across data sets, in a way that might be expected to generate different results if subjects have credibility concerns. Instead, both data sets exhibit the same qualitative choice patterns. For these reasons, we conclude that credibility concerns are unlikely to be a key mechanism driving the results.

If utility is nonlinear for the range of stakes in the experiment, this violates Assumption (2), and could generate choices that are qualitatively in line with the sub-additivity findings. The needed curvature, however, is convex (we provide a proof in Appendix E), which is unlikely to be true in general in the population. Indeed, Dohmen et al. (2011) implement an incentivized lottery choice experiment in a representative sample of the the German population, and find that more than 85 percent of individuals are risk neutral or risk averse. In our data we can also check directly whether a proxy for curvature of utility is important for exhibiting subadditivity. Specifically, we use a survey measure of risk attitudes that has been shown in previous research to correlate with measures of risk aversion based on real stakes lottery experiments (Dohmen et al., 2011).[30] Using the

[29] One can construct examples in which this is the case, involving credibility that declines with distance of the promised payment date from the present, and letting perceived credibility vary with the payment amount. Somewhat counter-intuitively, however, these examples involve credibility increasing with the payment amount, which does not seem particularly plausible.

[30] The question asks: "How willing are you, in general, to take risks." Respondents can use an 11-point scale from completely unwilling (0) to completely willing (10). For a representative sample of German adults who responded to the the question, and also participated in real stakes lottery experiments,

survey measure, Columns (10) and (11) of Tables 2 and 3 compare the results across the sub-groups with less than or greater than the median willingness to take risks. A very similar sub-additivity pattern is observed for both groups. This suggests that curvature of utility is unlikely to drive the sub-additivity result.[31]

While the literature has typically assumed that subjects treat monetary rewards in discounting experiments as equivalent to consumption opportunities, it could be that Assumption (3) is not in fact valid. If so, this introduces potentially complicated motives related to borrowing or saving, which could jeopardize ability to infer time discounting from choices in experiments using monetary rewards (for a discussion see Cohen et al., 2020). For example, in terms of borrowing, it has been hypothesized that subjects might engage in arbitrage between the interest rates offered in the experiment and market borrowing rates. Subjects might accept any IRR in the experiment higher than the market borrowing rate, e.g., 5%, giving the appearance of being patient, when in fact they have a high discount rate of 20%, and then plan to use market borrowing to finance immediate consumption. In our data, however, we do have a way to observe whether individuals might be thinking about arbitrage: After subjects had completed the experiment we asked whether they had thought at all during the experiment about current market interest rates or returns, a seeming pre-requisite for engaging in arbitrage between experiment and market interest rates.[32] The majority of subjects in both data sets stated that they had not not done so (67% in the CSS data, and 66% in the SOEP data). Column (12) of Table 2 and Table 3

answers to the question were robustly correlated with the degree of risk aversion measured in the experiments (Dohmen et al., 2011).

[31] A caveat is that the survey measure does not involve specified stakes and probabilities, and thus could potentially capture other determinants of willingness to take risks besides curvature of utility, e.g., risk perceptions. In the SOEP data, however, we have an additional survey question asking about willingness to invest in a hypothetical lottery that does have given stakes and probabilities. Specifically, the 1,214 individuals who remain in the SOEP panel until the 2009 wave and answer a question about how much of 100,000 Euro they would invest in a lottery that doubles the investment, or returns only half, with equal probability. Responses are elicited in increments of 20,000 Euro, starting from zero, up to the full amount. About 75 percent of individuals indicate strong risk aversion, choosing to invest zero rather than invest 20,000 or more, suggesting that convex utility is unlikely to explain the subadditivity results. Furthermore, we estimate interval regressions as in Table 3, separately for individuals who invest nothing in the lottery, compared to those who invest at least some money. We find that IRRs are significantly lower for longer time horizons than for subintervals, regardless of curvature of utility. Specifically, in our subsample comparing T012 to T06, IRRs are lower for T012 by -5.83 and -3.85 percentage points, for more and less risk averse, respectively ($p < 0.001$; $p < 0.001$). In our second subsample comparing T01 to T012, IRRS are lower for T012 than T01 by -18.86 and -19.49, for more and less risk averse samples, respectively ($p < 0.001$; $p < 0.001$). For the subsample comparing time horizons of equal length, T1213 to T01, we find no significant differences in IRRs across horizons, regardless of risk aversion.

[32] The exact wording is provided in Appendix G.

shows that the sub-additivity pattern is present for individuals who did not think about market interest rates. This helps establish that the sub-additivity result is not driven by subjects engaged in arbitrage strategies, or other motives related to borrowing.

Another way that Assumption (3) might be violated is if individuals plan to save experiment payments rather than spend them upon receipt; in this case also, choices in the experiment do not reflect choices between different timings of consumption. We therefore measured plans to save. Specifically, we asked individuals after the experiment what they planned to do with the immediate payment should they receive it, with four possible responses: "Spend all of it;" "spend most of it and save a little;" "spend a little and save most of it;" "save all of it."[33] We see that in both data sets a substantial proportion of individuals plan to spend most or all of the money immediately (67% plan to spend all or most in the CSS data, and in the SOEP data, where the immediate payment is twice as large, the corresponding proportion is about 50%). Column (14) checks whether the subadditivity pattern is present among individuals who do plan to spend the money right away. We find that this is the case in both data sets. Thus, the subadditivity result holds focusing on individuals who satisfy Assumption 3.

Assumption (4) is that preferences are stationary over time. Halevy (2015) argues that one source of non-stationarity in preferences could be the possibility of randomly arising expenses, combined with credit constraints. More generally, as noted recently by Strack and Taubinsky (2021), non-stationarity could arise if for some reason people have different preferences at different calendar dates, i.e., "taste shocks." Columns (16) and (17) of both tables show that the sub-additivity pattern is similar, however, regardless of whether or not individuals report being credit constrained, indicating that this source of non-stationarity does not appear to drive the results. Furthermore, we see similar patterns regardless of calendar dates, suggesting the results are unlikely to be driven by time-varying taste shocks. Discounting is similar across short time horizons regardless of whether they are 6 months apart (today vs. 6 and 6 vs. 12) or one year apart (today vs. 1 month vs. 12 months vs. 13 months). Also, we see similar results for the SOEP and CSS data, even though these were collected at different times during the calendar year, late winter and early summer, respectively. This provides additional evidence that the results are not driven by patterns in how preferences change with calendar date. In summary,

[33] The exact wording is provided in Appendix G.

while some of the identifying assumptions (1) to (4) may be violated for some individuals, this does not appear to drive the pattern of subadditivity at the aggregate level found in both datasets.

A remaining concern could be that the subadditivity result is driven by some type of bias or decision-making heuristic specific to the structure on intertemporal choice experiments involving money. We consider two possibilities, but conclude that neither can explain the aggregate findings.

The first possibility is that exponential growth bias might explain the results. This bias is a tendency for individuals to fail to appreciate the influence of compounding, and thus overestimate the interest rate implied by a delayed payment of a given size (see, e.g., Stango and Zinman, 2009). If individuals ignore the effect of compounding in our experiments, this tends to make the implicit interest rates offered for longer time horizons seem relatively more attractive, and works in favor of finding more patient choices in longer horizons, the same as subadditivity. Such a bias, however, has only a modest implication for behavior in our setting, due to the time horizons and parameters we use, and is thus is unable to explain the substantial magnitude of subadditivity that we observe. For example, the median choice behavior in the $T_{0,12}$ horizon in the CSS data implies an annual IRR of 27.5 with semi-annual compounding; ignoring compounding would instead imply an IRR about about 30. While this brings the discounting implied by choices in the long horizon closer to what we observe in the shorter horizons, it is still far away; median IRRs are 37.5 and 40 for $T_{0,6}$ and $T_{6,12}$, respectively.[34]

We also investigated a second possible explanation for the findings, based on a rule of thumb heuristic: A "double-double" heuristic such that an individual requires double the amount of money to wait twice as long. We focus on the CSS data where it is possible to see the full subadditivity pattern. We find that relatively few individuals exhibit the heuristic, however, with only 15 percent switching for exactly double the amount in $T_{0,12}$ versus $T_{0,6}$, and find even fewer who switch for exactly double in both $T_{6,12}$ and $T_{0,6}$ compared to T_{012}. If we exclude such individuals, we find similarly strong evidence of subadditivity, in that

[34] The impact of exponential growth bias increases with the amount of the delayed payment, all else equal. Another way to see that exponential growth bias would only have a modest effect in our setting is to note that on average across the 20 delayed payments offered in the $T_{0,12}$ horizon of the CSS data, exponential growth bias increases the perceived annual IRR by only about 2.2 percentage points, which is less than is needed to move the IRR one category higher in our elicitation based on intervals of 2.5 percentage points.

median annual IRR's are 10 percentage points lower for T_{012} compared to $T_{0,6}$ or $T_{6,12}$, the same difference as we observe when we use the full sample.[35]

4 Individual-level analysis

4.1 Prevalence of nonadditivity

We focus our individual-level analysis on the CSS data because it contains information about choices for all three time horizons for each individual. We categorize individuals as falling into "additive," "subadditive", and "superadditive" categories. The categorization takes into account that the experiments measure IRRs in intervals of 2.5 percentage points, and is deliberately conservative in terms of favoring additivity: Someone is classified as sub-additive if the squared upper-bound of the IRR for the long time horizon is lower than the product of the lower bounds of the short horizon IRRs, and vice versa for super-additivity. The residual group is classified as additive. The analysis excludes individuals for whom right-censoring makes it ambiguous whether additivity is satisfied or not.[36] In our sample, 26 percent of individuals fall into this ambiguous category. This censoring is a caveat to the exact percentages of different types the we calculate, because in the absence of censoring, one category or another might gain more members. What is clear, however, is that this censoring does not change the conclusion that nonadditivity is a pervasive characteristic of intertemporal choice. Indeed, even if we assume that every single censored individual satisfies additivity, this would still leave 52 percent of the population being nonadditive.

The first row of Table 4 shows that 30 percent of individuals satisfy additivity according to our conservative measure. Thus, the group of individuals whose choices can be rationalized by some version of the DU model is a minority. By contrast, about 59 percent of individuals exhibit sub-additivity, and 11 exhibit superadditivity. Thus, nonadditivity, and particularly subadditivity, is by far the most prevalent choice pattern at the individual level.

The next row of Table 4 further subdivides the sample of additive individuals to indicate the percentages of the overall sample that can be classified as consistent with

[35] We thank an anonymous referee for suggesting that we check for this heuristic.

[36] The construction of the sample takes into account the fact that right-censoring of short time horizons does not prevent classifying someone as sub-additive, and right censoring of the long horizon does not prevent categorizing someone as super-additive.

particular shapes of discount functions. Our categorization distinguishes between strong and weak consistency with a given discount function, where both inequalities for IRRs must be strictly satisfied for the former, but only one must be strictly satisfied for the latter. For example, an individual can be additive and weakly consistent with declining discounting if $IRR^{T06} > IRR^{T612} = IRR^{T012}$ or $IRR^{T06} = IRR^{T012} > IRR^{T612}$.[37] We see that 16 percent of the sample are strongly or weakly classifiable as constant discounters, 5 percent as declining discounters, and 8 percent as increasing discounters. Thus, we find that when additivity is verified to hold, rather than assumed, the prevalences of individuals who can be rationalized by any given type of discount function in the DU model is relatively small.

The third row of Table 4 subdivides the group of nonadditive individuals according to whether their choices are in some sense consistent with present bias, future bias, or no bias, albeit not in ways that can be rationalized with a discount function in the DU model. For example, those classified as nonadditive but present biased have $IRR^{T06} > IRR^{T612} > IRR^{T012}$, indicating that choices are more impatient if they involve the present (near future), but also violate additivity. Those who violate additivity and have the opposite strict ranking of IRRs are classified as nonadditive and future biased. We also identify individuals who are nonadditive but weakly consistent with present bias and future bias, meaning that one of the IRR relationships can hold with equality. Another group of individuals violate additivity and have an inconsistent classification of bias type, depending on whether one compares short horizons to each other (SH comparisons), or compares a short to long horizon (OH or OSH comparison). Specifically, such individuals have IRRs for both short horizons either above or below the IRR for the long horizon. We group these individuals according to whether the comparison of the two short time horizons to each other (SH comparison) is by itself consistent with present bias ($IRR^{T06} > IRR^{T612}$), future bias ($IRR^{T06} < IRR^{T612}$), or zero bias ($IRR^{T06} = IRR^{T612}$). The rationale for such a classification is that, while the comparison of short to long time horizon (OH comparisons) confounds present or future bias with subadditivity or superadditivity, the comparison between two short horizons does not. We denote these categories as nonadditive with SH present bias, SH future bias, and SH no bias, respectively. Finally,

[37] Such individuals can be additive because of our conservative definition of nonadditivity. A rationale for treating such a case as weakly consistent with declining discounting is that the elicitation of IRRs in intervals might yield equality when in fact there is a small difference in the direction predicted by the discount function.

Table 4: Individual Types, CSS Data

Additivity types

	Additive	Subadditive	Superadditive	Total
Percent	29%	59%	12%	100%

Subdividing additive types by discounting type

	Constant	Declining	Increasing	Total
Percent	16%	5%	8%	29%

Subdividing nonadditive types by bias type

	Strong/ weak present	Strong/ weak future	No bias SH	Present SH	Future SH	Ambiguous	Total
Percent	6%	8%	17%	11%	17%	11%	71%

Notes: The first row shows the percentages of individuals in the CSS data who exhibit additivity, subadditivity, and super-additivity, excluding individuals for whom right-censoring prevents unambiguous classification of additivity. Percentages refer to the uncensored sample; the sample size is $N = 369$. The second row shows the percentages of the sample who are additive and consistent with a given shape discount function in the DU model. The third row shows the percentages of the sample who are nonadditive and exhibit a given type of bias: present bias, future, bias, or no bias. Some individuals who are additive have an ambiguous bias type due to censoring.

there is a residual category of individuals who violate additivity, but due to right-censoring, cannot be unambiguously classified as present, future, or no bias types even according to SH comparisons.

We see in the third row of Table 4 that about 17 percent of individuals are nonadditive and show signs of present bias in some form. About 25 percent are nonadditive and show signs of future bias, while 17 percent show no bias. A residual 17 percent violate additivity, but due to right censoring, cannot be classified in terms of bias. In summary, non-trivial fractions of the sample are both nonadditive and show choice patterns consistent with present bias or future bias.

We also checked robustness of additivity violations to an even more conservative approach that further expands the upper and lower bounds for each time horizon's IRR, effectively allowing for some larger "errors" in decision making, and favoring additivity. Specifically, we added and subtracted 2.5 percentage points from the regular upper and lower bounds, respectively. We then classified someone as sub-additive only if the square of the wider upper bound IRR for the long horizon is less than the product of the wider lower bound IRR's for the short horizons, and vice versa for super-additive. Again, we exclude individuals for whom right-censoring makes classification ambiguous.[38] While this approach mechanically favors additivity, we see that the majority still violates additivity, and sub-additivity remains a key feature of the distribution of types. Specifically, about 52 percent of individuals violate additivity, and about 42 percent of individuals are sub-additive.[39]

4.2 Individual level correlates of nonadditivity

As a second step in our individual level analysis, we explore the relationship of additivity violations to demographics and other individual-level characteristics. Table 5 presents the results of Probit estimations, where the dependent variable is equal to 1 if an individual violates additivity and 0 otherwise. The sample excludes individuals for whom right-censoring leads to an ambiguity in assessing additivity violations. The independent variables include the same demographics used for the sub-group analysis in the section on aggregate results, with the addition of a quadratic in age. In the final column we also control for the traits used in the robustness checks on the aggregate results.

The results in Table 5 show that additivity violations are not significantly more or less likely based on variation in the different demographics, and are not significantly related to education or cognitive ability. The one significant relationship is with the indicator for thinking about market interest rates: Those who thought about market interest rates are significantly less likely to violate additivity. This goes in the opposite direction of the hypothesis that arbitrage strategies lead to apparent additivity violations. Instead, one explanation could be that this pattern captures greater familiarity with and exposure

[38] Note that the combination of allowing for errors, but addressing the impact of censoring on categorization, leads to a smaller sample. For example, consider individuals who were unambiguously subadditive without errors, but had right censoring for one of the short time horizons. Allowing for errors, some of these are no longer unambiguously subadditive, but on the other hand, right-censoring prevents classifying them as unambiguously additive. The resulting sample involves $N = 349$.

[39] Notably, this robustness check also indicates that exponential growth bias cannot explain the prevalence of subadditivity at the individual level. The bias can change perceived IRR's in $T_{0,12}$ by at most 5 percentage points (for the three highest delayed payments that we offer), but we still find prevalent subadditivity when we allow for errors of this magnitude. The "double-double" heuristic also cannot explain the finding that most individuals are subadditive; if we focus on individuals for whom additivity type is unambiguous, and exclude those who exhibit the heuristic in $T_{0,12}$ and $T_{0.6}$, 68 percent of individuals are subadditive. Excluding those who exhibit the heuristic for both short horizons compared to the long horizon, 65 percent of individuals are subadditive.

to financial decision making, which helps people reduce additivity violations.[40] Notably, however, subadditivity is still the aggregate pattern even among those who thought about market interest rates, as seen in Column (13) of Tables 2 and 3. Thus, the bias is reduced but not eliminated by financial sophistication. These results underline that nonadditivity is a pervasive feature of intertemporal choice, and also provide some suggestive evidence that it is not entirely driven by a mechanism of cognitive mistakes.

4.3 Relationship of discounting types and nonadditivity to economic outcomes

The final part of our analysis is exploratory. We investigate how nonadditivity, and measures of bias like present bias, are related to various economic outcomes. First, we consider outcomes that have been related frequently to experimental indicators for time preference or discounting types in the literature, due to hypotheses that these might depend on impatience or present bias: BMI, smoking, drinking, poor nutrition, poor health, life satisfaction, and having an overdrawn account. We also include a second set of outcomes, household income and household wealth, as overall proxies for financial success and quality of financial decision making. It is not clear ex ante how additivity should be related to the first set of outcomes, if at all, as the mechanisms underlying nonadditivity are conceptually different from those thought to underlie, e.g., present bias.[41] Given that nonadditivity involves distortions in intertemporal choice depending on how options are framed, however, it seems plausible that nonadditivity might be associated with less financial success and lower quality of financial decision making.

Table 6 shows OLS regressions of outcomes on an indicator for nonadditive with additive as the omitted category. We include controls for demographics, and the indicator for thinking about market interest rates, since we have seen that additivity violations are correlated with this proxy for financial sophistication. The results in Table 6 show

[40] A multinomial logit that treats subadditivity and superadditivity as separate categories shows that individuals who thought about market interest rates were significantly less likely to be subadditive than additive.

[41] The implications of nonaddivity for outcomes depends on how the consumption opportunities associated with different choices are framed. Nonadditive types could make different choices about a given outcome, e.g., investment in preventing health behaviors, depending on how choices are framed in terms of subintervals or intervals as a whole. We do not observe the nature of the choice architectures facing individuals in our sample, and these may even vary across individuals for a given outcome, making predictions of how nonadditivity should be related for specific outcomes to a certain extent ambiguous.

Table 5: Relationship of non-additive discounting to individual characteristics

	Non-additive			
	(1)	(2)	(3)	(4)
Female	0.01	0.01	0.00	-0.01
	(0.05)	(0.05)	(0.05)	(0.05)
Age	-0.01	-0.01	-0.01	-0.00
	(0.01)	(0.01)	(0.01)	(0.01)
Age^2	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Educated		-0.08	-0.07	-0.04
		(0.05)	(0.06)	(0.06)
IQ			-0.00	0.02
			(0.03)	(0.03)
Thought int. rate				-0.18***
				(0.05)
Credit constrained				0.04
				(0.07)
Willingness to take risks				-0.00
				(0.01)
Plans to save experiment payment				-0.02
				(0.05)
Observations	361	361	319	313

Notes: Marginal effects from Probit estimates. Sample excludes individuals for whom right-censoring prevents assessing additivity violations. Characteristics include an indicator for being highly educated (having taken the *Abitur*, a college entrance exam in Germany); the average of an individual's score on two different IQ tests; an indicator for financial sophistication, proxied by reporting having thought about market interest rates during the choice experiments; indicator for self-reported credit constraints; self-reported willingness to take risks in general on an 11-point scale from completely unwilling to completely willing; self-reported plans to save most or all of the immediate payment in the experiment. Robust standard errors are in parentheses. *, ** indicates significance at 10 and 5 percent level..

Table 6: Relationships of economic outcomes to nonadditivity

	BMI	Smoker	Drinking	Poor nutrition	Health	Happiness	Overdrawn	Ln income	Ln wealth
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Nonadditive	-0.28	0.05	0.05	0.20**	-0.37	-0.27	0.04	-0.13*	-1.22*
	(0.55)	(0.06)	(0.12)	(0.08)	(0.26)	(0.20)	(0.06)	(0.08)	(0.69)
Female	-1.55***	-0.13**	-0.55***	-0.38***	-0.01	0.09	-0.01	-0.09	-0.83
	(0.49)	(0.05)	(0.11)	(0.08)	(0.24)	(0.19)	(0.05)	(0.07)	(0.62)
Age	0.28***	0.01	-0.01	-0.01	-0.05	-0.03	0.02***	0.01	0.32***
	(0.06)	(0.01)	(0.02)	(0.01)	(0.03)	(0.02)	(0.01)	(0.01)	(0.11)
Age^2	-0.00***	-0.00*	0.00	0.00	0.00	0.00	-0.00***	-0.00	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Educated	-0.22	-0.13**	0.24*	-0.34***	0.54**	0.33*	-0.06	0.07	1.68***
	(0.57)	(0.06)	(0.14)	(0.09)	(0.24)	(0.18)	(0.06)	(0.09)	(0.64)
IQ	-0.11	0.04	0.03	0.01	0.34*	0.09	0.05	-0.02	0.13
	(0.34)	(0.03)	(0.07)	(0.06)	(0.18)	(0.16)	(0.04)	(0.06)	(0.42)
Thought int. rate	-0.30	0.02	0.05	-0.18**	0.10	0.03	-0.05	0.07	0.29
	(0.54)	(0.05)	(0.12)	(0.08)	(0.24)	(0.20)	(0.05)	(0.07)	(0.66)
Constant	18.85***	0.35*	1.48***	2.98***	9.28***	8.62***	-0.06	7.43***	-0.27
	(1.24)	(0.18)	(0.39)	(0.29)	(0.73)	(0.54)	(0.15)	(0.26)	(2.60)
Observations	319	326	326	325	326	326	311	310	252
Adjusted R^2	0.124	0.061	0.079	0.135	0.132	0.013	0.074	0.015	0.123

Notes: Notes: OLS estimates. BMI denotes body mass index; Smoker is an indicator for smoking more than occasionally; Drinking is an index for frequency of drinking (combining information about wine, spirits, beer, and mixed drinks), with higher values indicating more frequent drinking; Poor Nutrition is self-reported degree of poor nutrition, on a 4-point scale, with higher values indicating worse nutrition; Health is self-reported satisfaction with health; Happiness measures subjective well-being; Overdrawn is days per year that current deposit account is overdrawn; Ln income is the natural log of household income; Ln wealth is the natural log of household wealth, except that individuals with zero or negative wealth are coded as having a zero value for the dependent variable. Nonadditive is an indicator for violating additivity. Control variables include an indicator for being highly educated (having taken the *Abitur*, a college entrance exam in Germany); the average of an individual's score on two different IQ tests; an indicator for financial sophistication, proxied by reporting having thought about market interest rates during the choice experiments. The sample excludes individuals for whom censoring makes additivity unambiguous. Robust standard errors are in parentheses. *, ** indicates significance at 10 and 5 percent level.

that nonadditivity is not significantly related to any of the first set of outcomes, except for being associated with poor nutrition. The lack of a relationship is not surprising, for reasons discussed above. We do see, however, that nonadditivity is associated with lower income and wealth, with the relationships being marginally significant.[42] Nonadditivity means that individuals' choices are influenced by how options are structured or framed. The overall effect may be to lead to lower quality financial decision making and worse financial outcomes as captured by income and wealth.[43]

One interesting additional question is whether the relationship of nonadditivity to these financial outcomes might be capturing the fact that some of these individuals exhibit signs of future or present bias. We therefore add indicators for present bias and future bias to the regressions, with results shown in the appendix in Table A.1. In this case the coefficient on nonadditivity can be interpreted as the relationship of nonadditivity to outcomes among those who have no bias or ambiguous bias type. We see that results for the relationship of nonadditivity to financial outcomes are similar in terms of point estimates, and still marginally significant for wealth although not for income. Thus, the negative relationship of nonadditivity to financial outcomes appears to not just reflect a correlation with other forms of bias like present bias.[44] Notably, the indicators for present and future bias are significantly related to only a few of the economic outcomes: Both are associated with a higher probability of smoking, future bias is associated with a lower probability of drinking, and present bias is associated with worse health. These null results should be interpreted with caution, however, given that such bias types make up only a modest proportion of the sample.

---

[42] Log wealth is coded as a zero in the data for individuals with zero or negative wealth. Being nonadditive is associated with about a 10 percentage point higher probability of falling into this zero category ($p = 0.12$; result is the marginal effect from a probit regression with the same controls as Column (9)).

[43] We also investigated regressing financial outcomes on separate indicators for subadditivity and superadditivity (results are available upon request). We find that both subadditivity and superadditivity are associated with lower wealth, both with similar size point estimates to the nonadditivity indicator in Table 6, although coefficients are not individually significant. These indicators also have similar size negative coefficients to each other in a regression explaining income, although the coefficient is not significant for superadditivity. Thus it appears to be deviations from additivity, rather than a particular type of nonadditivity, that are associated with worse financial outcomes.

[44] In regressions with separate indicators for subadditivity and superadditivity, adding indicators for bias type has little impact on the results (results available upon request).

5 Methodological implications

5.1 Method variance

One implication of our findings is that different traditional approaches to categorizing groups or individuals into different discounting types within the DU model, or estimating time preference parameters for the model, can be expected to lead to systematically different results due to prevalent subadditivity. For example, with subadditivity, OH designs will tend to find declining discounters, OSH to find increasing discounters, and SH to find something "in-between." The findings of CTB designs, which have typically combined both OH, SH, and sometimes OSH approaches, will depend on the particular types and numbers of time horizons used.

We can use our data to illustrate how OH, SH, and OSH designs will tend to have systematically different results on categorizing discounting types. Table 7 shows the following patterns for both the CSS and SOEP data: The aggregate finding with OH comparisons is declining discounting, and the modal categorization at the individual level is declining discounting (roughly 60 percent); the aggregate result with SH comparisons is constant discounting, with a roughly uniform distribution at the individual level across declining, constant, and increasing types; with the OSH approach, which is only possible in the CSS data, the aggregate finding is instead increasing discounting, and the modal pattern at the individual level is increasing discounting (roughly 65 percent). Note that in the CSS data, the same group of individuals is being categorized as declining, constant, or increasing at the aggregate level, depending on the measure used. These patterns can be understood as arising from the way that the different designs are influenced by subadditivity.

We also find evidence of this same method variance in our survey of the previous literature. As shown in Figure 2, among the 24 studies in our survey that report (non-mixed) aggregate results from OH designs, 21 (88 percent) find evidence interpreted as declining discounting, while 3 (12%) find evidence consistent with constant discounting. By contrast, among the 10 studies that use SH designs and report aggregate results on discounting types, only half find declining discounting, with the rest finding constant discounting (OSH designs are almost never analyzed in the literature). This particular pattern of mixed results across studies has received little attention in the literature, but can be explained by pervasive subadditivity. Notably, although OH designs are prone

Table 7: Results by Measurement Approach

Approach	Horizons	Sample	Aggregate result	Discounting types (percent)			Med. parameter est. (quasi-hyperbolic)	
				constant	declining	increasing	β	δ
OH	$T_{0,6}$ vs. $T_{0,12}$	CSS	declining	22.39	60.70	16.92	0.97	0.92
SH	$T_{0,6}$ vs. $T_{6,12}$	CSS	constant	27.89	34.37	37.75	1.00	0.81
OSH	$T_{0,12}$ vs. $T_{6,12}$	CSS	increasing	15.92	19.15	64.93	1.05	0.81
OH	$T_{0,6}$ vs. $T_{0,12}$	SOEP	declining	25.92	54.28	19.80	0.9998	0.92
OH	$T_{0,1}$ vs. $T_{0,12}$	SOEP	declining	12.97	76.06	10.97	0.9959	0.94
SH	$T_{0,1b}$ vs. $T_{12,13}$	SOEP	constant	32.91	32.91	34.18	1.0000	0.71
OSH	n.a.	SOEP	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

Notes: Aggregate results for OH: Declining means that the average annual IRR for the short horizon is longer than for the long horizon, difference significant at least at the five-percent level. Individual-level classification for OH: Constant means IRRs are the same for both horizons, declining means the IRR is higher for the shorter horizon, increasing means IRR is higher for the longer horizon. Aggregate results for SH: Constant means no significant difference in annual IRR across the earlier and later horizons. Individual-level classifications for SH: Constant means IRRs are the same for both horizons, declining means a higher IRR for the early horizon, increasing means a higher IRR for the later horizon. Aggregate results for OSH: Increasing means higher average annual IRR for the shorter horizon than the long horizon, with the difference significant at least at the five-percent level. Individual-level classification for OSH: Constant means the same IRRs for both horizons, declining means a higher IRR for the longer horizon, increasing means a higher IRR for the shorter horizon. Parameters for the quasi-hyperbolic model are calculated assuming linear utility. For the SOEP data, entries for OSH are n.a. because this type of comparison was not part of the experimental design.

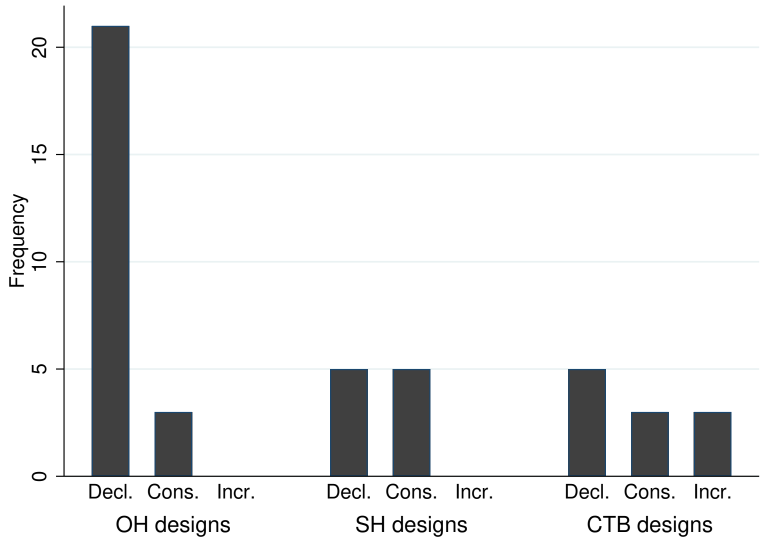
to confound declining discounting and subadditivity, OH designs have been much more commonly used than SH designs, by a ratio of about 2 to 1 (see Figure Figure 2). This is also the case considering only the most recent five years in our survey, 2015 to 2020.[45] Our results point to one advantage of SH designs that is not typically discussed, which is that the categorization of individuals as present biased or future biased in SH designs will not be affected by nonadditivity, since SH designs hold time horizon length constant.

Our findings also imply a method variance in different approaches to estimating quantitative time preference parameters. If additivity is violated, this can influence the estimates that such studies find, and different design choices can lead to systematically different parameter estimates. We focus here on intuition, but in Appendix F we provide a more detailed demonstration of how preference parameter estimates in different designs can be influenced by subadditivity.

One type of study has sought to estimate the exponential discount rate from a single time horizon (see Matousek et al., 2022, for a meta-analysis). While the choice of time horizon length is irrelevant in the DU model, the horizon length could matter if individuals are nonadditive. For example, as we show in the appendix, if individuals are subbaditive,

[45] As summarized in Table A, studies using OH designs are more frequent than SH designs by a ratio of about 2 to 1 whether looking at the entire survey period, the years since 2010, or the years since 2015.

Figure 2: Frequencies of aggregate findings of constant, declining, and increasing discounting in the literature, by design type



Notes: Frequencies are drawn from the literature survey presented in Table A.1, and include studies that provide aggregate results on discounting pattern. Cons., decl., and incr. indicate aggregate findings of constant, declining, and increasing discounting patterns, respectively. Aggregate results for OH and SH designs are based on comparing the magnitudes of average discount rates for short versus long, and early versus late time horizons, respectively. Results for CTB designs are categorized based on whether the average parameter estimate for β in the quasi-hyperbolic model is equal to 1 (constant), less than 1 (declining), or greater than 1 (increasing). The figure excludes a few studies that have mixed results on aggregate discounting pattern across, e.g., different parameters (1 from OH, 2 from SH, and 1 from the CTB design categories).

using shorter horizons can lead to higher estimates of the exponential discount rate than using longer horizons. This is consistent with what we see in our data, e.g., as reported in Section 3, in the CSS data the implied median annual discount rate goes from 27.5% for the 12-month horizon to 37.5% for the 6-month horizon. It is also in line with the general tendency for studies that use shorter time horizons to find higher exponential discount rates, noted by Frederick et al. (2002).

Another type of study has used OH or SH designs to estimate parameters of the quasi-hyperbolic model. In OH designs, both β and δ are inferred from comparing two time horizons of different lengths, and thus nonadditivity can influence estimates through a channel that is outside of the DU model. Also, the particular lengths of time horizons chosen can matter systematically for results. In SH designs, by contrast, β estimates will not be affected by nonadditivity, because these are based on two horizons of the same length, and the length of horizon will not affect estimates. The level of δ, however, will be influenced by nonadditivity, and vary with the length of time horizons that is used. We show these results in more detail in the appendix. Our data illustrate empirically how preference parameters can vary across measures. As seen in Table 7, the parameter β is below 1 for OH measures, implying present bias, but equal to 1 for SH measures, implying constant discounting. For our OSH comparison, the parameter is greater than 1, implying future bias. The parameter δ also varies across the measures. As expected based on subadditivity, by far the lowest estimate, 0.71, is for the SH comparison in the SOEP data, which uses the shortest length time horizons, i.e., $T_{0,1}$ vs. $T_{12,13}$.

Our results also highlight the potential for parameter estimates to vary across studies using CTB designs, depending on the particular combinations of time horizons used. The main innovation in the CTB approach was to have subjects face a range of different interest rates for allocating money between earlier and late periods, thereby calibrating curvature of utility and relaxing the traditional assumption in the literature of linear utility. A feature of the way this design has typically been implemented, however, which has received less attention, is that studies have typically used four or more time horizons for the same individual, rather than only two horizons as is the case in OH or SH designs.[46] Indeed, most CTB studies choose time horizons that can allow both OH and SH comparisons, but these are pooled for the purpose of estimating preference parameters. The findings from the 11 CTB studies in our literature survey that report aggregate results are mixed, with 5 (45%) finding declining discounting on average, 3 (27%) finding constant discounting, and the remaining 3 finding increasing discounting.[47].

If additivity is violated, then the impact on preference parameter estimates in CTB designs depends on details such as whether there are equal numbers of horizons that involve the present, and that do not involve the present. With equal numbers, estimates of the short-term discount factor, β, will not be affected by nonadditivity, while estimates of the long-term discount factor, δ, will be affected, and vary systematically with the lengths of

[46] For example, the seminal paper, Andreoni and Sprenger (2012), had nine time horizons, and Andreoni et al. (2015) had four. Many of the more recent studies using the CBT design have adopted a set of four horizons similar to those in Andreoni et al. (2015).

[47] Our survey overlaps with the meta-analysis of Imai et al. (2021) on CTB designs, but they include some additional papers that were not easily accessible to us. They report that about 50% of CTB studies find present bias at the aggregate level (estimated β less than 1), so conclusions are very similar to our survey.

time horizons chosen. If the design involves unequal numbers, e.g., more horizons that start in the future than horizons that start in the present, however, then nonadditivity can influence estimates of both β and δ, and both can depend on the lengths of time horizons used.[48]

5.2 Modifying designs to allow testing additivity

A second main methodological implication of our findings is a value of building-in the possibility to test additivity in intertemporal choice experiments. For SH, OH, or OSH designs, without a third time horizon, appropriately chosen to test additivity, it is unclear how to interpret results. If additivity is violated, then patterns of, e.g., present bias in the SH comparison cannot be rationalized by declining discounting in the DU model. Some other type of model of present bias is needed. Results from our CSS data illustrate that testing for additivity can be quite revealing: Among those who exhibit present bias patterns in the SH comparison, only 25 percent satisfy additivity; for those exhibiting a present biased choice pattern in the OH comparison, the corresponding fraction is 16 percent. Thus, for both types of designs, adding an additional time horizon to test for additivity might substantially change conclusions about how many individuals are consistent with a given traditional discount function in the DU model. CTB designs can also benefit from including the combinations of time horizons needed to assess whether additivity is violated. If additivity is violated, then this calls into question the interpretation of behavior in the framework of the DU model with quasi-hyperbolic discounting. While CTB designs have typically included several time horizons, the set needs to include appropriately chosen subintervals in order to allow testing additivity.

[48] The set of time horizons used by Andreoni et al. (2015) has been widely adopted in subsequent literature: today to 5 weeks; today to 9 weeks; 5 weeks to 10 weeks; 5 weeks to 14 weeks. The δ parameter is identified from the two horizons that do not involve the present, but these are of different lengths, with one being a subinterval of the other, so violation of additivity can influence the estimate of δ and the estimate can be affected by the particular lengths being used. β is identified by contrasting horizons that involve the present versus those that do not. In this particular set of horizons, the effect of nonadditivity will be differenced out and will not matter for the estimate of β. If, however, the design included additional time horizons starting in the future, then the estimate of β could be influenced by nonadditivity.

6 Conclusion

This paper provides evidence that nonadditivity is a robust and pervasive feature of intertemporal choice in the general population. The analysis addresses potentially important confounds, which have been raised about intertemporal choice experiments, and the findings suggest that these do not drive the subadditive choice pattern. The findings have implications for economic theory, underlining the value of research on models that can incorporate both subadditivity and other biases that have traditionally been modeled in the DU framework. They also raise caveats about the ways that time horizon effects in intertemporal choice experiments have traditionally been interpreted. This latter implication may help explain mixed results from previous studies regarding the prevalence of discounting types and the levels of preference parameter estimates. A methodological implication is that future intertemporal choice experiments should include appropriate time horizons to allow testing additivity.

A final point relates to the potential policy relevance of the findings. One implication concerns the on-going debate about policy interventions designed to address self-control problems (e.g., Camerer et al., 2003). Such policies are often motivated with reference to models with declining (hyperbolic or quasi-hyperbolic) discounting, and the impact of such policies depends crucially on the distribution of discounting types in the population. Intertemporal choice experiments have been a commonly used tool for measuring the prevalence of declining discounting. Our results raise caveats, however, about the typical experimental design for measuring the distribution of types, which has used choices over two time horizons, elicited in the present, to try to trace out the shape of a time discount function. The observed impact of time horizon on choices may reflect a mechanism that is outside of the hyperbolic or any other standard discounting model. Instead, combining multiple types of time horizon comparisons, or using alternative types of experiments, may be better suited to identify present bias. A second policy-related implication of the paper is that a different bias in decision making, subadditivity, is also worth consideration by policy makers it its own right, as it affects many people in the population. Subadditivity implies that framing time horizons in more narrow or more broad ways my influence how people make choices. This can be relevant when policy makers design and advertise different types of packages of benefits or services to be received or payed for over time, and potentially for assessing the welfare effects of different ways that firms might describe choices to consumers.

References

Andersen, S., G. Harrison, M. Lau, and E. Rutstroem (2011): "Discounting Behavior and the Magnitude Effect: Evidence from and Field Experiment in Denmark," *Economica*, 80(320), 670–697.

Augenblick, N., and M. Rabin (2019): "An Experiment on Time Preference and Misprediction in Unpleasant Tasks.," *Review of Economic Studies*, 86(3).

Baron, J. (2000): "Can we use human judgments to determine the discount rate?," *Risk Analysis*, 20(6), 861–868.

Benhabib, J., A. Bisin, and A. Schotter (2010): "Present-bias, quasi-hyperbolic discounting, and fixed costs," *Games and Economic Behavior*, 69(2), 205–223.

Camerer, C., S. Issacharoff, G. Loewenstein, T. O'Donoghue, and M. Rabin (2003): "Regulation for Conservatives: Behavioral Economics and the Case for" Asymmetric Paternalism"," *University of Pennsylvania Law Review*, 151(3), 1211–1254.

Cohen, J., K. M. Ericson, D. Laibson, and J. M. White (2020): "Measuring time preferences," *Journal of Economic Literature*, 58(2), 299–347.

Coller, M., and M. Williams (1999): "Eliciting individual discount rates," *Experimental Economics*, 2(2), 107–127.

Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2010): "Are Risk Aversion and Impatience Related to Cognitive Ability?," *American Economic Review*, 100(3), 1238–1260.

Ebert, J., and D. Prelec (2007): "The fragility of time: Time-insensitivity and valuation of the near and far future," *Management Science*, 53(9), 1423–1438.

Frederick, S., G. Loewenstein, and T. O'donoghue (2002): "Time discounting and time preference: A critical review," *Journal of Economic Literature*, 40(2), 351–401.

Fudenberg, D., and D. Levine (2012): "Timing and Self-Control," *Econometrica*, 80(1), 1–42.

Goebel, J., M. M. Grabka, S. Liebig, S. Kroh, C. S. David Richter, and J. Schupp (2019): "The German Socio-Economic Panel (SOEP)," *Jahrbücher für Nationalökonomie und Statistik*, 239(2), 345–360.

Halevy, Y. (2015): "Time consistency: Stationarity and time invariance," *Econometrica*, 83(1), 335–352.

Harrison, G. W., M. I. Lau, and M. B. Williams (2002): "Estimating Individual Discount Rates in Denmark: A Field Experiment," *American Economic Review*, 92(5), 1606–1617.

Henrich, J., S. J. Heine, and A. Norenzayan (2010): “Most people are not WEIRD,” *Nature*, 466(7302), 29–29.

Laibson, D. (1997): “Golden eggs and hyperbolic discounting,” *Quarterly Journal of Economics*, 112(2), 443–477.

Matousek, J., T. Havranek, and Z. Irsova (2022): “Individual discount rates: a meta-analysis of experimental evidence,” *Experimental Economics*, 25(1), 318–358.

McAlvanah, P. (2010): “Subadditivity, patience, and utility: The effects of dividing time intervals,” *Journal of Economic Behavior and Organization*, 76(2), 325–337.

Meier, S., and C. Sprenger (2010): “Present-biased preferences and credit card borrowing,” *American Economic Journal: Applied Economics*, 2(1), 193–210.

O’Donoghue, T., and M. Rabin (1999): “Doing it now or later,” *American Economic Review*, 89(1), 103–124.

Ok, E., and Y. Masatlioglu (2007): “A Theory of (Relative) Discounting,” *Journal of Economic Theory*, pp. 214–245.

Phelps, E., and R. Pollak (1968): “On second-best national saving and game-equilibrium growth,” *Review of Economic Studies*, 35(2), 185–199.

Read, D. (2001): “Is time-discounting hyperbolic or subadditive?,” *Journal of Risk and Uncertainty*, 23(1), 5–32.

Read, D., M. Airoldi, and G. Loewe (2005): “Intertemporal tradeoffs priced in interest rates and amounts: A study of method variance,” *working paper, London School of Economics and Political Science*.

Read, D., and P. Roelofsma (2003): “Subadditive versus hyperbolic discounting: A comparison of choice and matching,” *Organizational Behavior and Human Decision Processes*, 91(2), 140–153.

Rubinstein, A. (2003): ““Economics and Psychology?” The Case of Hyperbolic Discounting,” *International Economic Review*, 44(4), 1207–1216.

Samuelson, P. A. (1937): “A note on measurement of utility,” *The review of economic studies*, 4(2), 155–161.

Scholten, M., and D. Read (2006): “Discounting by intervals: A generalized model of intertemporal choice,” *Management Science*, 52(9), 1424–1436.

Stango, V., and J. Zinman (2009): “Exponential growth bias and household finance,” *The Journal of Finance*, 64(6), 2807–2849.

Strack, P., and D. Taubinsky (2021): “Dynamic Preference “Reversals” and Time Inconsistency,” Discussion paper, National Bureau of Economic Research.

Thompson, S. K. (2006): “Targeted Random Walk Designs,” *Survey Methodology*, 32(1), 11–24.

Vieider, F. M. (2021): “Noisy coding of time and reward discounting,” *Working Paper, Ghent University*.

Zauberman, G., B. Kim, S. Malkoc, and J. Bettman (2009): "Discounting time and time discounting: Subjective time perception and intertemporal preferences," *Journal of Marketing Research*, 46(4), 543–556.

Appendix

A Literature survey

We conducted a survey of the literature in economics on intertemporal choice experiments. The main purposes of our survey were twofold: (1) To document that the assumption of additivity continues to be largely unexamined in the economics literature; (2) to report the frequency of different types of experimental designs, and how findings vary with design choice in ways that can be explained by subadditivity.

We started our survey in the year 2000 because that was the publication date of seminal papers on subadditive discounting (Barron, 2000; Read, 2001). Our literature survey has substantial overlap with two recent meta-analyses, Matousek et al. (2022) and Imai et al. (2021), which consider single-horizon experiments measuring the exponential discount factor, and experiments using CTB designs to measure parameters of the quasi-hyperbolic model, respectively. The main difference in the selection of studies is that the former includes studies before 2000, and the latter includes some studies that were not easily accessible to us online. Notably, the proportion of CTB studies in Imai et al. (2021) finding evidence of present bias on average is roughly 50%, similar to what we find in our survey, so the difference in the sample does not strongly influence conclusions. Cohen et al. (2020) provide another survey of studies on time discounting; the set of studies overlaps with ours, but they consider a time period that starts earlier and ends earlier, and the survey is not focused on documenting results by OH versus SH design. Frederick et al. (2002) provide a survey of studies before 2000, of which the vast majority are OH designs finding evidence of declining discounting.

Table A.1: Intertemporal choice experiments from recent decades

Study	Year	Design type	Tests for nonadditivity	SH aggregate finding	OH aggregate finding	CTB aggregate finding
Harrison et al. (2002)	2002	OH	No		Declining	
Rubinstein (2003)	2003	SOH	No			
McClure et al. (2004)	2004	OH;SH	No			
Coller et al. (2005)	2005	OH, SH, SOH	No		Constant	
Eckel et al. (2005)	2005	OH	No			
Harrison et al. (2005)	2005	Other	No			
Andersen et al. (2006)	2006	OH	No			
Khwaja et al. (2007)	2007	OH	No		Declining	
McClure et al. (2007)	2007	SH;OH	No	Declining	Declining	
Slonim et al. (2007)	2007	SH	No	Declining		
Andersen et al. (2008)	2008	OH	No		Declining	
Chabris et al. (2008)	2008	OH	No		Declining	
Chabris et al. (2008b)	2008	OH	No		Declining	
Frederick and Loewenstein (2008)	2008	OH	No			
Abdellaoui et al. (2009)	2009	OH	No		Declining	
Booij and van Praag (2009)	2009	Other	No			
Brown et al. (2009)	2009	Other	No			
Chabris et al. (2009)	2009	OH	No			
Engle-Warnick et al. (2009)	2009	OH	No			
Andersen et al. (2010)	2010	OH	No		Declining	
Bauer and Chytilova (2010)	2010	Single horizon	No			
Benhabib et al. (2010)	2010	OH	No		Declining	
Dohmen et al., (2010)	2010	Single horizon	No			
Manzini et al. (2010)	2010	OH	No			
McAlvanah (2010)	2010	SH;OH;SOH	Yes	Declining	Declining	
Tanaka et al. (2010)	2010	OH	No		Declining	
Ifcher and Zaghamee (2011)	2011	OH	No		Declining	
Takeuchi (2011)	2011	Other	No			
Andreoni and Sprenger (2012a)	2012	CTB	No			Constant
Andreoni and Sprenger (2012b)	2012	OH	No			
Bauer et al. (2012)	2012	SH	No	Declining		
Burks et al. (2012)	2012	OH	No		Constant	
Carlsson et al. (2012)	2012	OH,SH	No	Constant	Declining	
Coller et al. (2012)	2012	OH	No		Declining	
Duquette et al. (2012)	2012	Single horizon	No			
Laury et al. (2012)	2012	Other	No			
Abdellaoui et al. (2013)	2013	OH	No		Declining	
Abdellaoui et al. (2013b)	2013	Other	No			
Andersen et al. (2013)	2013	OH	No			
Bauer and Chytilova (2013)	2013	SH	No	Declining		
Field et al. (2013)	2013	Single horizon	No			
Finke and Huston (2013)	2013	Single horizon	No			
Meier and Sprenger (2013)	2013	OH,SH	No			
Sutter et al. (2013)	2013	OH,SH	No	Constant	Declining	
Andersen et al. (2014)	2014	OH	No		Declining	
Jackson and Yariv (2014)	2014	Other	No			
Montiel-Olea and Strzalecki (2014)	2014	Other	No			
Alan and Ertac (2015)	2015	SH	No			
Andreoni et al. (2015)	2015	CTB	No			Declining
Augenblick et al. (2015)	2015	CTB	No			Constant
Aycinena et al. (2015)	2015	CTB	No			Increasing
Deck and Jahedi (2015)	2015	Single horizon	No			
Deck and Jahedi (2015b)	2015	Single horizon	No			
Halevy (2015)	2015	SH	No	Constant		
Meier and Sprenger (2015)	2015	OH,SH	No			
Miao and Zhong (2015)	2015	CTB	No			
Newell and Siikamaki (2015)	2015	Single horizon	No			
Swada and Kuroishi (2015)	2015	CTB	No			Declining
Attema et al. (2016)	2016	SH,OH,SOH	No	Mixed	Mixed	
Carvalho et al. (2016)	2016	CBT	No			
Freeman et al. (2016)	2016	OH	No			
Olivola and Wang (2016)	2016	Other	No			
Schreiber and Weber (2016)	2016	SH	No	Constant		
Andreoni et al. (2017)	2017	Single horizon	No			
Cassar et al. (2017)	2017	OH	No		Constant	
Cubitt et al. (2017)	2017	OH	No			
Janssens et al. (2017)	2017	CTB	No			Mixed
Alan and Ertac (2018)	2018	CTB	No			
Andersen et al. (2018)	2018	OH	No		Declining	
Aycenina and Rentschler (2018)	2018	CTB	No			Increasing
Banerji et al. (2018)	2018	CTB	No			Declining
Bradford et al. (2017)	2018	OH,SH	No			
Brocas et al. (2018)	2018	CTB	No			Constant
Gine et al., (2018)	2018	SH	No	Constant		
Luehramnn et al. (2018)	2018	CTB	No		Declining	
Amasino et al. (2019)	2019	OH	No			
Augenblick and Rabin (2019)	2019	Other	No			
Bradford et al. (2019)	2019	OH	No		Declining	
Aycinena et al. (2020)	2020	CTB	No			Increasing
Balakrishnan et al. (2020)	2020	CTB	No			Declining
DeJarnette et al. (2020)	2020	OH	No			
Abebe et al. (2021)	2021	CTB	No			Declining
Vieider (2021)	2021	SH, OH, SOH	Yes	Mixed	Declining	

Notes: The survey is composed of 83 studies that use intertemporal choice experiments to study time discounting. Studies that do not report corresponding aggregate results have empty entries.

B Additional results on nonadditivity, bias type, and economic outcomes

Table A.1: Relationships of economic outcomes to nonadditivity and bias type

	BMI	Smoker	Drinking	Poor nutrition	Health	Happiness	Overdrawn	Ln income	Ln wealth
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Nonadditive	-0.36	0.03	0.08	0.18**	-0.33	-0.22	0.04	-0.11	-1.26*
	(0.53)	(0.06)	(0.12)	(0.08)	(0.26)	(0.20)	(0.06)	(0.08)	(0.70)
Present biased	0.82	0.14**	-0.11	0.04	-0.51*	-0.13	0.06	0.03	-1.27
	(0.61)	(0.07)	(0.14)	(0.10)	(0.28)	(0.23)	(0.07)	(0.08)	(0.84)
Future biased	0.40	0.12*	-0.22*	0.13	-0.20	-0.26	-0.01	-0.14	0.78
	(0.59)	(0.06)	(0.13)	(0.10)	(0.29)	(0.23)	(0.06)	(0.09)	(0.72)
Female	-1.55***	-0.13**	-0.55***	-0.38***	-0.02	0.09	-0.01	-0.09	-0.78
	(0.49)	(0.05)	(0.11)	(0.08)	(0.24)	(0.19)	(0.05)	(0.07)	(0.62)
Age	0.28***	0.01	-0.01	-0.01	-0.05	-0.03	0.02***	0.01	0.32***
	(0.06)	(0.01)	(0.02)	(0.01)	(0.03)	(0.02)	(0.01)	(0.01)	(0.11)
Age^2	-0.00***	-0.00*	0.00	0.00	0.00	0.00	-0.00***	-0.00	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Educated	-0.30	-0.15**	0.26*	-0.34***	0.58**	0.35*	-0.07	0.08	1.77***
	(0.58)	(0.06)	(0.13)	(0.09)	(0.25)	(0.18)	(0.06)	(0.09)	(0.63)
IQ	-0.08	0.05	0.03	0.01	0.33*	0.09	0.05	-0.02	0.10
	(0.34)	(0.03)	(0.07)	(0.06)	(0.18)	(0.16)	(0.04)	(0.06)	(0.41)
Thought int. rate	-0.31	0.01	0.08	-0.19**	0.09	0.06	-0.05	0.09	0.13
	(0.52)	(0.06)	(0.12)	(0.08)	(0.25)	(0.20)	(0.05)	(0.07)	(0.67)
Constant	18.61***	0.30*	1.55***	2.94***	9.42***	8.70***	-0.07	7.45***	-0.13
	(1.27)	(0.18)	(0.40)	(0.29)	(0.72)	(0.54)	(0.15)	(0.27)	(2.61)
Observations	319	326	326	325	326	326	311	310	252
Adjusted R^2	0.123	0.072	0.081	0.134	0.135	0.011	0.071	0.021	0.138

Notes: Notes: OLS estimates. BMI denotes body mass index; Smoker is an indicator for smoking more than occasionally; Drinking is an index for frequency of drinking (combining information about wine, spirits, beer, and mixed drinks), with higher values indicating more frequent drinking; Poor Nutrition is self-reported degree of poor nutrition, on a 4-point scale, with higher values indicating worse nutrition; Health is self-reported satisfaction with health; Happiness measures subjective well-being; Overdrawn is days per year that current deposit account is overdrawn; Ln income is the natural log of household income; Ln wealth is the natural log of household wealth. Nonadditive is an indicator for violating additivity. Present bias and Future bias are indicators for having some form of corresponding bias pattern, as defined in Table 4. Control variables include an indicator for being highly educated (having taken the *Abitur*, a college entrance exam in Germany); the average of an individual's score on two different IQ tests; an indicator for financial sophistication, proxied by reporting having thought about market interest rates during the choice experiments. The sample excludes individuals for whom censoring makes additivity unambiguous. Robust standard errors are in parentheses. *, ** indicates significance at 10 and 5 percent level.

C Predictions for Fixed-Cost Specification of the Quasi-Hyperbolic Model

The fixed-cost version of the quasi-hyperbolic model allows this model to predict a "magnitude effect," a tendency for measured impatience to decrease as stake sizes increase, *ceteris paribus*. Besides predicting a magnitude effect, however, this model has the same properties as the quasi-hyperbolic model: There is an extra cost of waiting, above and beyond exponential discounting, when comparing the present to the next future period. When comparing two adjacent, future periods, however, the extra cost applies to both periods, and thus cancels out, leaving discounting to be governed solely by the exponential discount rate. The quasi-hyperbolic model with fixed costs thus predicts $IRR^{T_{0,6}} > IRR^{T_{0,12}} > IRR^{T_{6,12}}$ if the present extends further than two days into the future, and invariance of the IRR to time horizon, $IRR^{T_{0,6}} = IRR^{T_{0,12}} = IRR^{T_{6,12}}$, if it does not, the same as the standard quasi-hyperbolic model.

More formally, in this model a payment Z_t received in the future is discounted by $\Delta_t = \delta^t - \frac{b}{Z_t}$, where $\delta = \frac{1}{1+\rho}$ is the standard exponential discount factor and $b > 0$ is a fixed cost of having a payment arrive in the future, which goes to zero as the stakes in the experiment increase. A payment received at $t = 0$ is not discounted, i.e., $\Delta_0 = 1$. The present value of a future payment is thus given by $\Delta_t Z_t = \delta^t Z_t - b$. Importantly, between-period discounting in the future is the same as in the exponential model (and as in the quasi-hyperbolic model with variable costs). I.e., the Euler equation for consumption in future periods $t+1$ and $t+2$ is $\delta^{t+1} Z_{t+1} - b = \delta^{t+2} Z_{t+2} - b \Leftrightarrow Z_{t+1} = \delta Z_{t+2}$. Maintaining the assumption that the present extends more than two days into the future, the indifference conditions implied by choices are given by

$$\begin{aligned}
\left(1+\tfrac{\rho}{2}\right)100 + \left(1+\tfrac{\rho}{2}\right)b &= Z^{T_{0,6}} \\
\left(1+\tfrac{\rho}{2}\right)^2 100 + \left(1+\tfrac{\rho}{2}\right)^2 b &= Z^{T_{0,12}} \\
\left(1+\tfrac{\rho}{2}\right)100 &= Z^{T_{6,12}}
\end{aligned}$$

Substituting into (2) yields

$$\begin{aligned}
IRR^{T_{0,6}} &= 2\left[\left(1+\tfrac{b}{100}\right)\left(1+\tfrac{\rho}{2}\right) - 1\right] \\
IRR^{T_{0,12}} &= 2\left[\left(1+\tfrac{b}{100}\right)^{\frac{1}{2}}\left(1+\tfrac{\rho}{2}\right) - 1\right] \\
IRR^{T_{6,12}} &= 2\left[\left(1+\tfrac{\rho}{2}\right) - 1\right]
\end{aligned}$$

These equations imply $IRR^{T_{0,6}} > IRR^{T_{0,12}} > IRR^{T_{6,12}}$. If, instead, the present-bias falls within the window of two days from the present, all terms involving b are eliminated, the conditions reduce to those in (1), and the prediction implies invariance of IRR with respect to time horizon, $IRR^{T_{0,6}} = IRR^{T_{0,12}} = IRR^{T_{6,12}}$. Qualitative predictions are similar using the alternative specification of the model, discussed by Benhabib et al. (2010), where exponential discounting is applied to b as well as the future payment.

D Demonstration of transitivity violation

To see more formally why the aggregate choice pattern in the CSS data implies a violation of transitivity, note that $T_{0,6}$, $T_{0,12}$, and $T_{6,12}$ pose individuals with choices between different bundles that involve payments received at 0, 6, or 12 months. In the data, we observe for each time horizon the late payment amounts $Z^{T_{0,6}}$, $Z^{T_{0,12}}$, and $Z^{T_{6,12}}$ such that an individual is indifferent between the bundle involving the early payment and the bundle involving the late payment. The table below shows these particular bundles for each horizon:

$T_{0,6}$		$T_{0,12}$		$T_{6,12}$	
$A:$	$(100, 0, 0)$	$A:$	$(100, 0, 0)$	$D:$	$(0, 100, 0)$
$B:$	$(0, Z^{T_{0,6}}, 0)$	$C:$	$(0, 0, Z^{T_{0,12}})$	$E:$	$(0, 0, Z^{T_{6,12}})$

where individuals have preferences $A \sim B$, $A \sim C$, and $D \sim E$ when presented with each choice pair in isolation. For now, we are assuming that utility is linear, and $u(x) = x$.

An inconsistency arises, however, when one compares across choice pairs. From $T_{0,12}$, we observe the $Z^{T_{0,12}}$ that makes an individual indifferent between receiving 100 at time 0 or $Z^{T_{0,12}}$ in 12 months

$$Z^{T_{0,12}} = 100 \left(1 + \frac{IRR^{T_{0,12}}}{2}\right)^2. \tag{A.1}$$

Combining the observed IRRs for $T_{0,6}$ and $T_{6,12}$, we can also construct a late payment $Z^{T'_{0,12}}$ that should make the individual indifferent between 100 at time 0 or waiting 12 months to get $Z^{T'_{0,12}}$

$$Z^{T'_{0,12}} = 100 \left(1 + \frac{IRR^{T_{0,6}}}{2}\right) \left(1 + \frac{IRR^{T_{6,12}}}{2}\right). \tag{A.2}$$

For the typical individual, however, the differences in IRRs that we observe for short and long time horizons imply:

$$\left(1 + \frac{IRR^{T_{0,12}}}{2}\right)^2 < \left(1 + \frac{IRR^{T_{0,6}}}{2}\right) \left(1 + \frac{IRR^{T_{6,12}}}{2}\right). \tag{A.3}$$

This is a violation of additivity in the direction of subadditivity (see condition (4) in Section 2.3), and implies in combination with (A.1) and (A.2) that:

$$Z^{T_{0,12}} < Z^{T'_{0,12}}. \tag{A.4}$$

To see how this translates into a violation of transitivity, consider what the 6-month measures imply about the following choices:

$A:$	$(100, 0, 0)$	$C':$	$(0, 0, Z^{T'_{0,12}})$
$C':$	$(0, 0, Z^{T'_{0,12}})$	$C:$	$(0, 0, Z^{T_{0,12}})$

By the definition of $Z^{T'_{0,12}}$, we have $A \sim C'$, and given $Z^{T'_{0,12}} > Z^{T_{0,12}}$ and monotonicity of utility in money, we have $C' \succ C$. Transitivity then requires $A \succ C$, but behavior in $T_{0,12}$ implies $A \sim C$, a violation of transitivity.

There is also a violation of transitivity allowing for a nonlinear (monotonic) utility function $u(\cdot)$, as long as (A.3) still holds. As discussed in the next appendix, (A.3) holds given the behavior we observe, unless the unobserved utility function is characterized by sufficiently extreme convexity.

E Implications of nonlinear utility for apparent additivity violations

In this section we show that nonlinear utility can rationalize the observed aggregate results without violating additivity, but only if utility is sufficiently convex.

The standard DU model involves the following intertemporal utility function

$$U(Z_t, ...Z_T) = \sum_{k=0}^{T-t} D(k)u(Z_{t+k}) \tag{A.1}$$

Where the discount function, D, is given by:

$$D(k) = \prod_{n=0}^{k-1} D(n) \tag{A.2}$$

As pointed out by Read (2001), regardless of the shape of the discounting function (i.e., whether the discount factor is constant or non-constant over time), discounting over an interval k can be expressed as the product of discounting over the course of subintervals of k.

$$D(k) = \prod_{0}^{t} D(n) \prod_{t+1}^{k} D(n) \tag{A.3}$$

In our study we measure discounting between today and 12 months, and over the subintervals today to 6 months, and 6 months to 12 months. Treating 6 months as a period we have:

$$D(2) = \prod_{0}^{1} D(n) \prod_{1}^{2} D(n) \tag{A.4}$$

We measure the discount factors by seeing the delayed payment that makes an individual indifferent to an earlier payment of 100 and the delayed payment.

$$D(2) = \tfrac{U(100)}{U(Z_{0,2})} \qquad D(01) = \tfrac{U(100)}{U(Z_{0,1})} \qquad D(12) = \tfrac{U(100)}{(U(Z_{1,2})} \tag{A.5}$$

Substitution into A.8 yields:

$$\frac{U(100)}{U(Z_{0,2})} = \frac{U(100)}{U(Z_{0,1})} \cdot \frac{U(100)}{U(Z_{1,2})} \tag{A.6}$$

Plugging in the median delayed payments that induce indifference in our experiments, the following must hold if additivity is satisfied:

$$\frac{U(100)}{U(130)} = \frac{U(100)}{U(120)} \cdot \frac{U(100)}{U(120)} \tag{A.7}$$

Now we determine whether a concave shape of $U()$ can rationalize the observed delayed payments and satisfy additivity. Suppose that $U()$ is always positive, continuous and monotonic. Taking logs:

$$\ln U(100) - \ln U(130) = 2\ln U(100) - 2\ln U(120) \tag{A.8}$$

$$\ln U(120) = 0.5 \ln U(100) + 0.5 \ln U(130) \tag{A.9}$$

Now suppose $U()$ is concave. $\ln$ is a concave function, so $\ln U()$ is concave. By concavity, the following is true:

$$\ln U(115) > 0.5 \ln U(100) + 0.5 \ln U(130) \tag{A.10}$$

Given monotonicity of $U()$, this implies:

$$\ln U(120) > 0.5 \ln U(100) + 0.5 \ln U(130) \tag{A.11}$$

Which implies a contradiction to A.13. Thus, for any concave (positive, monotonic) utility function the observed median delayed payments are incompatible with the condition for additivity as expressed by A.11 above. The same follows trivially for linear utility. For a sufficiently convex $U()$, however, we have $\ln U()$ being convex, and A.11 could hold, i.e., the observed aggregate results could be consistent with additivity.

F Implications of nonadditivity for standard approaches to estimating time preference parameters

In this appendix we derive formulas for preference parameters under different types of experimental designs, and discuss the implications of nonadditivity.

F1 Single time horizon estimates of exponential discount rate

A standard approach to estimating the constant discount factor (or equivalently discount rate) in the DU model is to assume linear utility and consider the delayed payment needed to induce waiting over a single time horizon. The length of time horizon chosen does not matter in the context of the DU model, as additivity is assumed to hold. We demonstrate here that if individuals are subadditive according to the functional form proposed by Read (2001), then the discount factor obtained from the DU model will increase with the length of the time horizon used.

Consider the time horizon from today to 6 months. Observing the minimum amount needed to make the individual willing to wait, $Z^{T_{0,6}}$, implicitly defines the annual discount factor, δ, assuming constant discounting and semi-annual compounding:

$$100 = \delta^{\frac{1}{2}} Z^{T_{0,6}} \tag{A.1}$$

$$\delta = \left(\frac{100}{Z^{T_{0,6}}} \right)^2 \tag{A.2}$$

The discount factor must be invariant to the time horizon used, if the model is correct, so one could as well use the time horizon today to 12 months to calculate the annual discount factor:

$$100 = \delta Z^{T_{0,12}} \tag{A.3}$$

$$\delta = \left(\frac{100}{Z^{T_{0,12}}} \right) \tag{A.4}$$

To be consistent with the DU model the observed delayed payments must vary with time horizon length in the way that will yield the same annual δ.

Now suppose instead that the individual is subadditive according to the functional form proposed by Read (2001). Time preference is captured by the same discount factor δ, but nonlinear (concave) perception of time duration, captured by an additional parameter σ, can cause discounting to depend on time horizon length. Specifically, the indifference relation and implied annual discount factor for the shorter time horizon are given as follows:

$$100 = \delta^{\frac{1}{2}^{\sigma}} Z'^{T_{0,6}} \tag{A.5}$$

$$\delta = \left(\frac{100}{Z'^{T_{0,6}}} \right)^{\frac{1}{\frac{1}{2}^{\sigma}}} \tag{A.6}$$

Note that because $0 < \sigma \leq 1$, the exponent in (A.6) is smaller than the exponent in (A.2), the formula for δ under additivity. This means that the delayed payment that satisfies A.21 is larger than the one satisfying A.17, i.e., $Z'^{T_{0,6}} > Z^{T_{0,6}}$. With the delayed payment in the denominator, this means that subadditivity is decreasing the discount factor that will be obtained if one applies the DU model with constant discounting to this time horizon. Now consider behavior in the longer time horizon with subadditivity:

$$100 = \delta^{1^{\sigma}} Z'^{T_{0,12}} \tag{A.7}$$

$$\delta = \left(\frac{100}{Z'^{T_{0,12}}} \right) \tag{A.8}$$

The exponent in (A.8) is the same as in (A.4), so the delayed payment is the same under subadditivity as additivity for this particular horizon, and the δ estimate is also the same (for any horizon longer than 12 months, subadditivity would lead to a higher estimate of δ than additivity). It follows immediately that the estimated δ is smaller for today to 6 months than for today to 12 months under subadditivity, illustrating that δ increases with time horizon length.

F2 Estimates of parameters of the quasi-hyperbolic model using OH designs

Consider an OH design contrasting behavior over a horizon from today to 6 months versus a horizon from today to 12 months. The observed delayed payments needed to induce waiting for each of the time horizons give two indifference relations:

$$100 = \beta \delta^{\frac{1}{2}} Z^{T_{0,6}} \tag{A.9}$$

$$100 = \beta \delta Z^{T_{0,12}} \tag{A.10}$$

Solving for the preference parameters yields:

$$\beta = 100 \frac{Z^{T_{0,12}}}{(Z^{T_{0,6}})^2} \tag{A.11}$$

$$\delta = \left(\frac{Z^{T_{0,6}}}{Z^{T_{0,12}}} \right)^2 \tag{A.12}$$

Now suppose the individual has the same constant discount factor, but is subadditive rather than present-biased, using again the functional form proposed by Read (2001). This implies that behavior is derived from the following indifference relations:

$$100 = \delta^{\frac{1}{2}^{\sigma}} Z'^{T_{0,6}} \tag{A.13}$$

$$100 = \delta Z^{T_{0,12}} \tag{A.14}$$

Equating the right hand sides implies a relationship between the discount factor and the ratio of the delayed payments needed for indifference for each of the horizons.

$$\delta^{\frac{1}{2}^{\sigma}} Z'^{T_{0,6}} = \delta Z^{T_{0,12}} \tag{A.15}$$

$$\delta = \left(\frac{Z'^{T_{0,6}}}{Z^{T_{0,12}}} \right)^{\frac{1}{1-\frac{1}{2}^{\sigma}}} \tag{A.16}$$

The exponent on the right hand side of (A.16) is greater than the exponent in (A.12), given $0 < s < 1$. Thus, for a given δ, we know that:

$$\frac{Z'^{T_{0,6}}}{Z^{T_{0,12}}} > \frac{Z^{T_{0,6}}}{Z^{T_{0,12}}} \tag{A.17}$$

This implies that if an individual is subadditive, a higher value of δ will be calculated if one applies the quasi-hyperbolic model. Also, we know that $Z'^{T_{0,6}} > Z^{T_{0,6}}$, Thus, subadditivity fosters finding lower values of β if the quasi-hyperbolic model is used, and can lead to finding $\beta < 1$ even if the individual is not present biased. Using different time horizons will also influence the calculated δ and β through the channel of subadditivity, e.g., using today to 24 months instead of today to 12 months will lead to inferring higher δ and lower β.

F3 Estimates of parameters of the quasi-hyperbolic model using SH designs

Consider an SH design contrasting behavior over a horizon from today to 6 months versus a horizon from 6 months to 12 months. The observed delayed payments needed to induce waiting for each of the time horizons give two indifference relations:

$$100 = \beta \delta^{\frac{1}{2}} Z^{T_{0,6}} \tag{A.18}$$

$$100 = \delta^{\frac{1}{2}} Z^{T_{6,12}} \tag{A.19}$$

Solving yields:

$$\beta = \frac{Z^{T_{6,12}}}{Z^{T_{0,6}}} \tag{A.20}$$

$$\delta = \left(\frac{100}{Z^{T_{6,12}}} \right)^{2} \tag{A.21}$$

Note that due to additivity of the DU model, the length of the time horizon used for the SH comparison should not matter for the calculated β or δ.

If the individual is instead subadditive with constant discounting, then we have:

$$100 = \delta^{\frac{1}{2}^{\sigma}} Z'^{T_{0,6}} \tag{A.22}$$

$$100 = \delta^{\frac{1}{2}^{\sigma}} Z'^{T_{6,12}} \tag{A.23}$$

It is immediately obvious that $Z'^{T_{0,6}} = Z'^{T_{6,12}}$. Thus, if the individual is a constant discounter who is subadditive, the beta calculated assuming the quasi-hyperbolic model will be equal to 1, implying no present bias. This is true regardless of the time horizon lengths chosen. The δ calculated will be smaller due to subadditivity. Furthermore, since the calculated δ is identified from a single time horizon in SH designs (the one shifted into the future), our results from above for single time horizon designs imply that the calculated δ will increase with the time horizon length used for the SH comparison.

G Experiment Instructions

In the following we present a translation of the German instructions. Instructions were presented to the interviewer on the screen of the laptop computer, and were read aloud to the subjects by the interviewer.

Screen 1

Now that the interview is over we invite you to participate in a behavioral experiment, which is important for economic science. The experiment involves financial decisions, which you can make in any way you want to. The questions are similar to those asked in the questionnaire with the exception that THIS TIME YOU CAN EARN REAL MONEY!

I will first explain the decision problem to you. Then you will make your decisions. A chance move will then determine whether you actually earn money.

Every 7th participant wins!

HOW MUCH MONEY YOU WILL EARN AND AT WHICH POINT IN TIME WILL DEPEND ON YOUR DECISIONS IN THE EXPERIMENT.

If you are among the winners, your amount will be paid by check. In this case the check will be sent to you by post.

Screen 2

Participants were then shown a choice table for the respective experiment as an example. The table was printed on a green piece of paper and was handed to participants for them to study.

The experimenter continued explaining how the experiment would work.

The interviewer gave the following explanation:

In each row you see two alternatives. You can choose between

- A fixed amount of 100 Euro (column A "today")
- and a somewhat higher amount, which will be paid to you only "in 12 months" (column B).

Payment "today" means that the check you get by post can be cashed immediately.
Payment "in 12 months" means that the check you get can be cashed only in 12 months.

You start with row 1 and then you go down from row to row. In each row you decide between 100 Euro today (column A) and a higher amount (column B); please always keep the timing of the payments in mind. The amount on the left side always remains the same, only the amount on the right side increases from row to row.

Which row on one of the tables will be relevant for your earnings will be determined by a random device later.

Screen 3

As you can see, you can earn a considerable amount of money. Therefore, please carefully consider your decisions.

Can we start now?

If the participant agreed, the experiment started. If not, the experimenter said the following:

The experiment is the part of the interview where you can earn money! Are you sure that you DO NOT WANT TO PARTICIPATE?

If the participant still did not want to participate, the experiment was not conducted and the participant answered a few final questions. In case the subject wanted to participate the experiment began.

Participants studied their table. The experimenter asked for the subject's decision in each row, whether they preferred the option in Column A or B, starting with the first row. In case a participant preferred the higher, delayed amount the experimenter asked:

You have decided in favor of the higher amount of X in X months. Can we assume that this implies that for all higher amounts you also prefer the later payment, meaning that for all remaining rows all higher amounts will be selected (i.e., Column B).

If the participant did not agree, he kept on deciding between columns A and B.

Once the first table was completed, the second table was presented to the participant. The experimenter then said:

Now there is a second table. Please look at the table. You will do the same as before but please note that the dates of payment and also the payments on the right side of the table have changed.

For the second and third tables, the same procedure as with the first table was followed.

When the tables were completed, participants answered some additional questions that included:

"When you made your decisions in the experiment, did you think about current interest rates or returns?" *Respondents could answer "yes," "no," or "no reply."*

"In case you were to receive the 100 Euro payment today from the experiment, what would you do with the money in the next week?" *Respondents could choose "spend everything," "spend most," "spend some but save most," "save all," or "no reply."*

Then it was determined whether the participant was among those who would be paid.

Participants could choose their "lucky number" between 1 and 7. They could then press on one out of seven fields on the computer, which represented numbers from 1 to 7. If they hit "their" number they won, otherwise they did not win. In case they won, it was determined which of the tables was selected and which row of the respective table. This was done again by pressing on fields presented to participants on the computer screen. In the end subjects who had won were informed that they would be sent the check by mail.

H Choices for time horizons in the CSS and SOEP datasets

Table A.1: Choices in the $T_{0,12}$ horizon of the CSS data

	Column A		Column B
1)	€100 today	or	€102.5 in 12 months
2)	€100 today	or	€105.1 in 12 months
3)	€100 today	or	€107.6 in 12 months
4)	€100 today	or	€110.3 in 12 months
5)	€100 today	or	€112.9 in 12 months
6)	€100 today	or	€115.6 in 12 months
7)	€100 today	or	€118.3 in 12 months
8)	€100 today	or	€121.0 in 12 months
9)	€100 today	or	€123.8 in 12 months
10)	€100 today	or	€126.6 in 12 months
11)	€100 today	or	€129.4 in 12 months
12)	€100 today	or	€132.3 in 12 months
13)	€100 today	or	€135.1 in 12 months
14)	€100 today	or	€138.1 in 12 months
15)	€100 today	or	€141.0 in 12 months
16)	€100 today	or	€144.0 in 12 months
17)	€100 today	or	€147.0 in 12 months
18)	€100 today	or	€150.1 in 12 months
19)	€100 today	or	€153.1 in 12 months
20)	€100 today	or	€156.3 in 12 months

Table A.2: Choices in the $T_{0,6}$ horizon of the CSS data

	Column A		Column B
1)	€100 today	or	€101.3 in 6 months
2)	€100 today	or	€102.5 in 6 months
3)	€100 today	or	€103.8 in 6 months
4)	€100 today	or	€105.0 in 6 months
5)	€100 today	or	€106.3 in 6 months
6)	€100 today	or	€107.5 in 6 months
7)	€100 today	or	€108.8 in 6 months
8)	€100 today	or	€110.0 in 6 months
9)	€100 today	or	€111.3 in 6 months
10)	€100 today	or	€112.5 in 6 months
11)	€100 today	or	€113.8 in 6 months
12)	€100 today	or	€115.0 in 6 months
13)	€100 today	or	€116.3 in 6 months
14)	€100 today	or	€117.5 in 6 months
15)	€100 today	or	€118.8 in 6 months
16)	€100 today	or	€120.0 in 6 months
17)	€100 today	or	€121.3 in 6 months
18)	€100 today	or	€122.5 in 6 months
19)	€100 today	or	€123.1 in 6 months
20)	€100 today	or	€125.0 in 6 months

Table A.3: Choices in the $T_{6,12}$ horizon of the CSS data

	Column A		Column B
1)	€100 in 6 months	or	€101.3 in 12 months
2)	€100 in 6 months	or	€102.5 in 12 months
3)	€100 in 6 months	or	€103.8 in 12 months
4)	€100 in 6 months	or	€105.0 in 6 months
5)	€100 in 6 months	or	€106.3 in 12 months
6)	€100 in 6 months	or	€107.5 in 12 months
7)	€100 in 6 months	or	€108.8 in 12 months
8)	€100 in 6 months	or	€110.0 in 12 months
9)	€100 in 6 months	or	€111.3 in 12 months
10)	€100 in 6 months	or	€112.5 in 12 months
11)	€100 in 6 months	or	€113.8 in 12 months
12)	€100 in 6 months	or	€115.0 in 12 months
13)	€100 in 6 months	or	€116.3 in 12 months
14)	€100 in 6 months	or	€117.5 in 12 months
15)	€100 in 6 months	or	€118.8 in 12 months
16)	€100 in 6 months	or	€120.0 in 12 months
17)	€100 in 6 months	or	€121.3 in 12 months
18)	€100 in 6 months	or	€122.5 in 12 months
19)	€100 in 6 months	or	€123.1 in 12 months
20)	€100 in 6 months	or	€125.0 in 12 months

Table A.4: Choices in the $T_{0,12}$ horizon of the SOEP data

	Column A		Column B
1)	€200 today	or	€205.0 in 12 months
2)	€200 today	or	€210.1 in 12 months
3)	€200 today	or	€215.2 in 12 months
4)	€200 today	or	€220.4 in 12 months
5)	€200 today	or	€225.6 in 12 months
6)	€200 today	or	€230.9 in 12 months
7)	€200 today	or	€236.3 in 12 months
8)	€200 today	or	€241.7 in 12 months
9)	€200 today	or	€247.2 in 12 months
10)	€200 today	or	€252.8 in 12 months
11)	€200 today	or	€258.4 in 12 months
12)	€200 today	or	€264.1 in 12 months
13)	€200 today	or	€269.8 in 12 months
14)	€200 today	or	€275.6 in 12 months
15)	€200 today	or	€281.5 in 12 months
16)	€200 today	or	€287.4 in 12 months
17)	€200 today	or	€293.4 in 12 months
18)	€200 today	or	€299.4 in 12 months
19)	€200 today	or	€305.6 in 12 months
20)	€200 today	or	€311.7 in 12 months

Table A.5: Choices in the $T_{0,6}$ horizon of the SOEP data

	Column A		Column B
1)	€200 today	or	€202.5 in 6 months
2)	€200 today	or	€205.0 in 6 months
3)	€200 today	or	€207.5 in 6 months
4)	€200 today	or	€209.9 in 6 months
5)	€200 today	or	€212.4 in 6 months
6)	€200 today	or	€214.9 in 6 months
7)	€200 today	or	€217.4 in 6 months
8)	€200 today	or	€219.9 in 6 months
9)	€200 today	or	€222.4 in 6 months
10)	€200 today	or	€224.8 in 6 months
11)	€200 today	or	€227.3 in 6 months
12)	€200 today	or	€229.8 in 6 months
13)	€200 today	or	€232.3 in 6 months
14)	€200 today	or	€234.8 in 6 months
15)	€200 today	or	€237.3 in 6 months
16)	€200 today	or	€239.8 in 6 months
17)	€200 today	or	€242.2 in 6 months
18)	€200 today	or	€244.7 in 6 months
19)	€200 today	or	€247.2 in 6 months
20)	€200 today	or	€249.7 in 6 months

Table A.6: Choices in the $T_{0,1}$ horizon of the SOEP data

	Column A		Column B
1)	€200 today	or	€200.4 in 1 month
2)	€200 today	or	€200.8 in 1 month
3)	€200 today	or	€201.2 in 1 month
4)	€200 today	or	€201.6 in 1 month
5)	€200 today	or	€202.0 in 1 month
6)	€200 today	or	€202.4 in 1 month
7)	€200 today	or	€202.8 in 1 month
8)	€200 today	or	€203.2 in 1 month
9)	€200 today	or	€203.6 in 1 month
10)	€200 today	or	€203.9 in 1 month
11)	€200 today	or	€204.3 in 1 month
12)	€200 today	or	€204.7 in 1 month
13)	€200 today	or	€205.1 in 1 month
14)	€200 today	or	€205.4 in 1 month
15)	€200 today	or	€205.8 in 1 month
16)	€200 today	or	€206.1 in 1 month
17)	€200 today	or	€206.5 in 1 month
18)	€200 today	or	€206.8 in 1 month
19)	€200 today	or	€207.2 in 1 month
20)	€200 today	or	€207.5 in 1 month

Table A.7: Choices in the $T_{0,1b}$ horizon of the SOEP data

	Column A		Column B
1)	€200 today	or	€200.8 in 1 month
1)	€200 today	or	€201.6 in 1 month
2)	€200 today	or	€202.4 in 1 month
3)	€200 today	or	€203.2 in 1 month
4)	€200 today	or	€203.9 in 1 month
5)	€200 today	or	€204.7 in 1 month
6)	€200 today	or	€205.4 in 1 month
7)	€200 today	or	€206.1 in 1 month
8)	€200 today	or	€206.8 in 1 month
9)	€200 today	or	€207.5 in 1 month
10)	€200 today	or	€208.2 in 1 month
11)	€200 today	or	€208.9 in 1 month
12)	€200 today	or	€209.6 in 1 month
13)	€200 today	or	€210.2 in 1 month
14)	€200 today	or	€210.8 in 1 month
15)	€200 today	or	€211.5 in 1 month
16)	€200 today	or	€212.1 in 1 month
17)	€200 today	or	€212.7 in 1 month
18)	€200 today	or	€213.3 in 1 month
19)	€200 today	or	€213.9 in 1 month

Table A.8: Choices in the $T_{12,13}$ horizon of the SOEP data

	Column A		Column B
1)	€200 in 12 months	or	€200.8 in 13 months
2)	€200 in 12 months	or	€201.6 in 13 months
3)	€200 in 12 months	or	€202.4 in 13 months
4)	€200 in 12 months	or	€203.2 in 13 months
5)	€200 in 12 months	or	€203.9 in 13 months
6)	€200 in 12 months	or	€204.7 in 13 months
7)	€200 in 12 months	or	€205.4 in 13 months
8)	€200 in 12 months	or	€206.1 in 13 months
9)	€200 in 12 months	or	€206.8 in 13 months
10)	€200 in 12 months	or	€207.5 in 13 months
11)	€200 in 12 months	or	€208.2 in 13 months
12)	€200 in 12 months	or	€208.9 in 13 months
13)	€200 in 12 months	or	€209.6 in 13 months
14)	€200 in 12 months	or	€210.2 in 13 months
15)	€200 in 12 months	or	€210.8 in 13 months
16)	€200 in 12 months	or	€211.5 in 13 months
17)	€200 in 12 months	or	€212.1 in 13 months
18)	€200 in 12 months	or	€212.7 in 13 months
19)	€200 in 12 months	or	€213.3 in 13 months
20)	€200 in 12 months	or	€213.9 in 13 months

References

Abdellaoui, M., A. E. Attema, and H. Bleichrodt (2010): "Intertemporal tradeoffs for gains and losses: An experimental measurement of discounted utility," *The Economic Journal*, 120(545), 845–866.

Abdellaoui, M., H. Bleichrodt, et al. (2013): "Sign-dependence in intertemporal choice," *Journal of Risk and Uncertainty*, 47(3), 225–253.

Abdellaoui, M., H. Bleichrodt, O. l'Haridon, and C. Paraschiv (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.

Abebe, G., A. S. Caria, and E. Ortiz-Ospina (2021): "The selection of talent: Experimental and structural evidence from ethiopia," *American Economic Review*, 111(6), 1757–1806.

Alan, S., and S. Ertac (2015): "Patience, self-control and the demand for commitment: Evidence from a large-scale field experiment," *Journal of Economic Behavior & Organization*, 115, 111–122.

——— (2018): "Fostering patience in the classroom: Results from randomized educational intervention," *Journal of Political Economy*, 126(5), 1865–1911.

Amasino, D. R., N. J. Sullivan, R. E. Kranton, and S. A. Huettel (2019): "Amount and time exert independent influences on intertemporal choice," *Nature human behaviour*, 3(4), 383–392.

Andersen, S., G. Harrison, M. Lau, and E. Rutstroem (2008): "Eliciting risk and time preferences," *Econometrica*, 76(3), 583–618.

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2006): "Elicitation using multiple price list formats," *Experimental Economics*, 9(4), 383–405.

——— (2010): "Preference heterogeneity in experiments: Comparing the field and laboratory," *Journal of Economic Behavior & Organization*, 73(2), 209–224.

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2013): "Discounting behaviour and the magnitude effect: Evidence from a field experiment in Denmark," *Economica*, 80(320), 670–697.

——— (2014): "Discounting behavior: A reconsideration," *European Economic Review*, 71, 15–33.

——— (2018): "Multiattribute utility theory, intertemporal utility, and correlation aversion," *International Economic Review*, 59(2), 537–555.

Andreoni, J., M. Callen, M. Y. Khan, K. Jaffar, and C. Sprenger (2016): "Using preference estimates to customize incentives: an application to polio vaccination drives in Pakistan," Discussion paper, National Bureau of Economic Research.

Andreoni, J., M. A. Kuhn, and C. Sprenger (2015): "Measuring time preferences: A comparison of experimental methods," *Journal of Economic Behavior & Organization*, 116, 451–464.

ANDREONI, J., AND C. SPRENGER (2012a): “Estimating Time Preferences from Convex Budgets,” *American Economic Review*, 102(7), 3333–3356.

——— (2012b): “Risk Preferences Are Not Time Preferences,” *American Economic Review*, 102(7), 3357–3376.

ATTEMA, A. E., H. BLEICHRODT, Y. GAO, Z. HUANG, AND P. P. WAKKER (2016): “Measuring discounting without measuring utility,” *American Economic Review*, 106(6), 1476–94.

AUGENBLICK, N., M. NIEDERLE, AND C. SPRENGER (2015): “Working over time: Dynamic inconsistency in real effort tasks,” *Quarterly Journal of Economics*, pp. 1067–1115.

AUGENBLICK, N., AND M. RABIN (2019): “An Experiment on Time Preference and Misprediction in Unpleasant Tasks.,” *Review of Economic Studies*, 86(3).

AYCINENA, D., S. BLAZSEK, L. RENTSCHLER, AND C. SPRENGER (2022): “Intertemporal choice experiments and large-stakes behavior,” *Journal of Economic Behavior & Organization*, 196, 484–500.

AYCINENA, D., C. V. S. DE ECONOMIA EXPERIMENTAL, G. GUATEMALA, S. BLAZSEK, L. RENTSCHLER, AND B. SANDOVAL (2015): “Risk, discounting and demand for intrahousehold control for recipients of conditional cash transfers,” .

AYCINENA, D., AND L. RENTSCHLER (2018): “Discounting and digit ratio: Low 2D: 4D predicts patience for a sample of females,” *Frontiers in behavioral neuroscience*, 11, 257.

BALAKRISHNAN, U., J. HAUSHOFER, AND P. JAKIELA (2020): “How soon is now? Evidence of present bias from convex time budget experiments,” *Experimental Economics*, 23(2), 294–321.

BANERJI, A., J. GOTO, H. ISHIZAKI, T. KUROSAKI, K. LAL, S. PAUL, S. YASUYUKI, S. TSUDA, ET AL. (2018): “Entrepreneurship in micro and Small enterprises: Empirical findings from resurveys in northeastern areas of Delhi, India,” Discussion paper.

BAUER, M., AND J. CHYTILOVÁ (2010): “The impact of education on subjective discount rate in Ugandan villages,” *Economic development and cultural change*, 58(4), 643–669.

——— (2013): “Women, children and patience: Experimental evidence from i ndian villages,” *Review of Development Economics*, 17(4), 662–675.

BENHABIB, J., A. BISIN, AND A. SCHOTTER (2010): “Present-bias, quasi-hyperbolic discounting, and fixed costs,” *Games and Economic Behavior*, 69(2), 205–223.

BOOIJ, A. S., AND B. M. VAN PRAAG (2009): “A simultaneous approach to the estimation of risk aversion and the subjective time discount rate,” *Journal of Economic Behavior & Organization*, 70(1-2), 374–388.

BRADFORD, W. D., C. COURTEMANCHE, G. HEUTEL, P. MCALVANAH, AND C. RUHM (2017): “Time preferences and consumer behavior,” *Journal of Risk and Uncertainty*, 55(2), 119–145.

Bradford, W. D., P. Dolan, and M. M. Galizzi (2019): "Looking ahead: Subjective time perception and individual discounting," *Journal of Risk and Uncertainty*, 58(1), 43–69.

Brocas, I., J. D. Carrillo, and J. Tarrasó (2018): "How long is a minute?," *Games and Economic Behavior*, 111, 305–322.

Brown, A. L., Z. E. Chua, and C. F. Camerer (2009): "Learning and visceral temptation in dynamic saving experiments," *The Quarterly Journal of Economics*, 124(1), 197–231.

Burks, S., J. Carpenter, L. Goette, and A. Rustichini (2012): "Which measures of time preference best predict outcomes: Evidence from a large-scale field experiment," *Journal of Economic Behavior and Organization*, 84(1), 308–320.

Carlsson, F., H. He, P. Martinsson, P. Qin, and M. Sutter (2012): "Household decision making in rural China: Using experiments to estimate the influences of spouses," *Journal of Economic Behavior & Organization*, 84(2), 525–536.

Carvalho, L. S., S. Meier, and S. W. Wang (2016): "Poverty and economic decision-making: Evidence from changes in financial resources at payday," *American economic review*, 106(2), 260–84.

Cassar, A., A. Healy, and C. Von Kessler (2017): "Trust, risk, and time preferences after a natural disaster: experimental evidence from Thailand," *World Development*, 94, 90–105.

Chabris, C., D. Laibson, C. Morris, J. Schuldt, and D. Taubinsky (2008a): "Individual laboratory-measured discount rates predict field behavior," *Journal of Risk and Uncertainty*, 37(2), 237–269.

Chabris, C. F., D. Laibson, C. L. Morris, J. P. Schuldt, and D. Taubinsky (2008b): "Measuring intertemporal preferences using response times," Discussion paper, National Bureau of Economic Research.

Chabris, C. F., C. L. Morris, D. Taubinsky, D. Laibson, and J. P. Schuldt (2009): "The allocation of time in decision-making," *Journal of the European Economic Association*, 7(2-3), 628–637.

Cohen, J., K. M. Ericson, D. Laibson, and J. M. White (2020): "Measuring time preferences," *Journal of Economic Literature*, 58(2), 299–347.

Coller, M., G. W. Harrison, and E. E. Rutström (2003): "Are discount rates constant? Reconciling theory and observation," Discussion paper, Citeseer.

——— (2012): "Latent process heterogeneity in discounting behavior," *Oxford economic papers*, 64(2), 375–391.

Cubitt, R., R. McDonald, and D. Read (2018): "Time matters less when outcomes differ: Unimodal vs. cross-modal comparisons in intertemporal choice," *Management Science*, 64(2), 873–887.

Deck, C., and S. Jahedi (2015a): "An experimental investigation of time discounting in strategic settings," *Journal of Behavioral and Experimental Economics*, 54, 95–104.

——— (2015b): "Time discounting in strategic contests," *Journal of Economics & Management Strategy*, 24(1), 151–164.

DeJarnette, P., D. Dillenberger, D. Gottlieb, and P. Ortoleva (2020): "Time lotteries and stochastic impatience," *Econometrica*, 88(2), 619–656.

Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2010): "Are Risk Aversion and Impatience Related to Cognitive Ability?," *American Economic Review*, 100(3), 1238–1260.

Duquette, E., N. Higgins, and J. Horowitz (2012): "Farmer discount rates: Experimental evidence," *American Journal of Agricultural Economics*, 94(2), 451–456.

Eckel, C., C. Johnson, and C. Montmarquette (2005): "Saving decisions of the working poor: Short-and long-term horizons," in *Field experiments in economics*. Emerald Group Publishing Limited.

Engle-Warnick, J., J. Héroux, and C. Montmarquette (2009): "Willingness to pay to reduce future risk," *CIRANO-Scientific Publications 2009s-37*.

Field, E., R. Pande, J. Papp, and N. Rigol (2013): "Does the classic microfinance model discourage entrepreneurship among the poor? Experimental evidence from India," *American Economic Review*, 103(6), 2196–2226.

Finke, M. S., and S. J. Huston (2013): "Time preference and the importance of saving for retirement," *Journal of Economic Behavior & Organization*, 89, 23–34.

Frederick, S., and G. Loewenstein (2008): "Conflicting motives in evaluations of sequences," *Journal of Risk and Uncertainty*, 37(2), 221–235.

Frederick, S., G. Loewenstein, and T. O'donoghue (2002): "Time discounting and time preference: A critical review," *Journal of Economic Literature*, 40(2), 351–401.

Freeman, D., P. Manzini, M. Mariotti, and L. Mittone (2016): "Procedures for eliciting time preferences," *Journal of Economic Behavior & Organization*, 126, 235–242.

Gine, X., J. Goldberg, D. Silverman, and D. Yang (2018): "Revising Commitments: Field Evidence on the Adjustment of Prior Choices," *Economic Journal*, 128(608), 159–188.

Halevy, Y. (2015): "Time consistency: Stationarity and time invariance," *Econometrica*, 83(1), 335–352.

Harrison, G. W. (2005): "Field experiments and control," in *Field Experiments in Economics*. Emerald Group Publishing Limited.

Harrison, G. W., M. I. Lau, and M. B. Williams (2002): "Estimating individual discount rates in Denmark: A field experiment," *American economic review*, 92(5), 1606–1617.

Ifcher, J., and H. Zarghamee (2011): "Happiness and time preference: The effect of positive affect in a random-assignment experiment," *American Economic Review*, 101(7), 3109–29.

Imai, T., T. A. Rutter, and C. F. Camerer (2021): "Meta-analysis of present-bias estimation using convex time budgets," *The Economic Journal*, 131(636), 1788–1814.

Jackson, M. O., and L. Yariv (2014): "Present bias and collective dynamic choice in the lab," *American Economic Review*, 104(12), 4184–4204.

Janssens, W., B. Kramer, and L. Swart (2017): "Be patient when measuring hyperbolic discounting: Stationarity, time consistency and time invariance in a field experiment," *Journal of Development Economics*, 126, 77–90.

Khwaja, A., D. Silverman, and F. Sloan (2007): "Time preference, time discounting, and smoking decisions," *Journal of health economics*, 26(5), 927–949.

Laury, S. K., M. M. McInnes, and J. Todd Swarthout (2012): "Avoiding the curves: Direct elicitation of time preferences," *Journal of Risk and Uncertainty*, 44(3), 181–217.

Luhrmann, M., M. Serra-Garcia, and J. Winter (2018): "The impact of financial education on adolescents' intertemporal choices," *American Economic Journal: Economic Policy*, 10(3), 309–32.

Manzini, P., M. Mariotti, and L. Mittone (2010): "Choosing monetary sequences: Theory and experimental evidence," *Theory and Decision*, 69(3), 327–354.

Matousek, J., T. Havranek, and Z. Irsova (2022): "Individual discount rates: a meta-analysis of experimental evidence," *Experimental Economics*, 25(1), 318–358.

McAlvanah, P. (2010): "Subadditivity, patience, and utility: The effects of dividing time intervals," *Journal of Economic Behavior & Organization*, 76(2), 325–337.

McClure, S. M., D. I. Laibson, G. Loewenstein, and J. D. Cohen (2004): "Separate neural systems value immediate and delayed monetary rewards," *Science*, 306(5695), 503–507.

Meier, S., and C. D. Sprenger (2013): "Discounting financial literacy: Time preferences and participation in financial education programs," *Journal of Economic Behavior and Organization*, 95, 159–174.

——— (2015): "Temporal stability of time preferences," *Review of Economics and Statistics*, 97(2), 273–286.

Miao, B., and S. Zhong (2015): "Risk preferences are not time preferences: Separating risk and time preference: Comment," *American Economic Review*, 105(7), 2272–86.

Montiel Olea, J. L., and T. Strzalecki (2014): "Axiomatization and measurement of quasi-hyperbolic discounting," *The Quarterly Journal of Economics*, 129(3), 1449–1499.

Newell, R. G., and J. Siikamäki (2015): "Individual time preferences and energy efficiency," *American Economic Review*, 105(5), 196–200.

Olivola, C. Y., and S. W. Wang (2016): "Patience auctions: The impact of time vs. money bidding on elicited discount rates," *Experimental Economics*, 19(4), 864–885.

Read, D. (2001): "Is time-discounting hyperbolic or subadditive?," *Journal of Risk and Uncertainty*, 23(1), 5–32.

Rubinstein, A. (2003): ""Economics and Psychology?" The Case of Hyperbolic Discounting," *International Economic Review*, 44(4), 1207–1216.

Sawada, Y., Y. Kuroishi, Y. Sawad, and Y. Kuroishi (2015): "How does a natural disaster affect people's preference? The case of a large scale flood in the Philippines using the convex time budget experiments," *Disaster risks, social preferences, and policy effects: Field experiments in selected ASEAN and East Asian countries*, (34), 27–56.

Schreiber, P., and M. Weber (2016): "Time inconsistent preferences and the annuitization decision," *Journal of Economic Behavior & Organization*, 129, 37–55.

Slonim, R., J. Carlson, and E. Bettinger (2007): "Possession and discounting behavior," *Economics Letters*, 97(3), 215–221.

Sutter, M., M. G. Kocher, D. Glätzle-Rützler, and S. T. Trautmann (2013): "Impatience and uncertainty: Experimental decisions predict adolescents' field behavior," *American Economic Review*, 103(1), 510–531.

Takeuchi, K. (2011): "Non-parametric test of time consistency: Present bias and future bias," *Games and Economic Behavior*, 71(2), 456–478.

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

Vieider, F. M. (2021): "Noisy coding of time and reward discounting," *Working Paper, Ghent University.*