NBER WORKING PAPER SERIES

COMPLEXITY AND HYPERBOLIC DISCOUNTING

Benjamin Enke
Thomas Graeber
Ryan Oprea

Working Paper 31047
http://www.nber.org/papers/w31047

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2023, Revised January 2024

This paper supersedes Enke's and Graeber's earlier working paper "Noisy cognition and intertemporal choice." This research was supported by the National Science Foundation under Grant SES-1949366 and was approved by Harvard and UC Santa Barbara IRB. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Complexity and Hyperbolic Discounting
Benjamin Enke, Thomas Graeber, and Ryan Oprea
NBER Working Paper No. 31047
March 2023, Revised January 2024
JEL No. D03

ABSTRACT

A large literature shows that people discount financial rewards hyperbolically instead of exponentially. While discounting of money has been questioned as a measure of time preferences, it continues to be highly relevant in empirical practice and predicts a wide range of real-world behaviors, creating a need to understand what generates the hyperbolic pattern. We provide evidence that hyperbolic discounting reflects mistakes that are driven by the complexity of evaluating delayed payoffs. In particular, we document that hyperbolicity (i) is strongly associated with choice inconsistency and cognitive uncertainty,(ii) increases in overt complexity manipulations and (iii) arises nearly identically in computationally similar tasks that involve no actual payoff delays. Our results suggest that even if people had exponential discount functions, complexity-driven mistakes would cause them to make hyperbolic choices. We examine which experimental techniques to estimate present bias are (not) confounded by complexity.

Benjamin Enke
Department of Economics
Harvard University
Littauer M008
Cambridge, MA 02138
and CESifo
and also NBER
enke@fas.harvard.edu

Thomas Graeber
Harvard Business School
Harvard University
tgraeber@hbs.edu

Ryan Oprea
UC Santa Barbara
Santa Barbara, CA 93106
roprea@gmail.com

1 Introduction

Over the last four decades behavioral economists have gathered significant evidence that people do not discount future financial flows exponentially as suggested by standard economic models. Instead, researchers consistently document a *hyperbolic pattern*: extreme short-run impatience, decreasing rates of impatience as time intervals increase, and present bias in structural estimates (Cohen et al., 2020). Initially the literature interpreted these results in the money-early-or-later-paradigm as direct evidence for non-exponential time *preferences*, but in recent years this interpretation has been questioned: because of the fungibility of money, impatience over monetary rewards need not directly measure true time preferences.

Why does the hyperbolic pattern emerge in the valuation of financial flows, and how should we interpret these deviations from standard exponential predictions? Answering this question is of first-order importance for two primary reasons. First, many of the most important intertemporal choices in modern economies are made over money rather than direct consumption. From a positive perspective, understanding anomalies like hyperbolicity is therefore crucial, regardless of their normative connection to underlying preferences. Second, as we discuss in detail in Section 2, despite the theoretical ambiguities associated with the money-early-or-later paradigm, it continues to be the dominant paradigm in empirical practice. Appendix Table A.1 illustrates the breadth of applications in an overview of recent publications in the profession's top general interest and field journals that rely on this paradigm. This continued popularity likely reflects that money-based measures are highly predictive of a wide range of real-world economic behaviors and outcomes including savings, educational investment, exercise, misconduct in school, tax filing, food choice and wealth inequality (see references below), suggesting that they capture an ecologically important component of time discounting.[1]

In this paper we provide evidence that hyperbolic discounting reflects systematic mistakes made in response to the complexity of evaluating intertemporal tradeoffs. Our point of departure is the observation that any intertemporal decision problem requires decision makers to mentally aggregate a number of constituent components of the problem (delays, rewards and one's discount factor) into a decision. This process of valuation and aggregation may be difficult for decision makers. For example, introspection suggests that determining one's present value for "$80 in two years" is non-trivial. If precisely valuing delayed payoffs is indeed complex in this sense,[2] decision-makers may approximate instead: they may calculate values for delayed payments noisily or heuristically rather than precisely and properly.

Importantly, the mistakes that result from noisy or heuristic approximations need not can-

[1]Another reason for the money paradigm's popularity is doubtless that, unlike consumption-based measures, it is logistically sufficiently simple to be fielded as part of large-scale data-collection exercises or field experiments.

[2]We use the word "complex" as a shorthand for "cognitively difficulty and / or costly". In particular, when we say that valuing a delayed payment is "complex", we mean that its discounted present value is not transparent to the decision maker because optimally aggregating its disaggregated components (delays and payments) into a value or decision is costly or difficult. If these costs and difficulties are sufficiently severe, the decision maker may be induced to use a less optimal procedure instead (Simon, 1955), producing mis-valuations.

cel out. Instead, they can produce systematic distortions that *look like* hyperbolic preferences. Intuitively, if the cognitive process of aggregating and trading off different problem components is difficult, observed decisions may become *insufficiently sensitive* ("attenuated") with respect to components of the choice problem: decision makers may not (know how to) fully integrate parameters such as the magnitude of the time delay into their decisions. This is relevant because – relative to exponential discounting – hyperbolic discounting can be interpreted descriptively as just such an insensitivity to the size of the payment delay. In a nutshell, then, our hypothesis is that the complexity of valuing delayed payments induces people to rely on noisy or heuristic processes, which generates an insensitivity of decisions to the delay, and, hence, hyperbolicity.

Complexity and hyperbolicity. We pursue a multi-pronged experimental approach that tests what we take to be the key implications of a complexity-based explanation. First, if hyperbolicity is indeed a pattern of complexity-driven mistakes, it should be correlated with independent evidence of imperfect decision making. In a first set of treatments we therefore gather data on two empirical proxies for imperfect or noisy decision-making: (i) we repeat some of the choices in order to measure inconsistencies in repetitions of the same problem (*choice inconsistency*) and (ii) we ask subjects to indicate how likely they think it is that they made suboptimal choices (*cognitive uncertainty*). We then examine whether these indices of error-prone choice predict hyperbolicity.

Second, if hyperbolicity is driven by complexity, we would expect it to intensify when we increase the computational difficulty of making intertemporal tradeoffs. In a second set of treatments, we make intertemporal problems more complex by deliberately describing them in a mathematically more complicated way and examine whether this increases the degree of hyperbolicity in subjects' choices.

Finally, the reverse should also be true: if hyperbolicity is driven by complexity, we would expect it to continue to appear even when scope for competing explanations (such as true time preferences or self-control problems) is experimentally removed from the decision problem. Therefore, in a third set of treatments, we compare how people value delayed payments to the way they value computationally similar "atemporal mirrors" of the same tasks. In an atemporal mirror, subjects are asked to value dollar amounts that are paid immediately (removing scope for time preferences or self-control problems) but are "shrunk" some number of times at a fixed rate prior to payment (effectively, experimentally inducing an exponential reward function). Thus, in addition to asking subjects to value, e.g., \$50 paid in 12 months, we also ask them to value \$50 shrunk 12 times, each time by 4%. We then study whether this descriptively similar task produces evidence of hyperbolicity, even when the canonical explanations in the literature – such as intertemporal fungibility of money, non-standard time preferences, or payment risk – are by design ruled out.

We find four pieces of evidence that hyperbolic discounting over money, indeed, represents complexity-driven valuation mistakes. First, we show that, in standard intertemporal decisions, both cognitive uncertainty and choice inconsistencies are strongly correlated with hyperbolic discounting. Specifically, as we hypothesized, both empirical proxies for noisy decisions strongly predict an *insensitivity* of intertemporal decisions with respect to the delay. This insensitivity implies a striking "flipping" pattern: subjects who are more inconsistent or more cognitively uncertain are *less* patient over short horizons but *more* patient over long horizons, and therefore more hyperbolic. Even within the choices of a given subject, decisions exhibit more pronounced hyperbolicity precisely when the subject expresses high cognitive uncertainty. Overall, our data suggest that 85% of decreasing impatience in our setting is driven by valuation errors.

Second, we document that experimentally increasing the difficulty of processing and aggregating delays and payments leads to a joint increase in cognitive uncertainty, choice inconsistency and hyperbolicity, again confirming the link between complexity, mistakes and hyperbolicity.

Third, we find that hyperbolicity strongly arises in the valuation of atemporal mirrors, and to a similar degree as in true intertemporal choice. Again, the intuition is that subjects' valuations are *insufficiently sensitive* to the number of discounting steps. As a result, subjects discount atemporal payments more than the experimentally induced discount factor δ specifies when payments only need to be discounted one or two times, which directly mirrors the extreme short-run impatience that arises in intertemporal choice. However, in the same tasks, subjects discount atemporal payments that need to be discounted a large number of times considerably *less strongly* than they should, given the induced discount factor. Thus, even though we experimentally induce a fixed discount factor, subjects' revealed per-period "impatience" strongly decreases in the number of discounting steps, replicating the hyperbolicity widely observed in intertemporal choice experiments. Decomposing the fraction of hyperbolicity that appears in mirrors relative to true intertemporal days, we find that a strikingly similar fraction of diminishing impatience can be attributed to valuation mistakes as in our cognitive uncertainty and choice inconsistency exercises.

Finally, because in some of our treatments we measure each subject's behavior in both atemporal mirrors and intertemporal choice, we can show that behavior in the former *predicts* behavior in the latter. We find correlations across the two choice problems of 0.34, which means outright valuation mistakes in atemporal mirrors serves as one of the strongest predictors of intertemporal choice ever measured in the literature. This is evidence that behaviors in the two settings may be driven by a common behavioral mechanism, which can only be the difficulty of aggregating and trading off different problem components, the property the two decision tasks share.

We interpret these four pieces of distinct but highly consistent evidence as strongly suggesting that the hyperbolic pattern is mostly a consequence of the difficulty of properly aggregating problem components of intertemporal choice problems. Even if people had exponential discount functions, our results suggest that mistakes would produce behavior that looks hyperbolic due to the complexity of the valuation problem.

Direct evidence for the insensitivity mechanism. The extant literature (Read, 2001) has proposed a direct test for the idea that intertemporal decisions are insufficiently sensitive to variation in the time delay: so-called *subadditivity* designs. These involve a particularly stark documentation of insensitivity in which subjects violate transitivity by making less patient choices when a single time interval is decomposed into two separate intervals. If our hypothesis – that hyperbolicity results from insensitivity, driven by complexity-driven mistakes – was true then we should observe tight links also between subadditivity and the experimental measures discussed above. In our data, we consistently find that this is the case: (i) subadditivity effects are strongly correlated with cognitive uncertainty; (ii) they increase in the complexity manipulation; and (iii) they arise with equal strength in the atemporal mirrors. We conclude that most of hyperbolic discounting in our experiment is a consequence of insensitivity to variation in time intervals, which, in turn, is driven by the complexity of aggregating and trading off problem components.

Implications for estimating present bias. If hyperbolicity in the valuation of financial flows largely reflects the confounding effects of complexity, then this raises the question of whether – or which – estimates of present bias are potentially confounded. Here, it is important to distinguish between (i) structural estimates of present bias (as in the quasi-hyperbolic $\beta-\delta$ model) that can be identified from the hyperbolic shape of the empirical discount function; and (ii) treatment-based estimates that rely on experimental variation in front-end delays.

Our evidence suggests that structural estimates of present bias that are identified from the hyperbolic shape of the empirical discount function are severely inflated due to model misspecification (resulting from ignoring the effects of complexity). The intuition is that estimates of β are largely identified off of the insensitivity of the discount function, which we have shown is driven by complexity-derived mistakes. For instance, in our atemporal mirrors experiment, we structurally estimate $\hat{\beta} = 0.85$ – an estimate that would conventionally be interpreted as strong evidence of present bias, but here is unambiguously a measure of valuation mistakes. Similarly, in our intertemporal choice experiments over time-dated monetary rewards, we find that structural estimates of present bias are strongly correlated with cognitive uncertainty and choice inconsistencies, again suggesting that complexity spuriously inflates these estimates.

As is well-understood in the literature (Cohen et al., 2020), *structural* estimates of β only measure "true" present bias under strong assumptions. More directly, *causal* estimates of present bias can be gathered using front-end delay designs, in which subjects tend to violate stationarity by acting in more patient ways when both earlier and later payments are moved into the future by a common delay. In our experiments, we never find any indication linking these effects to valuation mistakes: (i) in intertemporal decisions, front-end delay effects are present and sizable, but they are uncorrelated with cognitive uncertainty and don't amplify in the complexity manipulation; and (ii) no front-end delay effects appear in the atemporal mirrors. This suggests that when estimating true present bias over money, researchers should use front-end delay designs rather than rely on the hyperbolic shape of the empirical discount function, which confounds complexity-derived error with present bias.

Summary and related literature. Taken together, over a number of distinct empirical approaches – including atemporal mirrors, choice inconsistencies, cognitive uncertainty and experimental complexity manipulations – we document the same story. The classical hyperbolic pattern in intertemporal decisions over money occurs because people respond to the difficulty of processing and aggregating delays and payments by using imperfect valuation rules that are *insufficiently sensitive* ("attenuated") to variation in time intervals. These results clarify a question that has puzzled researchers for a long time: what is it that the hyperbolic pattern in money-early-versus-later experiments captures, and how should we interpret results from such experiments? We interpret our results as suggesting that (i) hyperbolicity is a consequence of the way people respond to complexity and (ii) the true signal about underlying time preferences measured using money-early-or-later decisions is small. However, we highlight that this failure to identify time preferences is not due to the *temporal* factors commonly discussed in the literature (such as fungibility of money or payments risks) but instead to generic responses to complexity that arise even when these factors are absent (e.g., in atemporal mirrors). In this regard, an interesting question is why money-based experimental measures are often predictive of relevant field behaviors. One possibility is that the noisy and heuristic procedures (complexity responses) that people deploy in money experiments are similar to (and thus predictive of) the procedures they use in field contexts. Another possibility is that there is some signal about true time preferences in money experiments, and that this drives the linkage to field behaviors.

Our paper most directly connects to a long literature documenting hyperbolicity in intertemporal choices over money, beginning with Thaler (1981) and summarized in Cohen et al. (2020). Measures of discounting using money continue to be very popular in practice, and are widely used as both response and explanatory variables in empirical work. Our paper clarifies how these experimental measures and the hyperbolicity they produce should be interpreted. Less directly connected is a more recent literature that attempts to directly measure true time preferences by studying people's intertemporal preferences over consumption rather than monetary payments (e.g., Augenblick et al., 2015; Augenblick, 2018; Augenblick and Rabin, 2019). Our paper differs in that our aim is not to measure time preferences but rather to understand monetary discounting behavior, which we show is shaped to a great extent by factors other than time preferences. Moreover, while the consumption literature is largely focused on measuring present bias, our focus is instead on understanding hyperbolicity – a pattern that is difficult to measure using the short-horizon designs used in most of the consumption literature.

Our paper connects to experimental work that documents various "cognitive effects" in intertemporal choice, such as that time discounting is sensitive to cognitive load, time pressure and framing (e.g., Ebert and Prelec, 2007; Imas et al., 2021; Dertwinkel-Kalt et al., 2021), and that noise or confusion can spuriously drive estimates of present bias or commitment demand (Chakraborty et al., 2017; Carrera et al., 2022). Related to us are also experimental papers showing that people have difficulty with exponential reasoning, suffering an exponential growth bias (Stango and Zinman, 2009; Goda et al., 2015). We too find errors in exponential reasoning, though unlike this literature on exponentially growing processes we study exponentially decaying ones and show that the associated errors are responsible for the classic hyperbolic pattern. Our paper also relates to theoretical work on how cognitive limitations such as cognitive noise

(Woodford, 2020) may affect intertemporal decision making (e.g., Gabaix and Laibson, 2022; Gershman and Bhui, 2020; Vieider, 2021b; Regenwetter et al., 2018). Our contribution to this body of research is to highlight the implications of complexity (and resulting noise) for understanding the hyperbolic pattern in the most-widely used experimental paradigm.

Our paper also links to recent experimental work on complexity and the non-standard behaviors it induces in other decision domains (e.g. Nielsen and Rehbeck, 2020; Ba et al., 2022; Augenblick et al., 2021; Oprea, 2020). For instance, the link between complexity and insensitivity that we show drives hyperbolicity is reminiscent of a seemingly-unrelated literature on choice under risk, where an emerging body of work has found that the generic difficulty of aggregating the constituent components of a risky choice problem generates an insensitivity to probabilities (e.g., Oprea, 2022; Enke and Graeber, 2023; Vieider, 2021a; Frydman and Jin, 2023; Enke and Shubatt, 2023; Khaw et al., 2021, 2022), producing systematic patterns like probability weighting. The common thread that runs through these lines of work is that complexity (and the noise it produces) cause insensitivity and, hence, famous behavioral anomalies.

Our paper is organized as follows. In Section 2 we review common anomalies in intertemporal choice and discuss our empirical strategies to explain them. Section 3 presents our experimental design. Sections 4 and 5 present the results. Section 6 discusses which classes of models can potentially rationalize our results and concludes.

2 Conceptual Background and Empirical Approaches

2.1 The Hyperbolic Pattern

Consider a simple intertemporal choice problem $D = (x_1, t_1; x_2, t_2)$ in which a decision maker must decide what dollar amount x_1 paid at t_1 (e.g., now) makes her indifferent to earning x_2 paid at $t_2 > t_1$ (e.g., in two months). Define $\Delta t \equiv t_2 - t_1$ (all time units are in months). The exponential discounted utility model constitutes the benchmark in the economics literature as the only model of intertemporal choice compatible with time-consistent behavior. In this model, the decision maker discounts rewards exponentially with an annual discount factor $\delta = 1 - \gamma$, where γ is approximately the constant discount rate. Throughout the paper we treat γ not as a preference parameter but rather as a descriptive empirical measurement of the per-period impatience that is implicit in choice. We then have:

$$(1-\gamma)^{t_1/12}x_1 = (1-\gamma)^{t_2/12}x_2 \qquad \Leftrightarrow \qquad \gamma = 1 - \left(\frac{x_1}{x_2}\right)^{12/\Delta t} = 1 - e^{-12 \cdot RRR/\Delta t}, \qquad (1)$$

where $RRR/\Delta t = ln(x_2/x_1)/\Delta t$ is the "interval-adjusted required rate of return" that the decision maker reveals through her choices. In the exponential model, the interval-adjusted RRR – and, hence, also γ – are constant, absent confounding factors. In what follows, we refer to the empirical measurement of γ as "implied annual impatience," which is an approximation of the average annual discount rate implied by a decision.

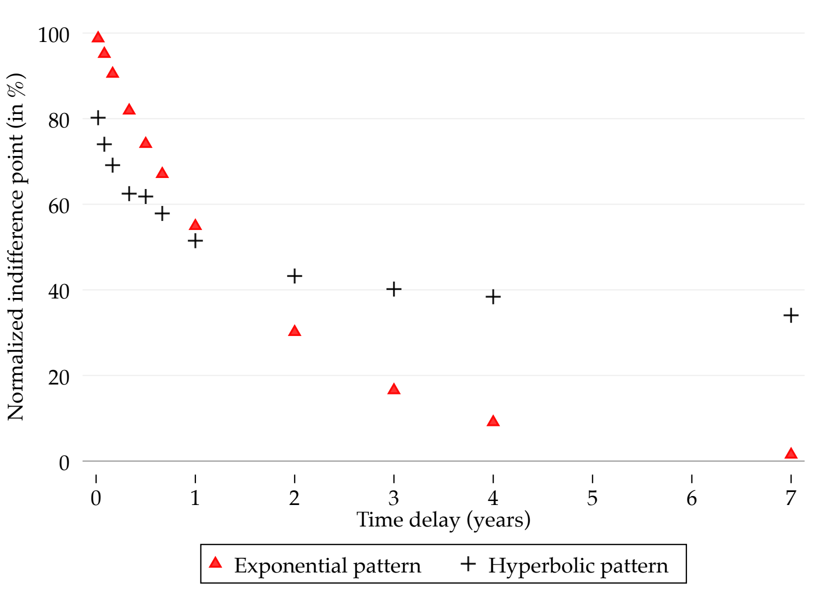

Figure 1: Qualitative illustration of classical hyperbolic discounting pattern. Figure shows the earlier payment that makes a decision-maker indifferent to a fixed later payment, expressed as a percentage of the later amount. Time delay describes the length of a delay, which can either start now or in the future.

Since Thaler (1981), behavioral economists have gathered significant evidence that decision makers behave in ways that are incompatible with the exponential discounting model when valuing financial flows (see Cohen et al. (2020) and Ericson and Laibson (2019) for reviews). Many of the core anomalies boil down to the observation that empirically observed discounting has a hyperbolic shape. Figure 1 summarizes the patterns. Throughout, we will highlight the observation that, within the set of strictly positive delays, observed discounting is insufficiently sensitive (attenuated) to variation in delays.

Extreme short-run impatience. Many studies show that decision makers tend to discount relatively short intervals at such an extremely steep rate that a constant discount rate would imply implausible and empirically counterfactual decisions over longer intervals. This tendency is sometimes confused with "present bias" in part because many early studies exclusively featured problems in which the sooner payoff date was immediate (i.e. $t_1 = 0$). More recent evidence (e.g., Kable and Glimcher, 2010) shows that, even when $t_1 > 0$, decision makers exhibit short-run impatience that is so pronounced that a constant discount factor would imply implausible and empirically counterfactual medium-run discounting behavior.

Decreasing impatience. Extreme short-run impatience is a component of a more general tendency for decision makers' revealed per-period impatience to decrease as the interval Δt becomes longer. Crucially, this implies that – relative to the best-fitting exponential model – decision-makers act "too impatiently" for short delays and "too patiently" for sufficiently long delays.

Sub-unitary estimates of β. It is common in the literature to summarize the hyperbolic pattern of discounting described above using structural estimates of the quasi-hyperbolic $\beta - \delta$ model (Laibson, 1997). This model postulates that decision makers put weight 1 on immediately paid

($t = 0$) rewards but weight $\beta\delta^t$ on delayed ($t > 0$) rewards. Structural estimates tend to find values of β significantly lower than 1 in most datasets, which is routinely interpreted as evidence of present bias. However, this evidence is indirect in the sense that it is typically identified from the hyperbolicity of discounting per se. By contrast, as discussed below, present bias can alternatively be estimated *causally* using front-end delay designs.

We will refer to this triplet of phenomena as the classical "hyperbolic pattern."

2.2 Interpretations of Hyperbolicity over Money

The traditional view and its criticisms. A number of explanations have been offered for these deviations from the exponential model. Traditionally, the dominant class of explanations have been *motivational* in nature: explanations rooted in preferences or internal conflicts that arise due to the fact that intertemporal choices involve the elapse of time. One category is *preference-based* explanations which argue that people simply do not have exponential, dynamically consistent time preferences, leading to non-exponential discounting behavior. This includes, for instance, hyperbolic and quasi-hyperbolic preference models (e.g., Loewenstein and Prelec, 1992; Laibson, 1997; O'Donoghue and Rabin, 1999) or temptation effects (Gul and Pesendorfer, 2001). Other authors have proposed that people have, in effect, "multiple selves" with divergent preferences at different dates, strategically vying for control (Fudenberg and Levine, 2006).

Recently, the literature has called many of these interpretations into question. Most importantly, the literature has raised concerns that because money is fungible, choices over delayed payment may not directly measure temporal motivations (e.g., time preferences). This has given rise to a literature attempting to measure time preferences and self-control problems using experiments in which subjects make decisions directly over real consumption (e.g., effort expenditures) rather than financial flows (e.g., Augenblick et al., 2015).

The continued popularity of the money-early-or-later paradigm. Despite the well-understood theoretical ambiguities of the money-based paradigm, it continues to be widely used in practice. Appendix Table A.1 provides an overview of recent top publications in economics that rely on the money-early-or-later paradigm. They include publications in most top general interest and top field journals, including the *AER, REStud, JPE, AEJ* and *JEEA*. The table also highlights the breadth of applications that researchers have pursued: (i) studies that use the money paradigm to measure "preferences" and to link them to ecological behaviors and outcomes (e.g. Epper et al., 2020; Sunde et al., 2022; Martinez et al., 2023); (ii) studies that use the paradigm to study the impact of randomized interventions or shocks on patience (e.g., Alan and Ertac, 2018); and (iii) studies that attempt to understand the psychological foundations of intertemporal decision making (e.g., Dertwinkel-Kalt et al., 2022; Fisher, 2021). Moreover, current working papers continue to rely on the paradigm (e.g., Abdellaoui et al., 2023; Brownback et al., 2023).

There are at least four reasons for this continued popularity. First, some researchers believe that money experiments actually do measure real time preferences because people narrowly bracket their choices and ignore the fungibility of money (e.g, Halevy, 2014; Andreoni et al., 2018; Balakrishnan et al., 2020). Second, as discussed in the Introduction, given that many decisions in modern economies involve monetary tradeoffs, measures of money discounting may contain important information about empirically-relevant behavioral tendencies, even if they don't identify deep time preferences. Third, the money-early-or-later paradigm is logistically much simpler than consumption-based experiments. This is of particular relevance given the rising prominence of lab-in-the-field experiments or large-scale data collection exercises that would be infeasible with consumption-based designs. A final reason is that despite the theoretical ambiguities, it is well-known that money discounting measures *do* predict many ecological outcomes and behaviors, underscoring their ecological relevance. This includes documented links with wealth inequality, income, savings, educational investment, exercise, misconduct in school, tax filing and food choice (e.g. Ashraf et al., 2006; Chabris et al., 2008; Meier and Sprenger, 2010; Mahajan and Tarozzi, 2012; Sutter et al., 2013; Falk et al., 2018; Epper et al., 2020; Sunde et al., 2022; Martinez et al., 2023; Brownback et al., 2023).

The relevance of the money-based discounting paradigm and the robustness of the hyperbolic pattern raise the question of what it is that intertemporal choice over monetary payments actually measures, and how hyperbolicity should be interpreted.

2.3 A Complexity-Based Account

We hypothesize that hyperbolicity arises because intertemporal decision making is inherently difficult (or complex) and prone to errors. In order to properly discount a delayed payment in the process of valuing it, a decision maker must (consciously or unconsciously) engage in an *aggregation procedure* that involves combining (i) one's discount factor; (ii) the time delay and (iii) the payments into a decision. Various steps of this aggregation process may be difficult and / or costly. In response, decision makers may fall prey to systematic mistakes.

In particular, recent discussions in the behavioral literature suggest that these mistakes may be systematic and have sufficient structure to produce the hyperbolic shape of the empirically-observed discounting function. Our hypothesis is that if the cognitive process of aggregating and trading off different problem components is difficult, observed decisions may become *insufficiently sensitive* ("attenuated") with respect to components such as the time delay, thus producing hyperbolicity. To take just one possible way this might occur, the difficulty of aggregating problem components may lead people to intuitively anchor their valuation in the middle of a "plausible range," and then imperfectly adjust up or down depending on what the precise delay in a given problem is. Such an anchoring-and-adjustment process is arguably easier than optimal choice and can generate the characteristic pattern in Figure 1: error-prone agents can simultaneously appear *more* impatient over short delays and *less* impatient over very long delays.

More generally, a small-but-growing theoretical intertemporal choice literature has shown that a number of error-prone evaluation rules are capable of producing decisions that are *insufficiently sensitive* to the time delay, hence producing hyperbolicity (Ebert and Prelec, 2007).

This literature posits specific cognitive processes, including models of random utility and random response (Regenwetter et al., 2018; Lu and Saito, 2018; He et al., 2019), imperfect time perception (e.g., Zauberman et al., 2009; Brocas et al., 2018), noisy mental simulations of future utils (Gabaix and Laibson, 2022; Gershman and Bhui, 2020), measurement error that could mechanically lead to a compressed / insensitive discounting functions if there are boundary effects (Gillen et al., 2019), or heuristic decision procedures (e.g., Rubinstein, 2003; Ericson et al., 2015; Read et al., 2013). Below in Section 6 we discuss which specific models are consistent with our results.

2.4 Providing Evidence for a Complexity Account

There are three main empirical implications of a complexity-based account of the hyperbolic pattern that are not shared by alternative explanations (i.e., time preferences, self-control problems). We use these three implications to propose three different tests of this hypothesis.

Connecting anomalies to signatures of error-prone decision processes. First, if complexity drives hyperbolicity, then hyperbolicity should be strongly correlated with independent evidence that subjects are not using optimal choice rules and are making aggregation mistakes. Thus, a first approach is to directly measure behavioral signatures of such errors in standard intertemporal choice problems, and to study whether this evidence predicts the severity of hyperbolic discounting.

The literature has proposed two empirical indicators that a decision was made using an imperfect (i.e., heuristic or noisy error-prone) choice rule: self-reported cognitive uncertainty and choice inconsistencies. First, we ask subjects how likely it is (in percentage terms) that their choice actually complies with their true tastes and preferences. This type of simple, unincentivized self-report about the optimality of choice has been shown to be highly predictive of an insensitivity of decisions with respect to parameters in other choice settings (Enke and Graeber, 2023; Arts et al., 2020). Second, we look for direct evidence of inconsistent, noisy decision-making, another indication of the use of an imperfectly rational decision procedure. Following the literature (e.g. Agranov and Ortoleva, 2017; Agranov et al., 2020; Khaw et al., 2021), we classify a decision as deriving from a noisy procedure if it is *different* from other choices made in repeated elicitations of an identical decision problem. Our empirical strategy is to study to what degree these behavioral signatures of error-prone decision making predict the incidence and severity of anomalies.

Manipulating complexity. Second, if hyperbolicity is a response to complexity, it should increase when the complexity of the choice problem exogenously increases. Thus, a second approach is to run experiments in which we deliberately describe payouts and delays in especially complicated ways, to make it more difficult (or more costly) to aggregate them into a decision. We then look for evidence that this increases the intensity of the hyperbolic pattern.

Reducing scope for competing explanations. Finally, if hyperbolicity is really driven by aggregation errors, it should continue to arise even after competing "motivational" explanations have been removed from the choice problem. Unlike the motivational explanations described above, complexity-based explanations do not rely in any special way on the actual elapse of time. They instead rely on the fact that intertemporal choice requires decision makers to aggregate multiple pieces of information, which requires intensive information processing. Because they do not depend on time, these sorts of explanations should produce hyperbolicity in decision problems that involve no actual temporal delay, but that require a similar type of reasoning.

Building on this observation, we can construct an immediately paid "atemporal mirror" M_D of choice problem D that replaces payment dates with a sequence of "steps" of payoff discounting. In each step of discounting, the payoff from the previous step is multiplied by an exogenously provided and known fixed factor $\delta < 1$. Thus, an atemporal mirror of D pays a deterministic amount $\delta^{t_1}x_1$ or $\delta^{t_2}x_2$ immediately. Instead of, e.g., valuing a payoff "$50 in two months," a decision maker evaluating a mirror is asked to value a payoff "$50 shrunk by δ two times." An atemporal mirror is therefore nothing more or less than an immediate dollar payment that has been deliberately described in such a way as to require a similar kind of information processing as is required in intertemporal choice.

Because atemporal mirrors involve no actual time delays, hyperbolicity in their evaluation must be driven by mistakes in aggregating the problem components. For instance, potential anomalies cannot be driven by non-exponential time preferences: an atemporal mirror *induces* exponential preferences. Thus, if hyperbolicity is present in such a setting, it means that we should expect people to exhibit hyperbolic behavior even if they had exponential preferences.

Our strategy is to first compare the way decision makers evaluate a set of intertemporal choice problems to the way they evaluate atemporal mirrors of those same problems. To whatever extent hyperbolicity arises in the evaluation of mirrors as in the evaluation of intertemporal payments, we have evidence that those anomalies can arise due to valuation errors alone. Second, we correlate behavior in the two types of problems across subjects. To whatever extent they are positively correlated we have evidence that hyperbolicity is driven in both cases by valuation errors, the only explanatory channel the two types of problems share.

An attractive feature of our multi-pronged research design is that it relies on three essentially orthogonal empirical approaches with different strengths and weaknesses: (i) directly measuring signatures of error-prone choice rules in standard intertemporal problems; (ii) experimentally manipulating complexity; and (iii) stripping away time delay from a standard intertemporal choice problem and inducing exponential preferences while retaining a similar degree of task difficulty.

Sessions	Description	Subjects
Delay & Mirror	18 tasks under *Mirror* & exactly repeated under *Delay* (order of treatments randomized)	500
Delay-M	12 delay tasks with elicitations of cognitive uncertainty and choice inconsistencies	645
Voucher-M	12 delay tasks with UberEats vouchers and elicitations of cognitive uncertainty and choice inconsistencies	500

Table 1: Overview of main experiments.

3 Experimental Design

3.1 Basic Setup

Table 1 provides an overview of the experimental design. Following standard methods used in the literature, the core tasks in our experiments are *multiple price lists* that ask subjects to evaluate a payment of x_2 at a time t_2 in terms of dollars paid at an earlier date $t_1 < t_2$. An example of the subject's decision screen is shown in Appendix Figure A.1. In each list, Option A is kept identical in every row, paying x_2 at date t_2. By contrast, Option B pays an amount x_1 that declines montonotically by \$2 in each row (ranging between x_2 and \$2), at date t_1. Non-negative discounting entails that subjects choose A in early rows of the list (or, with extreme preferences, never) and switch to B at some later row (we enforce single switching). The switching point between earlier and later payment yields a direct measure of the RRR and thereby implied annual impatience, γ.

We refer to these choice problems as the *Delay* treatment. In most cases we randomize (at the subject-list level) the delayed payment $x_2 \in \{\$40, \$42, \ldots \$52\}$. The experimental design includes three main types of price lists. First, "Now Lists," in which $t_1 = 0$ and t_2 varies across 1/4, 1, 2, 12, 24, 36, 48 and 84 months. Second, "Later Lists," which are identical to Now Lists except that the earlier payment is slightly delayed: $t_1 = 1$ or $t_1 = 1/4$ months. Finally, "Subadditivity/Front-End Delay (SA/FED) Lists", in which for some horizon T we assign subjects lists (t_1=0, $t_2 = T/2$), (t_1=T/2, $t_2 = T$) and (t_1=0, $t_2 = T$), maintaining a consistent x_2 across the three lists. We randomly assign T across subjects to be either 8 or 12. Dates t_1, t_2 represent months. Lists from each of these categories are included in every treatment, for every subject and randomly ordered at the subject level. Now and Later lists are used primarily to study the anomalies of extreme short-run discounting, decreasing impatience / hyperbolicity and sub-unitary β. SA/FED Lists are used to measure subadditivity and front-end delay effects.

3.2 Inducing Exponential Preferences and Removing Delay

Our first variation on this standard choice setting is to study companion problems in which we pay subjects an iteratively discounted version of the stated payoff immediately. In these tasks, discounting occurs though a known, exogenous discount factor, transforming A and B into "atemporal mirrors" of standard intertemporal choice tasks. This is framed to subjects as "shrinking" a payment t times. Each time a payment is "shrunk," it falls to $\delta < 1$ of its previous value, but a subject must reason through the consequence of this discounting in order to properly value it. The fact that atemporal mirrors are paid immediately is repeatedly emphasized to subjects in the instructions.

A choice list from treatment *Mirror* is displayed in Appendix Figure A.1. Each list asks subjects an exactly analogous sequence of binary choice questions as in the corresponding list from the *Delay* treatment. Option A (kept identical in each row of the list) is a dollar payment, paid out immediately but iteratively discounted some number of times. Option B is a dollar payment that involves strictly fewer iterations, and often none, which mimics an immediate or earlier payment. For example, in one row of a list, subjects are asked to choose between "Option A: $42 shrunk 12 times" and "Option B: $2". We again elicit a standard switching interval to calculate the implied "annual impatience".

The mirrors we implement include a single step of discounting for each month of discounting in the *Delay* problem it mirrors. Throughout the experiment, we set the per-period $\delta = 0.96$. This particular value was chosen because it corresponds to the estimated monthly discount factor δ in our intertemporal choice experiments. Our choice to induce a monthly (rather than yearly or daily) discount factor was largely guided by practicality and common sense. First, inducing a yearly discount factor would likely have contributed to participant confusion for delays of less than one year. Second, daily discount factors would have lead to a very large number of discounting steps for delays of several years.[3]

Every subject participated in both *Delay* and *Mirror*, in a random order. The upside of this within-subjects design is that it allows us to correlate behavior in the two types of problems across subjects. When we are not interested in correlating behavior across treatments, we take care to rule out contamination effects by only analyzing decisions from the treatment that a subject encountered first (the results are very similar when we also include the data from the second-assigned treatment, see Appendix Table A.3).

Because the treatments were designed to be compared to one another, we took great pains to use an identical interface and identical numbers. However, we were also careful to strongly differentiate the two treatments from one another using clear instructions. Importantly, to minimize cross-treatment contagion, subjects first assigned to *Mirror* did not know they would later

[3]We recognize that changing the units of steps (e.g., inducing an annual discount factor) might impact behavior because some problems will be easier under any given system (e.g., annual discounting for a delay of one year and monthly discounting for a delay of one month). However, we view this as squarely in line with our complexity-based hypotheses because there is indeed evidence that changing the unit of the time delay (e.g., from months to weeks) in real intertemporal choice causes systematic changes in measured time discounting (Read et al., 2005; Vieider, 2021b).

be making intertemporal choices, and vice versa.

The *Mirror* treatment is incentivized using real payments, but the *Delay* treatment is a purely hypothetical elicitation. This was unavoidable because our motivating questions in *Delay* require us to study choices regarding multi-year delays, which are infeasible to implement using real incentive schemes. There are strong reasons to believe that this is a benign design choice. Reviewing the literature, Cohen et al. (2020) conclude "there is little evidence of systematic differences between RRR in incentivized and unincentivized experiments." Still, to whatever degree hypothetical payments lead to, e.g., less careful decision making in *Delay* than in *Mirror*, we should expect this to work *against* the valuation errors hypothesis we are testing when we contrast the two treatments – we would expect hypotheticals, if anything, to exaggerate anomalies in unincentivized *Delay* observations relative to incentivized *Mirror* observations. We view this, therefore, as a conservative feature of our design. Moreover, below, we use the *Voucher-M* treatment to probe robustness to incentivizing elicitations.

3.3 Measuring Evidence of Noisy or Heuristic Decisions

As motivated in Section 2.2, in other treatments we gather auxiliary evidence that subjects are using decision rules that are error-prone. To do this, we implement treatments *Delay-M* and *Voucher-M*. In both of these treatments, we measure the following objects.

Cognitive Uncertainty. Adapting the methodology from Enke and Graeber (2023), after each choice list, we measure cognitive uncertainty (CU) as the subject's subjective probabilistic belief that their true valuation of the later payment is contained in their stated switching interval:

> *Your choices on the previous screen indicate that you value \$$x_2$ in t_2 somewhere between \$$a$ and \$$b$ in t_1. How certain are you that you actually value \$$x_2$ in t_2 somewhere between \$$a$ and \$$b$ in t_1?*

Participants answer this question by selecting a radio button between 0% and 100%, in steps of 5%, see Appendix Figure A.2. We interpret this question as measuring the participant's awareness that their decision procedure is noisy or heuristic.[4] The measure is not incentivized.

Choice inconsistencies. A standard way of measuring the noisiness of subjects' decision procedure is choice inconsistency in repetitions of the same choice problem. In our study, each subject completes two randomly selected choice lists twice. We generate a binary indicator that equals one if the subject's decisions on the two repeated trials are different from each other. We verify that our results continue to hold if we instead compute the absolute difference between

[4]We ensure that subjects do not misunderstand the question as referring to *external* uncertainty that they may not actually receive the reward. To this effect, our experiments include a comprehension check question that directly asks participants to indicate whether the CU elicitation question asks about (i) the subject's subjective probability of actually receiving the money or (ii) their certainty about their own valuation, given that they know they will receive the money with certainty.

the two decisions as our measure of inconsistency.

We collected these two pieces of data in two different treatments. In *Delay-M* we again use hypothetical monetary payments, which allows us to study multi-year delays. However, we pair this with a second incentivized treatment to show that our findings are robust to the inclusion of incentives. The *Voucher-M* treatment is identical to the *Delay* treatments in most respects, except (i) that we actually pay subjects for their choices using UberEats food delivery vouchers[5] and (ii) we do not study delays of more than one year (for feasibility reasons). In *Voucher-M*, payments are denominated in UberEats vouchers usable starting at date t_1 or $t_2 \leq 12$, respectively; these vouchers are valid for a period of only seven days from the starting date, which minimizes fungibility concerns. Subjects again complete multiple price lists, except that all payments refer to UberEats vouchers (of value between \$40 and \$50).

Participants' vouchers were directly credited to their personal UberEats accounts within 10 hours of completion of the study, such that subjects did not have to actively claim the voucher. The vouchers were always visible in their accounts, they could just not be used before the validity period. Because participants could always view vouchers in their account within a few hours of the study regardless of the precise validity period, there is no differential payment risk across vouchers with different time delays. Participants received automatic reminders 24 hours before a voucher became valid and 24 hours before it expired.

3.4 Procedures

All experiments were conducted on Prolific. Online Appendix D contains details on experimental instructions, visual display and screening questions used.

Subjects in the *Mirror & Delay* sessions were paid a \$6 base payment and had a 20% chance of being paid a bonus based on their choice from a randomly selected list and row of Mirror (or from a separate risk elicitation we included in our sessions). In *Delay-M*, subjects earned a flat \$4.50 payment. In *Voucher-M*, subjects received a \$4 base payment and voucher payments from a randomly selected list and row with 25% chance.

4 Complexity and Hyperbolic Discounting

Evidence from Atemporal Mirrors. We begin by examining whether hyperbolic discounting appears in *Mirror*. In analyzing this data, it is important to emphasize that there are at best weak reasons to expect similar "patience" *levels* in the two treatments. In *Mirror*, subjects face an induced discount factor of 0.96; in *Delay*, choices depend on subjects' individual discount factors, which may differ from 0.96. Our focus will therefore be on comparing the *severity of hyperbolicity* rather than comparing measured patience levels.

[5]UberEats is a takeout delivery service that can be used for a wide array of restaurants. It is widely available throughout the United States (Curry, 2021).

The top panels of Figure 2 provide an overview of the raw data by plotting, for both the *Delay* and *Mirror* treatments, the average switching point (expressed as a percentage of the "later" payment, x_2) as a function of the time interval or the number of discounting iterations.[6] Recall that in an exponential discounting framework with linear utility, these normalized switching points correspond to $\frac{x_1}{x_2} = \delta^{\Delta t}$. For the *Mirror* treatment, we overlay the indifference point that a payoff-maximizing subject would choose given the induced "monthly discount factor" ($\delta = 0.96$). Both panels pool data from Now Lists (the earlier date is immediate or paid with no discounting in *Mirror*) and from Later Lists (the earlier date is in one month or after one step of discounting). Appendix Figure A.3 shows that the results look very similar for both types of lists.

The bottom panels of Figure 2 transform the data in a straightforward way by computing implied annual impatience, $\hat{\gamma} = 1-(x_1/x_2)^{12/\Delta t} = 1-e^{-RRR\cdot 12/\Delta t}$, see eq. (1).

It is clear from these figures that subjects in *Delay* show extremely high impatience over short horizons. Our first finding is that subjects also show extreme discounting over the first few steps of discounting in *Mirror*, even though there is no delay in these problems – and even though subjects are incentivized to maximize an exponential discount function. Importantly, in *Mirror* (unlike in *Delay*) we can identify this behavior as a mistake: subjects discount payments made in $t_1 = 1$ or $t_1 = 2$ to a far greater degree than their true (experimentally induced) discount rate warrants.

A second classical pattern visible in Figure 2 is that indifference payments are a highly compressed function of the time interval. The bottom panels show that this insensitivity implies that implied annual impatience is sharply decreasing in the length of the interval. This pattern of decreasing impatience is a primary motivation for models of non-exponential time preferences like hyperbolic or quasi-hyperbolic discounting.

Our second finding is that in the atemporal mirrors we similarly observe that indifference payments in Figure 2 are too insensitive to the delay, relative to the experimentally-induced discount factor. This insensitivity implies a strong decrease in implied "annual" impatience as the number of discounting steps increases. Once again the figure highlights that this is starkly money-losing mistake: subjects' average switch points in Figure 2 are located above the normative benchmark for few iterations but below it for many iterations.

To compare magnitudes across treatments, Appendix Table A.2 presents regression evidence. In *Delay*, for each additional year, implied annual impatience decreases by 5.6 percentage points (pp). In *Mirror*, that effect is 4.8 pp, meaning that decreasing impatience in *Mirror* is 86% as strong as in *Delay*. This suggests, by a natural decomposition, that the vast majority of hyperbolicity is driven by valuation errors.

Result 1. *Subjects exhibit extreme short-run impatience and decreasing impatience when evaluating atemporal mirrors just as they do when evaluating delays. For mirrors, these are clear misvaluations.*

Results from Proxies for Choice Errors. Next, we turn to analyzing the data from treatments *Delay-M* and *Voucher-M*, where we elicited both cognitive uncertainty and potential choice

[6]We approximate switching points by computing the midpoint of the switching interval.

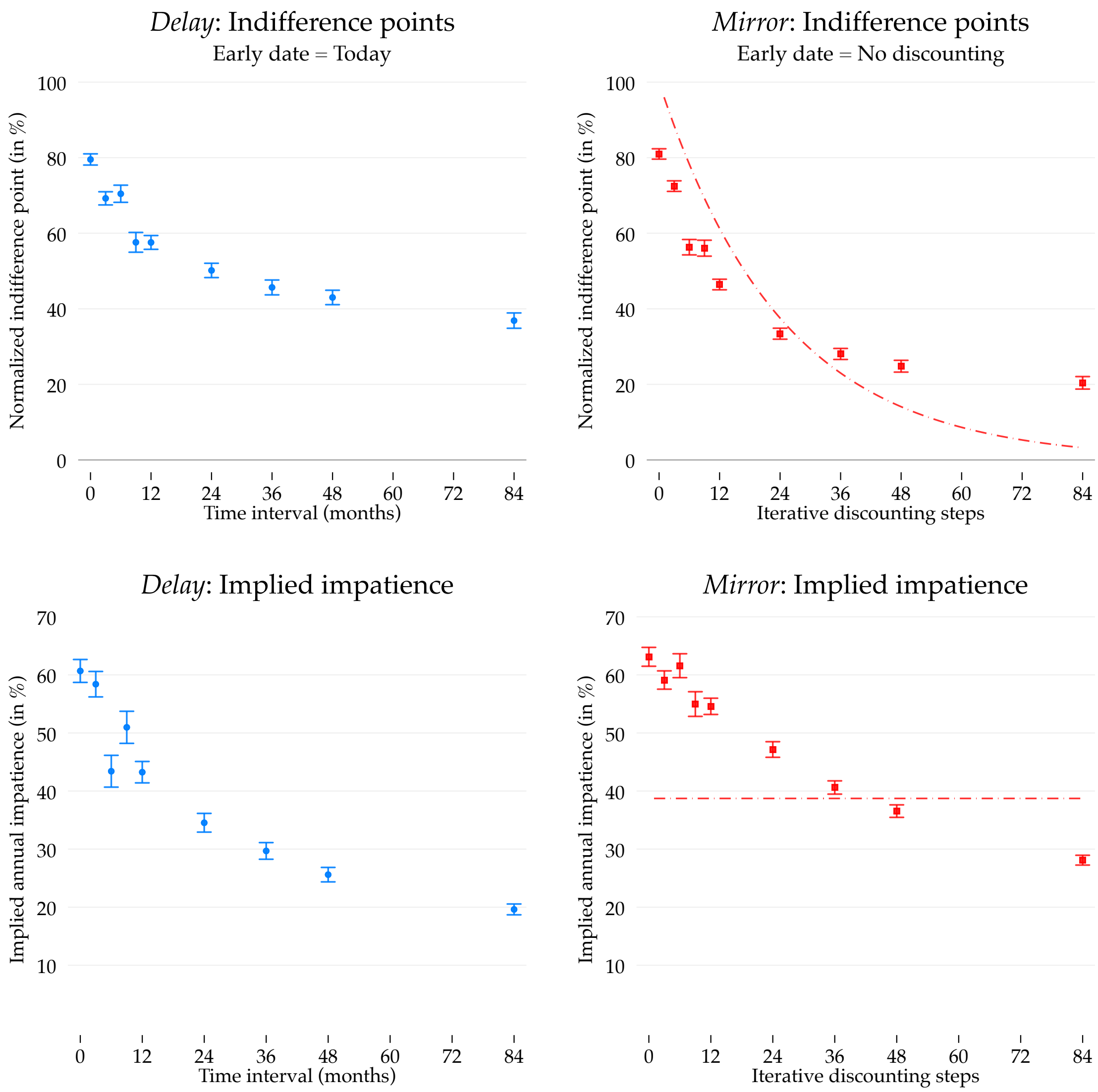

Figure 2: Top panels: Average normalized indifference points by time interval (*Delay*) or number of iterations (*Mirror*). Bottom panels: Average implied annual impatience $\hat{\gamma}$ by time interval or number of iterations. Left panels show *Delay* treatment (4,572 decisions from 254 participants). Right panels show *Mirror* treatment (4,428 decisions from 246 participants). In the *Mirror* panels, the dashed line represents payoff-maximizing decisions. The time interval in months and the number of iterations are rounded to the nearest multiple of three. Whiskers show standard error bars, computed based on clustering at the subject level.

inconsistencies. All findings are statistically highly significant – we only show figures in the main text and relegate econometric analyses to Appendix Table A.4.

The data from the *Delay-M* and *Voucher-M* treatments show strong *prima facie* evidence that subjects' decisions are driven by the use of error-prone valuation rules. In *Delay-M*, 75% of all decisions are associated with strictly positive CU and 60% of all repeated decisions show strictly positive inconsistency. In *Voucher-M*, the corresponding frequencies are 83% and 60%. These results strongly indicate wide-spread usage of imperfectly rational decision procedures in the data. We now investigate how variation in these measures predicts the strength of anomalies.

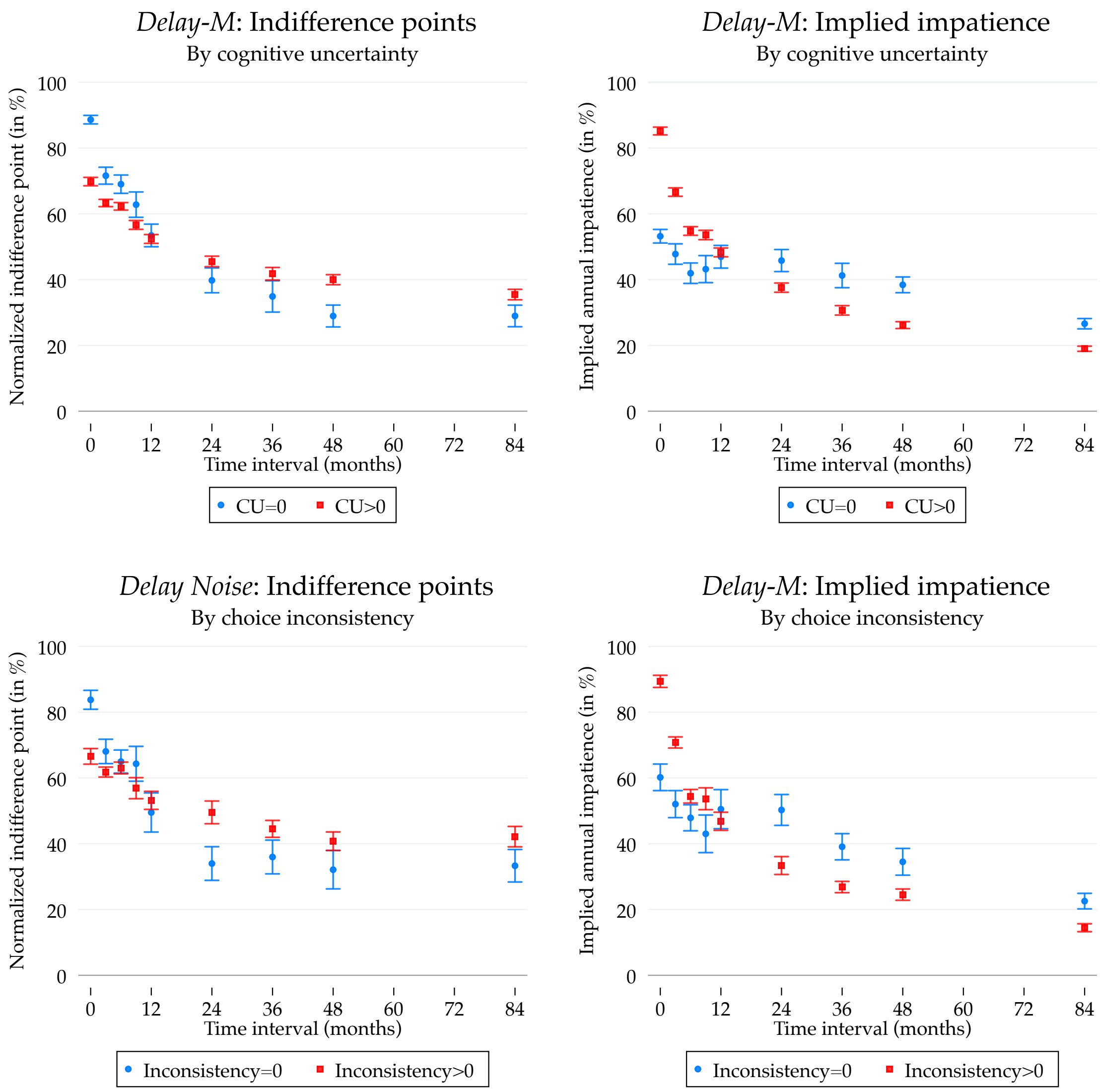

Figure 3: Normalized indifference points (left panels) and implied annual impatience (right panels) as a function of the time interval. The top panels include all decisions from *Delay-M*, and we split the sample according to whether or not a choice is associated with strictly positive CU (7,740 decisions by 645 subjects). The bottom panels include data from all decisions in *Delay-M* that were elicited twice (two repeated problems per subject for a total of 2,580 decisions from 645 subjects), and we split the sample according to whether or not decisions differed in a set of repeated choices. Time intervals are rounded to nearest multiple of three months. Whiskers show standard error bars, computed based on clustering at the subject level.

The left-hand panels of Figure 3 illustrate the raw data for the *Delay-M* treatment: the relationship between normalized indifference points (in percent) and time intervals. The panels split results based on the presence or absence of (i) measured CU in the decision (top panel) or (ii) choice inconsistency in the decision (bottom panel). The corresponding right-hand panels transform these data (as in the previous section) by computing the implied annual impatience $\hat{\gamma}$. All panels pool the data for Now and Later lists (the results are very similar looking at each of them separately). Figure 4 shows analogous results for the incentivized *UberEats* voucher experiments.

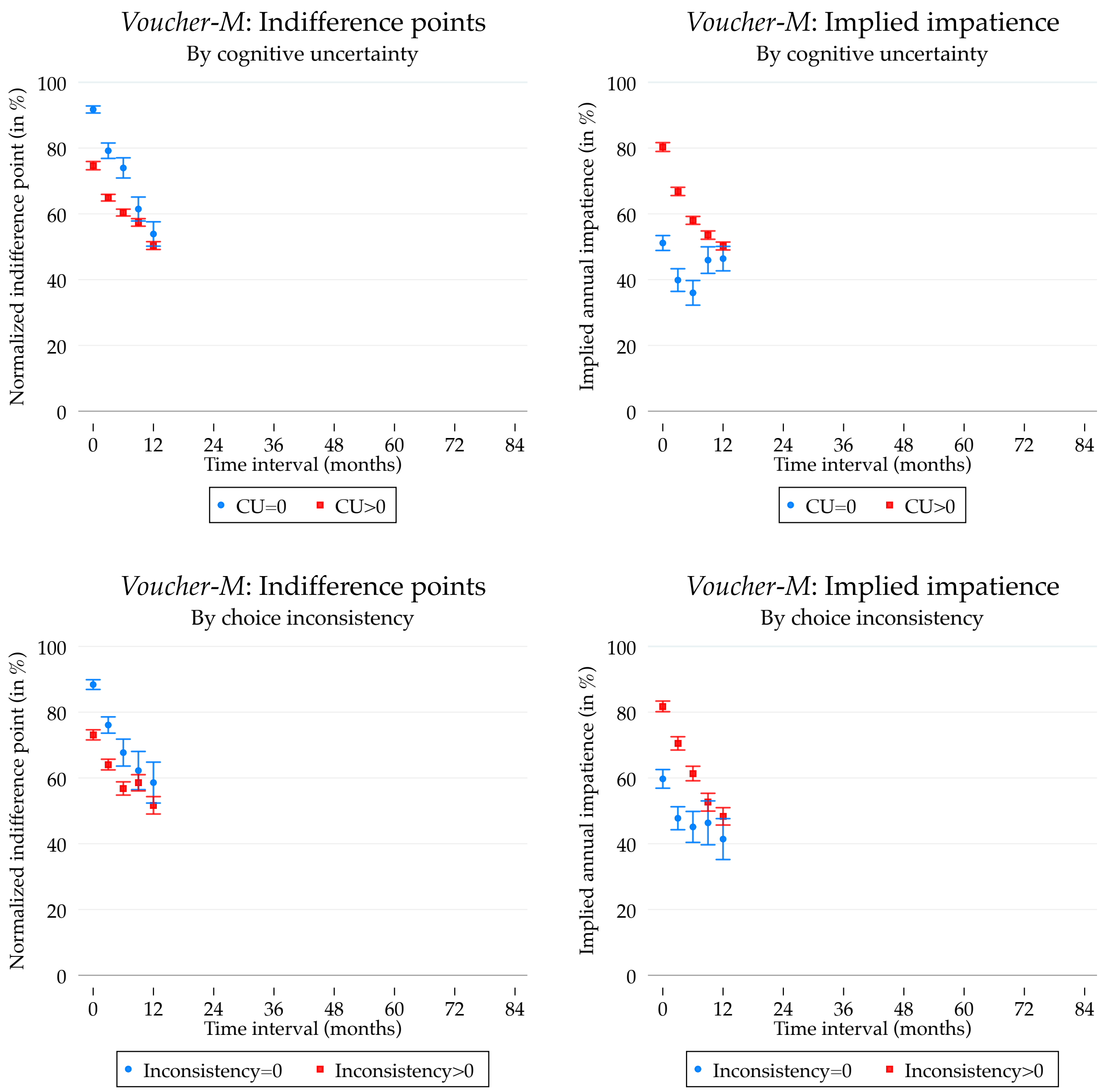

Figure 4: Normalized indifference points (left panels) and implied annual impatience (right panels) as a function of the time interval in *Voucher-M*. The top panels include all decisions, and we split the sample according to whether or not subjects indicate strictly positive CU (6,000 decisions from 500 subjects). The bottom panels include data from all decisions that were elicited twice (two repeated problems per subject, for a total of 2,000 decisions from 500 subjects), and we split the sample according to whether or not decisions differed in a set of repeated choices. Time intervals are rounded to nearest multiple of three months. Whiskers show standard error bars, computed based on clustering at the subject level.

The top panels illustrate that decisions associated with strictly positive CU are considerably less sensitive to variation in the time interval, making them look considerably more hyperbolic. This has two direct implications. First, CU is strongly predictive of short-run impatience. Second, implied annual impatience decreases much more rapidly in the time interval for uncertain than certain subjects. For instance, going from $\Delta t \approx 1$ to $\Delta t \approx 84$ months, the implied annual impatience drops by a factor of 4.5 for $CU > 0$, but only by a factor of 2 for $CU = 0$. In treatment *Delay-M*, this pattern implies that cognitively uncertain participants act as if they are *less* patient over relatively short horizons, yet *more* patient over relatively long horizons. These results do not hinge on splitting the sample into decisions with zero versus strictly positive CU. To show this, we split the sample into CU quartiles. We find that the effect of the time interval on decisions continuously decreases (in absolute terms) as CU increases, see Appendix Figure A.4.

Strikingly, the strong link between CU and insensitivity to delays is also present in within-subject comparisons. To show this, we normalize the CU data to have mean zero and standard deviation one for each subject, and then look at whether this pure within-subject measure still predicts choices. Appendix Table A.5 shows that this is the case.

The bottom panels of Figures 3 and 4 show analogous results for choice inconsistency. Notably, these patterns from cognitive uncertainty and choice inconsistency match exactly what we find in *Mirror*, where subjects are too "impatient" with few discounting iterations but too "patient" with many ones.

What fraction of decreasing impatience is driven by valuation mistakes? To quantify this, we compare the magnitudes in two sub-samples: (i) decisions that are associated with no CU and no choice inconsistency vs. (ii) decisions that reflect either strictly positive CU or choice inconsistency. We examine how strongly implied annual impatience increases in the evaluated time interval, akin to the regressions in Table A.4. We find that in the sample with no CU and no choice inconsistencies, the magnitude of decreasing impatience is only 10% of that in the comparison sample. This suggests that at least 90% of decreasing impatience is driven by valuation errors, rather than preferences. This conclusion is strikingly close, quantitatively, to the decomposition computed by comparing decreasing impatience in atemporal mirrors and time intervals.

Result 2. *Short-run impatience and decreasing impatience are strongly correlated with auxiliary evidence of valuation errors.*

Linkage between Atemporal and Intertemporal Decisions. Next, we show that the results from our atemporal mirrors and true intertemporal choice are driven by the same mechanism and that this mechanism is valuation errors. We show this in two ways.

First, we show that anomalies in *Mirror* and *Delay* are linked at the individual level. To do this, unlike in the analyses above, we leverage the within-subjects design of treatments *Delay* and *Mirror* to examine the within-subject relationship between behaviors across the two treatments. If there is a common behavioral mechanism behind the anomalies across treatments, behavior in the two cases should be correlated with each other. And since this shared mechanism can only be valuation errors in *Mirror* this serves as direct evidence that valuation errors drive anomalies in *Delay* as well.

We link subjects' decisions in those choice problems that are direct mirror images of each other, such as "\$40 in 6 months" vs. "\$40 shrunk 6 times". Thus, we compute a correlation coefficient for (500 subjects * 18 unique problems * 2 treatments =) 18,000 observations. In doing so, we take care to net out that component of the correlation that is mechanically driven by the fact that for longer intervals or a higher number of iterations subjects should be expected to state lower valuations. Thus, we compute the partial correlation between decisions, netting

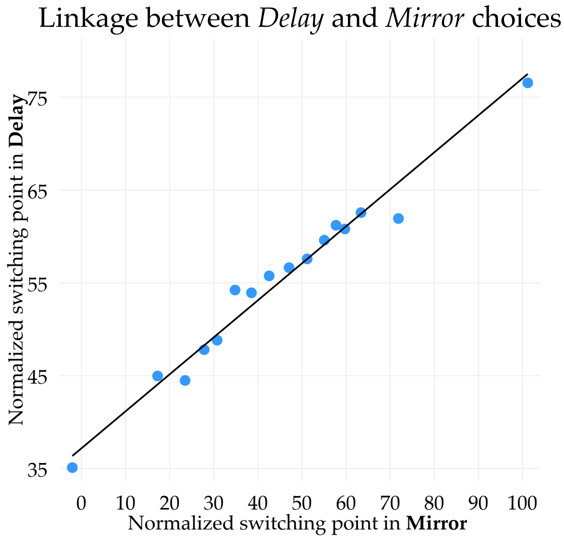

Figure 5: Binned scatter plot of normalized indifference points in structurally identical choice problems in *Delay* and *Mirror*. Partial correlation plot, controlling for fixed effects for each choice list type (each possible combination of t_1 and t_2). Based on 18,000 decisions by 500 subjects. The partial correlation is $r = 0.34$.

out fixed effects for each unique problem type (each possible combination of t_1 and t_2). As a result, the correlation captures how similar subjects' behavior is across the two treatments, holding fixed the nature of the choice problem.

We find a partial correlation of $r = 0.34$ ($p < 0.01$), see Figure 5. This correlation is remarkably high given that the absence of time preference-based variation in the *Mirror* treatment should produce correlations *close to zero* for rational decision makers. Instead, behavior in *Mirror* produces one of the strongest predictors of intertemporal choice ever documented in the literature (Cohen et al., 2020).[7]

A second piece of evidence that anomalies in the two domains are driven by a common mechanism is that we find identical evidence that choice inconsistencies predict hyperbolicity (as documented above) in *Mirror* and *Delay*. In our *Mirror* treatment (just as in *Delay*), we repeated one randomly-selected choice list for each subject. As shown in Appendix Table A.6, we find that choice inconsistencies are strongly predictive of "short-run impatience" and "decreasing impatience" in mirror valuations, just as they are of true intertemporal decisions. This further supports our interpretation that anomalies represent valuation mistakes.

Result 3. *Across subjects, valuation of time intervals is strongly correlated with valuation of atemporal mirrors. Moreover, valuation of delayed payments and atemporal mirrors are both strongly correlated with choice noise, suggesting behavior across the two domains is driven by a common mechanism (valuation errors).*

[7]For instance, the correlation between valuations of atemporal mirrors and delays is in the same ballpark as the correlation documented between identical intertemporal list choices made by the same subjects in elicitations delivered several months apart (Meier and Sprenger, 2015).

Manipulation of Task Difficulty. We interpret the correlations between anomalies and CU / choice inconcistency as evidence that anomalies arise due to error-prone responses to the aggregation problem implied by intertemporal choice. That is, we interpret the valuation errors we measure as a direct *response* to the fact that intertemporal choice problems are costly or difficult to evaluate. To provide direct evidence for this linkage, we ran an additional experiment that exogenously increases the cognitive difficulty of intertemporal choice. To the degree this manipulation jointly intensifies (i) our signatures of valuation errors and (ii) intertemporal choice anomalies, we have complementary causal evidence supporting our interpretation.

In treatment *Complex Payments/Delays*, for a subset of subjects, ($N = 153$), we express all of the payoffs in the price list as an algebraic expression (e.g., $40 is described as "$(4*8/2)+(8*9/2)-12"). For another subset ($N = 149$), we express all dates in the price list as algebraic expressions (e.g., 1 year is described as "in (6*2/3-3) years AND (3*6/2-9) months AND (5*4/2-10) days"). These interventions are always paired with time constraints of 25 seconds to make the relevant information processing constraints more likely to bind.

We find that this intervention significantly increases both of our measures of boundedly rational choice. Average CU rises from 21.7% in *Delay-M* to 35.2% for *Complex Payments/Delays*; choice inconsistency rises from 60.4% in *Delay-M* to 67.2% in *Complex Payments/Delays* (both comparisons are statistically significant at least at the 5% level, see Appendix Table A.7).

Next, we find that this manipulation simultaneously intensifies intertemporal choice anomalies. As Figure 6 shows, the decisions of subjects in *Complex Payments/Delays* evince stronger short-run impatience and flatter long-run impatience than those of subjects in *Delay-M* (see Appendix Tables A.7 and A.8 for regression evidence). Thus, the exogenous manipulation of task difficulty has the same effects as the patterns we observed correlationally for choice inconsistency and CU.

Result 4. *Short-run impatience and decreasing impatience become significantly more pronounced when complexity is exogenously increased.*

Mechanism: Complexity and Insensitivity. The key takeaway from the preceeding analysis is that complexity causes hyperbolicity becauses it induces an insensitivity of decisions with respect to the delay. To further sharpen this point, we consider a second canonical intertemporal choice anomaly, so-called subadditivity effects. Documentations of subadditivity are the standard method in the literature for measuring insensitivity to time delays. The subadditivity literature shows that impatience over a single time interval (t_1, t_3) tends to be considerably smaller than the total impatience people reveal when they are asked to make two decisions, one over interval (t_1, t_2) and one over (t_2, t_3), with $t_1 < t_2 < t_3$ (Read, 2001).[8] The resulting transitivity violations are direct evidence of insensitivity (i) because they involve people treating shorter intervals too much like they treat a longer interval, and (ii) because this cannot be confounded

[8] Formally, denote by $a_{i,j}$ the indifference point for the tradeoff over interval (t_i, t_j). Then, subadditivity means that there is less discounting (more patient indifference values) over the single long interval: $a_{1,3} > a_{1,2}a_{2,3}$, or, equivalently, $\gamma(a_{1,3}) > \gamma(a_{1,2}a_{2,3})$.

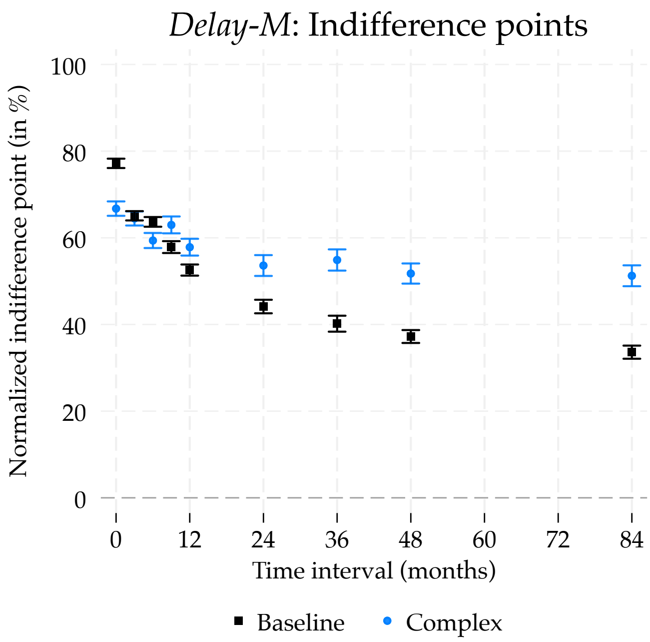

Figure 6: Normalized indifference points as a function of the time interval (rounded to nearest multiple of three months) in *Delay-M* and the two *Complex* manipulations, pooled for ease of readability (11,364 decisions from 947 subjects). Whiskers show standard error bars, computed based on clustering at the subject level.

with non-stationarities in discounting because the decisions involve the comparison of nested intervals.

To investigate whether complexity produces the insensitivities to the interval length that are typically observed in these tasks, we included in all of our experiments choice lists in which we asked subjects to complete tasks that have a subadditivity structure, where we varied (t_1, t_2, t_3) randomly to be $(0, 4, 8)$ or $(0, 6, 12)$.

Table 2 summarizes the evidence. Columns (1) and (2) show how implied annual impatience differs between the choice over interval (t_1, t_3) and the combined choices over (t_1, t_2) and (t_2, t_3), separately for treatments *Delay* and *Mirror*. We find strong evidence for subadditivity in both treatments: people are roughly 10pp less "patient" when a composite interval is broken up into two sub-intervals. Most importantly, the effect is *similarly strong* in atemporal mirrors and true delays, suggesting that all of the insensitivity of subadditivity is attributable to complexity-driven mistakes.

Columns (3)–(6) present the results on cognitive uncertainty in treatments *Delay-M* and *Voucher-M*. In both treatments, we find that the magnitude of subadditivity is strongly correlated with CU. Indeed, we find that subjects with $CU = 0$ exhibit no subadditivity at all. Thus, again, valuation errors seem to entirely explain the insensitivity of decisions to the interval. Finally, consistent with these correlational results, Appendix Table A.7 shows that subadditivity effects also become substantially more pronounced in our *Complex* treatments that increase the difficulty of intertemporal decision making, creating a third link to errors.

Table 2: Complexity and subadditivity

	Dependent variable: Implied annual impatience (in %)					
Phenomenon:	Subadditivity					
Treatment:	*Delay*	*Mirror*	*Delay-M*		*Voucher-M*	
	(1)	(2)	(3)	(4)	(5)	(6)
1 if one long interval	-7.57***	-9.93***	-8.58***	-3.55***	-9.39***	-1.14
	(1.38)	(1.16)	(0.63)	(1.34)	(0.60)	(1.60)
Cognitive uncertainty				0.47***		0.45***
				(0.06)		(0.08)
1 if one long interval × Cognitive uncertainty				-0.24***		-0.33***
				(0.06)		(0.06)
Payment amount FE	Yes	Yes	Yes	Yes	Yes	Yes
Task set FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	508	492	1948	1948	2000	2000
R^2	0.09	0.06	0.03	0.08	0.04	0.08

Notes. OLS estimates, robust standard errors (in parentheses) are clustered at the subject level. The data are restricted to problems that have a subadditivity structure. We combine the three choices that make up a subadditivity set into two observations according to fn. 8. Task set FE are fixed effects for each pair of tasks that have a subadditivity structure. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Thus, the atemporal mirrors, the measure of cognitive uncertainty and the experimental complexity manipulation all suggest that the errors subjects make in these valuations produce an insensitivity of decisions to the delay. Since the insensitities measured by subadditivity effects have been linked to hyperbolicity by the prior literature (Read, 2001), we take this as further evidence that hyperbolicity in the empirical discount function is primarily a consequence of complexity-driven inensitivities to delays.

Result 5. *Interval insensitivities, as measured by subadditivity effects, are entirely driven by valuation errors.*

5 Complexity and Estimates of Present Bias

Our findings so far suggest that the hyperbolic shape of the empirical discount function in monetary rewards is largely driven by mistakes in aggregating the components of an intertemporal choice problem. What guidance can this provide for efforts to measure *present bias*, perhaps the most often measured empirical object in the intertemporal choice literature?

Structural estimates of present bias. We begin by measuring present bias using the approach taken by structurally estimating the parameters of a $\beta - \delta$ model, as described in Section 2. Intuitively, in these model estimations, present bias is identified off the hyperbolicity of the empirical discount function, including especially the excess degree of short-run impatience not captured by the estimated exponential discounting parameter, δ. Because we know from the previous section that this hyperbolicity is largely driven by complexity-driven valuation errors, there are strong reasons to hypothesize that structural estimates of β will, likewise, be confounded by complexity and mistakes. Because in this section we will frequently distinguish between structural and experimental estimates of present bias, we will denote structural estimates by $\hat{\beta}_{ST}$.

We first examine whether there is evidence for $\beta_{ST} < 1$ in our *Mirror* treatment, in which exponential discounting is experimentally induced and present-biased motivations are removed by design. Recall that in all of our experiments, a subject is asked to state an amount x_1 in t_1 that makes her indifferent to x_2 in t_2. In a $\beta - \delta$ model with linear utility, we, hence, have:

$$\delta^{t_1} \cdot x_1 = \beta_{t_1=0} \cdot \delta^{t_2} \cdot x_2 \tag{2}$$

We estimate this model at the population level, amended by a mean-zero error term. In *Mirror*, we estimate $\hat{\beta}_{ST} = 0.85$ ($s.e. = 0.01$) and $\hat{\delta} = 0.96$ ($s.e. = 0.01$).[9] Valuation mistakes alone, therefore, induce behavior that *looks like* standard levels of present bias under the lens of standard estimation approaches. Intuitively, the reason for this result is that decisions in the *Mirror* treatment have a hyperbolic shape with high short-run "impatience", which gets attributed to a sub-unitary β_{ST}. Indeed, our estimates recover the true, induced δ of 0.96, suggesting that most of the distorting effects of valuation errors appear in the spurious estimate of β.

If errors indeed confound structural estimates of present bias, we should also see that – in traditional intertemporal choice experiments – decisions that are associated with stronger errors (cognitive uncertainty and choice inconsistencies) are associated with more pronounced estimated present bias. To investigate this, we turn to the data from the *Delay-M* treatment. As is well-understood in the literature, individual-level heterogeneity in discount factors renders population-level estimates of β_{ST} potentially biased (Weitzman, 2001; Jackson and Yariv, 2014). Thus, we estimate eq. (2) separately for each subject.[10] Figure 7 shows a binned scatter plot of the resulting individual-level estimates of $\hat{\beta}_{ST}$ against a summary index of signatures of valuation errors (the first principal component of CU and choice inconsistency). Estimated present bias is strongly concentrated in subjects with evidence of valuation errors (Spearman's $\rho = -0.28$, $p < 0.01$). Appendix Figure A.6 shows that quantitatively almost identical results hold when the estimation accounts for utility curvature (measured through separate lottery choice lists at the end of the experiment). In combination with the result of "present bias" in the atemporal

[9]Note that because all subjects are induced to have the same time preferences, estimates of β in atemporal mirrors do not run afoul of the aggregation concerns raised in the literature (Weitzman, 2001; Jackson and Yariv, 2014). Nonetheless, estimates at the individual level corroborate this result. As shown in Appendix Figure A.5, for the majority (62%) of subjects we estimate $\hat{\beta}_{ST} < 1$.

[10]Population-level estimates deliver similar results on how β varies with signatures of valuation errors. Appendix Table A.8 reports the results. For example, for $CU = 0$, we estimate $\hat{\beta}_{ST} = 0.87$, while for $CU > 0$ we get $\hat{\beta}_{ST} = 0.72$. In contrast, the estimates of δ are always very similar across the different sub-samples, suggesting that (as with our estimates from *Mirror*), valuation errors mostly influence the present bias β term in estimates of these models. These results show that *even if aggregation was not an issue*, valuation mistakes would still bias the estimation of present bias.

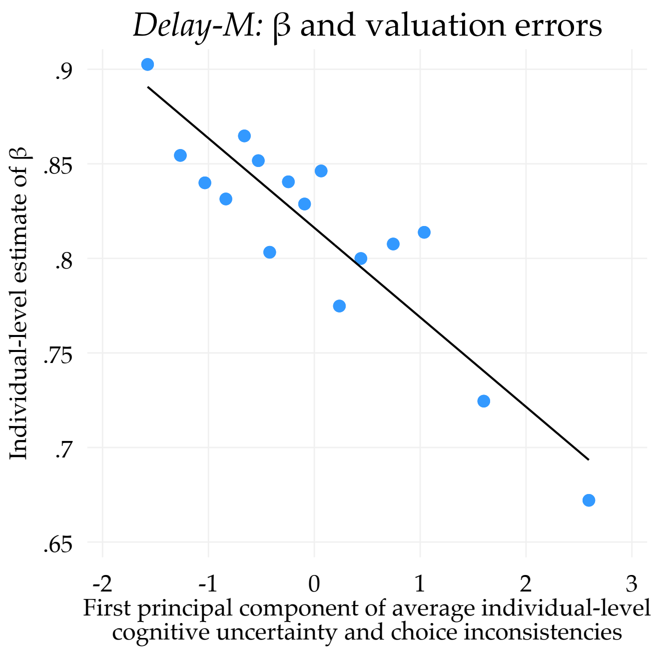

Figure 7: Binned scatter plot of individual-level $\hat{\beta}_{ST}$ in eq. (2) against first principal component of subject-level average CU and choice inconsistencies. Based on 643 subjects in *Delay-M*. The figure excludes two subjects for whom we estimate $\hat{\beta} > 2$.

mirrors, this strongly suggests that conventional structural estimates of present bias to a great extent pick up mistakes in valuation.

Causal Estimates of Present Bias. A standard way of causally identifying present bias in the literature is by measuring *front-end delay effects* (direct measurements of stationarity violations). Indeed, it is common in the literature to *define* genuine present bias through front-end delay effects (e.g., Chakraborty, 2021). In experimental documentations of these effects, subjects reveal lower discounting in evaluating $(t_1 + d, t_2 + d)$ than in (t_1, t_2), for $d > 0$. Some of our tasks feature such a front-end delay structure (with $t_1 = 0$ and d randomized between 4 and 6 months across subjects).

Table 3 summarizes the evidence on the link between valuation errors and front-end delay effects in our data. Columns (1) and (2) show that we find a statistically significant front-end delay effect in the *Delay* treatment but the *opposite* effect in the *Mirror* treatment. Thus, the mirror data provide no evidence that the front-end delay effect is an outgrowth of valuation errors. If anything, our results suggest that such errors might even work against the identification of these effects.

Columns (3)–(6) shows the results for treatments *Delay-M* and *Voucher-M*.[11] Again, we find

[11]Recall that we elicited only two randomly selected decisions per subject repeatedly. Given that these repeated decisions do not always occur for the choices in the SA/FED lists, we do not have access to a task-level measure of choice inconsistency that can be used to shed light on subadditivity or front-end delay effects. By contrast, the CU measure is available for each decision a subject makes.

Table 3: Complexity and front-end delay effects

	Dependent variable: Implied annual impatience (in %)					
Phenomenon:	Front-end delay					
Treatment:	*Delay*	*Mirror*	*Delay-M*		*Voucher-M*	
	(1)	(2)	(3)	(4)	(5)	(6)
1 if front end delay	-4.24**	3.79**	-3.07***	-2.51	-4.11***	-7.23***
	(1.85)	(1.69)	(0.99)	(1.53)	(1.09)	(2.12)
Cognitive uncertainty				0.38***		0.38***
				(0.06)		(0.07)
1 if front end delay × Cognitive uncertainty				-0.058		0.070
				(0.05)		(0.07)
Payment amount FE	Yes	Yes	Yes	Yes	Yes	Yes
Task set FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	508	492	2393	2393	2337	2337
R^2	0.07	0.02	0.02	0.07	0.02	0.08

Notes. OLS estimates, robust standard errors (in parentheses) are clustered at the subject level. The data are restricted to problems that have a front-end delay structure. Task set FE are fixed effects for each pair of tasks that have a front-end delay structure. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

evidence for the presence of front-end delay effects in intertemporal decisions. Importantly for our purposes, however, these effects are entirely uncorrelated with cognitive uncertainty, suggesting again that they have little to do with complexity and mistakes. Finally, we show in Appendix Table A.7 that the experimental complexity manipulation described in Section 4 does not amplify front-end delay effects, providing a third piece of evidence that front-end delay effects have little to do with the use of error-prone decision rules.

Thus, the atemporal mirrors, the measure of cognitive uncertainty and the experimental complexity manipulation all suggest that non-stationarity is *not* driven by mistakes. This result is squarely in line with the results discussed above. The main implication of complexity in intertemporal choice appears to be an *insensitivity to the length of the delay* – yet because the delay is held constant across the two choice problems that are designed to identify front-end delay effects, there is no reason to expect them to be impacted by complexity-driven insensitivity effects.

Reconciling structural and experimental estimates. How is it possible that valuation errors are strongly linked to structural estimates of present bias but not to causal, experimental estimates? To address this, we first document that the magnitude of present bias inferred from front-end delays is quantitatively considerably smaller than what is implied by the structural estimate $\hat{\beta}_{ST}$ reported above. To compute a causally-identified front-end-delay estimate (β_{FD}), we estimate eq. (2) only on those decision problems that have a front-end delay structure. In our two sets of intertemporal problems with a front-end delay structure, we estimate $\hat{\beta}_{FD} = 0.95$

and $\hat{\beta}_{FD} = 0.96$, respectively. While these estimates suggest strictly positive present bias, they are substantially larger (i.e. imply substantially less present bias) than the average individual-level structural estimate derived from estimating eq. (2) on the full dataset, $\hat{\beta}_{ST}^{ave} = 0.83$. A back-of-the-envelope calculation, hence, tentatively suggests that the structural estimate of 0.17 units of present bias can be roughly decomposed into 0.05 units attributable to true present bias (non-stationarity) and 0.12 units attributable to valuation errors.

Intuitively, the large difference between the structural estimate and the front-end delay estimate of present bias arises because the empirical discount function is substantially more hyperbolic than the magnitude of front-end delay effects would imply. Indeed, this discrepancy between the magnitude of front-end delay effects and of hyperbolic discounting is also highlighted in the review of Cohen et al. (2020). They call the coexistence of strongly decreasing impatience and relatively small front-end delay effects "contradictory patterns". Our results show that valuation errors are the main driver behind this discrepancy because they produce hyperbolicity but not front-end delay effects.[12]

To sum up, the severely inflated magnitude of structural estimates of present bias reflects model misspecification: conventional estimates of a $\beta - \delta$ model do not account for valuation mistakes, such that the error-induced hyperbolicity of the discount function gets spuriously attributed to β.

Result 6. *Structural estimates of present bias (that do not rely on causal experimental designs) are severely biased due to model misspecification resulting from omiting valuation errors. On the other hand, treatment-based estimates of present bias are unconfounded by complexity.*

6 Discussion

Table 4 summarizes the results from our paper across all of our treatments. The main takeaway is that regardless of how we operationalize and measure complexity effects and the mistakes they produce (through atemporal mirrors, choice inconsistency, cognitive uncertainty and exogenous treatment interventions), we consistently find that mistakes are strongly associated with short-run impatience, decreasing impatience, subadditivity and structural estimates of present bias. Indeed, across methods, we find strikingly similar *quantitative* evidence that each of these signatures of hyperbolicity is *primarily* attributable to such errors. We interpret this as evidence that intertemporal tradeoffs over money generate behavioral distortions in large part because they require a difficult cognitive act, which produces an insensitivity of decisions to delays. In contrast, treatment-based estimates of front-end delay effects ("true" present bias) are unconfounded by complexity.

Which models are consistent with the evidence? We now briefly discuss which of the models of noisy or heuristic decision procedures are, *prima facie*, consistent with the entirety

[12] Cohen et al. (2020) infer as much, attributing the remainder of hyperbolicity to the insensitivities described by subadditivity effects.

Table 4: Summary of results across experiments

	Short-run impatience	Decreasing impatience	Sub-additivity	Front-end delay effect	Estimated present bias
Present in atemporal mirrors?	✓	✓	✓	–	✓
More pronounced with cognitive uncertainty?	✓	✓	✓	x	✓
More pronounced with choice inconsistency?	✓	✓	n/a	n/a	✓
More pronounced in difficult problems?	✓	✓	✓	–	✓

Notes. "✓" means that an anomaly is present / more pronounced, "x" that it is not present / not more pronounced and "–" that the opposite is present / the anomaly is less pronounced. "n/a" means that data limitations do not allow us to assess a relationship.

of our evidence. Models of random utility (He et al., 2019; Lu and Saito, 2018) or imperfect time perception (Zauberman et al., 2009) cannot explain why hyperbolicity is also present in atemporal mirrors, or why subadditivity is strongly correlated with cognitive uncertainty.

The mirrors data in particular suggest that hyperbolicity reflects more the difficulty of *aggregation* than noisy perception or representation of individual components of intertemporal choice problems.[13] However, if one is willing to take the interpretation that models of "noisy cognition" capture the difficulty of aggregation rather than literally the noisy coding of dates or dollar amounts (which is how they are literally written), several models can explain some or even all of our evidence. For instance, if imperfect time perception reflects the noise that arises in the process of aggregating delays with other problem components, then the model of noisy coding of time in Vieider (2021b) seems to organize our data quite well because it predicts that cognitive noise generates both hyperbolicity and subadditivity.

The literature that models imperfect mental simulations of future utils (Gabaix and Laibson, 2022; Gershman and Bhui, 2020) can under some ancillary assumptions rationalize the result that hyperbolicity is correlated with cognitive uncertainty. However, it is not clear whether these models (which are explicitly motivated by the hypothesis of limited foresight) should apply to the atemporal mirrors; furthermore, they cannot explain the fact that subadditivity is strongly correlated with cognitive uncertainty.

Finally, any of a number of heuristic decision processes could be consistent with our data. For example, all of our data are consistent with the idea that subjects anchor their valuation of the delayed payment at some intermediate value (e.g., "50% of the delayed payment") and then heuristically and imperfectly adjust up or down depending on the precise delay in a problem, where the magnitude of the adjustment is inversely proportional to the noisiness of subjects decision process, such that a noiseless agent states their true valuation. The appendix of an earlier working paper version of this paper (Enke and Graeber, 2021) spells out such a model.[14]

[13] In principle, the mirrors data could be consistent with noisy coding of inputs (numbers). We are skeptical of this account, one of the reasons being that the recent neuroscience literature has provided much evidence that cognitive noise primarily arises in the process of aggregation rather than from imperfect number perception (e.g. Drugowitsch et al., 2016).

[14] A somewhat related possibility is that compression-to-the-center is driven by noise or measurement error in subjects' responses that bounces off the boundaries of zero and one. Such boundary effects can produce error distributions that are asymmetric and, hence, lead to compression to the center.

Which intertemporal aggregation problems are more difficult? Our data provide some indication of what it is about intertemporal choice that produces a difficult aggregation problem. One possibility, ex ante, is that it is difficult to introspectively evaluate or calculate one's own time preferences. However, the fact that hyperbolic discounting arises with near-full strength in atemporal mirrors instead suggests that complexity is driven by the difficulty of iterative discounting (successive multiplication). Appendix C provides some evidence in support of this idea. There, we document that both cognitive uncertainty and the variance of subjects' decisions strongly increase in the interval length or the number of iterations required to discount. For example, for very short delays average cognitive uncertainty is very small. This evidence suggests that repeated discounting is cognitively costly, producing noise that increases in the number of iterations required to discount a reward.

Money vs. consumption. The motivation of our paper is to understand intertemporal *financial* decision-making — a central type of intertemporal choice in any modern, money-based economy. We speculate that our findings may also extend to intertemporal consumption decisions. After all, intertemporal consumption choices require no less complex aggregation of problem components than intertemporal monetary choice, making it plausibly subject to similar barriers to preference recovery (Chakraborty et al., 2017; Carrera et al., 2022). Examining to what degree this is true will require new and different types of experiments, but seems like an important next step for the literature.

Broader takeaways and connection to decision making in other domains. An important takeaway from all of the experiments reported in this paper is that costly cognitive information processing (and the errors it induces) produces a *particular type of behavioral response*: an insufficient elasticity of decisions to variation in the main parameter of the problem, the length of the time interval.[15] This observation may suggest deep connections between intertemporal choice anomalies and other anomalies that have similarly been identified as growing out of complexity-derived mistakes. In two recent papers, Oprea (2022) and Enke and Graeber (2023), we show that some of the core anomalies behavioral economists have observed in the domain of risk (such as probability weighting or conservatism in belief updating) are similarly rooted in complexity and the errors it induces. In particular, an overarching message that emerges in the recent literature is that, when decisions involve non-trivial information processing, observed behavior is insufficiently sensitive ("attenuated") with respect to variation in objective problem parameters, including probabilities, deterministic frequencies, time delays, and atemporal discounting iterations. We view this generic insensitivity as a potentially unifying principle for behavioral economics anomalies. If true, this would suggest that many apparently distinct phenomena in behavioral economics might be parsimoniously united by models built to describe the way humans manage and respond to complexity.

[15] See Ebert and Prelec (2007) and Epper et al. (2019) for a related discussion.

References

Abdellaoui, Mohammed, Emmanuel Kemel, Amma Serwaah-Panin, and Ferdinand M Vieider, "Time for Tea: Measuring Discounting for Money and Consumption without the Utility Confound," Technical Report, Centre for Economic Policy Research 2023.

Agranov, Marina and Pietro Ortoleva, "Stochastic choice and preferences for randomization," *Journal of Political Economy*, 2017, *125* (1), 40–68.

__ , Paul J Healy, and Kirby Nielsen, "Non-Random Randomization," *Available at SSRN*, 2020.

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

Andreoni, James, Christina Gravert, Michael A Kuhn, Silvia Saccardo, and Yang Yang, "Arbitrage or narrow bracketing? On using money to measure intertemporal preferences," Technical Report, National Bureau of Economic Research 2018.

Arts, Sara, Qiyan Ong, and Jianying Qiu, "Measuring subjective decision confidence," *Working Paper*, 2020.

Ashraf, Nava, Dean Karlan, and Wesley Yin, "Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines," *The Quarterly Journal of Economics*, 2006, *121* (2), 635–672.

Attema, Arthur E, Han Bleichrodt, Yu Gao, Zhenxing Huang, and Peter P Wakker, "Measuring discounting without measuring utility," *American Economic Review*, 2016, *106* (6), 1476–1494.

Augenblick, Ned, "Short-term time discounting of unpleasant tasks," *Unpublished manuscript*, 2018.

__ and Matthew Rabin, "An experiment on time preference and misprediction in unpleasant tasks," *Review of Economic Studies*, 2019, *86* (3), 941–975.

__ , Eben Lazarus, and Michael Thaler, "Overinference from Weak Signals and Underinference from Strong Signals," *arXiv preprint arXiv:2109.09871*, 2021.

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

Ba, Cuimin, J Aislinn Bohren, and Alex Imas, "Over-and underreaction to information," *Available at SSRN*, 2022.

Balakrishnan, Uttara, Johannes Haushofer, and Pamela Jakiela, "How soon is now? Evidence of present bias from convex time budget experiments," *Experimental Economics*, 2020, *23*, 294–321.

Blumenstock, Joshua, Michael Callen, and Tarek Ghani, "Why do defaults affect behavior? Experimental evidence from Afghanistan," *American Economic Review*, 2018, *108* (10), 2868–2901.

Brocas, Isabelle, Juan D Carrillo, and Jorge Tarrasó, "How long is a minute?," *Games and Economic Behavior*, 2018, *111*, 305–322.

Brownback, Andy, Alex Imas, and Michael A Kuhn, "Time Preferences and Food Choice," Technical Report, National Bureau of Economic Research 2023.

Carrera, Mariana, Heather Royer, Mark Stehr, Justin Sydnor, and Dmitry Taubinsky, "Who chooses commitment? Evidence and welfare implications," *The Review of Economic Studies*, 2022, *89* (3), 1205–1244.

Carvalho, Leandro S, Stephan Meier, and Stephanie W Wang, "Poverty and economic decision-making: Evidence from changes in financial resources at payday," *American economic review*, 2016, *106* (2), 260–284.

Castillo, Marco, Jeffrey L Jordan, and Ragan Petrie, "Discount rates of children and high school graduation," *The Economic Journal*, 2019, *129* (619), 1153–1181.

Chabris, Christopher F., David Laibson, Carrie L. Morris, Jonathon P. Schuldt, and Dmitry Taubinsky, "Individual Laboratory-Measured Discount Rates Predict Field Behavior," *Journal of Risk and Uncertainty*, 2008, *37* (2-3), 237.

Chakraborty, Anujit, "Present bias," *Econometrica*, 2021, *89* (4), 1921–1961.

__ , Evan M Calford, Guidon Fenig, and Yoram Halevy, "External and internal consistency of choices made in convex time budgets," *Experimental economics*, 2017, *20* (3), 687–706.

Chapman, Jonathan, Mark Dean, Pietro Ortoleva, Erik Snowberg, and Colin Camerer, "Econographics," *Journal of Political Economy Microeconomics*, 2023, *1* (1), 115–161.

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

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

Curry, David, "Uber Eats Revenue and Usage Statistics (2021)," 2021.

Dertwinkel-Kalt, Markus, Holger Gerhardt, Gerhard Riener, Frederik Schwerter, and Louis Strang, "Concentration bias in intertemporal choice," *Review of Economic Studies*, 2021.

__ , __ , __ , __ , and __ , "Concentration bias in intertemporal choice," *The Review of Economic Studies*, 2022, *89* (3), 1314–1334.

Drugowitsch, Jan, Valentin Wyart, Anne-Dominique Devauchelle, and Etienne Koechlin, "Computational precision of mental inference as critical source of human choice suboptimality," *Neuron*, 2016, *92* (6), 1398–1411.

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

Enke, Benjamin and Cassidy Shubatt, "Quantifying Lottery Choice Complexity," *Working Paper*, 2023.

— **and Thomas Graeber**, "Cognitive uncertainty in intertemporal choice," Technical Report, National Bureau of Economic Research 2021.

— **and** — , "Cognitive Uncertainty," *Quarterly Journal of Economics*, 2023.

Epper, Thomas, Ernst Fehr, Helga Fehr-Duda, Claus Thustrup Kreiner, David Dreyer Lassen, Søren Leth-Petersen, and Gregers Nytoft Rasmussen, "Time discounting and wealth inequality," *American Economic Review*, 2020, *110* (4), 1177–1205.

— **, Helga Fehr-Duda et al.**, "Risk in time: The intertwined nature of risk taking and time discounting," Technical Report, Working Paper at the University of St. Gallen 2019.

Ericson, Keith Marzilli and David Laibson, "Intertemporal choice," in "Handbook of Behavioral Economics: Applications and Foundations 1," Vol. 2, Elsevier, 2019, pp. 1–67.

— **, John Myles White, David Laibson, and Jonathan D Cohen**, "Money earlier or later? Simple heuristics explain intertemporal choices better than delay discounting does," *Psychological science*, 2015, *26* (6), 826–833.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde, "Global evidence on economic preferences," *The Quarterly Journal of Economics*, 2018, *133* (4), 1645–1692.

Fisher, Geoffrey, "Intertemporal choices are causally influenced by fluctuations in visual attention," *Management Science*, 2021, *67* (8), 4961–4981.

Frydman, Cary and Lawrence J Jin, "On the Source and Instability of Probability Weighting," *Working Paper*, 2023.

Fudenberg, Drew and David K Levine, "A dual-self model of impulse control," *American economic review*, 2006, *96* (5), 1449–1476.

Gabaix, Xavier and David Laibson, "Myopia and discounting," Technical Report, National bureau of economic research 2022.

Gershman, Samuel J and Rahul Bhui, "Rationally inattentive intertemporal choice," *Nature Communications*, 2020, *11*.

Gillen, Ben, Erik Snowberg, and Leeat Yariv, "Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study," *Journal of Political Economy*, 2019.

Giné, Xavier, Jessica Goldberg, Dan Silverman, and Dean Yang, "Revising commitments: Field evidence on the adjustment of prior choices," *The Economic Journal*, 2018, *128* (608), 159–188.

Glätzle-Rützler, Daniela, Philipp Lergetporer, and Matthias Sutter, "Collective intertemporal decisions and heterogeneity in groups," *Games and Economic Behavior*, 2021, *130*, 131–147.

Goda, Gopi Shah, Matthew R Levy, Colleen Flaherty Manchester, Aaron Sojourner, and Joshua Tasoff, "The role of time preferences and exponential-growth bias in retirement savings," Technical Report, National Bureau of Economic Research 2015.

Gul, Faruk and Wolfgang Pesendorfer, "Temptation and self-control," *Econometrica*, 2001, *69* (6), 1403–1435.

Halevy, Yoram, "Some comments on the use of monetary and primary rewards in the measurement of time preferences," *Unpublished manuscript*, 2014.

Hanushek, Eric A, Lavinia Kinne, Philipp Lergetporer, and Ludger Woessmann, "Patience, risk-taking, and human capital investment across countries," *The Economic Journal*, 2022, *132* (646), 2290–2307.

He, Lisheng, Russell Golman, and Sudeep Bhatia, "Variable time preference," *Cognitive psychology*, 2019, *111*, 53–79.

Herz, Holger, Martin Huber, Tjaša Maillard-Bjedov, and Svitlana Tyahlo, "Time preferences across language groups: Evidence on intertemporal choices from the Swiss language border," *The Economic Journal*, 2021, *131* (639), 2920–2954.

Imas, Alex, Michael Kuhn, and Vera Mironova, "Waiting to choose: The Role of Deliberation in Intertemporal Choice," *American Economic Journal: Microeconomics*, 2021.

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

John, Anett and Kate Orkin, "Can simple psychological interventions increase preventive health investment?," *Journal of the European Economic Association*, 2022, *20* (3), 1001–1047.

Kable, Joseph W and Paul W Glimcher, "An "as soon as possible" effect in human intertemporal decision making: behavioral evidence and neural mechanisms," *Journal of neurophysiology*, 2010, *103* (5), 2513–2531.

Khaw, Mel Win, Ziang Li, and Michael Woodford, "Cognitive imprecision and small-stakes risk aversion," *The review of economic studies*, 2021, *88* (4), 1979–2013.

— , — , **and** — , "Cognitive Imprecision and Stake-Dependent Risk Attitudes," Technical Report, National Bureau of Economic Research 2022.

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

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

Lu, Jay and Kota Saito, "Random intertemporal choice," *Journal of Economic Theory*, 2018, *177*, 780–815.

Mahajan, Aprajit and Alessandro Tarozzi, "Time inconsistency, expectations and technology adoption: The case of insecticide treated nets," 2012.

Martinez, Seung-Keun, Stephan Meier, and Charles Sprenger, "Procrastination in the field: Evidence from tax filing," *Journal of the European Economic Association*, 2023, *21* (3), 1119–1153.

Meier, Stephan and Charles D Sprenger, "Temporal stability of time preferences," *Review of Economics and Statistics*, 2015, *97* (2), 273–286.

— **and Charles Sprenger**, "Present-Biased Preferences and Credit Card Borrowing," *American Economic Journal: Applied Economics*, 2010, *2* (1), 193–210.

Nielsen, Kirby and John Rehbeck, "When Choices are mistakes," *Available at SSRN 3481381*, 2020.

O'Donoghue, Ted and Matthew Rabin, "Doing it now or later," *American Economic Review*, 1999, *89* (1), 103–124.

Oprea, Ryan, "What Makes a Rule Complex?," *American Economic Review*, 2020, *110* (12), 3913–3951.

— , "Simplicity Equivalents," *Working Paper*, 2022.

Read, Daniel, "Is time-discounting hyperbolic or subadditive?," *Journal of risk and uncertainty*, 2001, *23* (1), 5–32.

— **, Christopher Y Olivola, and David J Hardisty**, "The value of nothing: Asymmetric attention to opportunity costs drives intertemporal decision making," *Management Science*, 2017, *63* (12), 4277–4297.

— **, Shane Frederick, and Marc Scholten**, "DRIFT: an analysis of outcome framing in intertemporal choice.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2013, *39* (2), 573.

—, —, **Burcu Orsel, and Juwaria Rahman**, "Four score and seven years from now: The date/delay effect in temporal discounting," *Management Science*, 2005, *51* (9), 1326–1335.

Regenwetter, Michel, Daniel R Cavagnaro, Anna Popova, Ying Guo, Chris Zwilling, Shiau Hong Lim, and Jeffrey R Stevens, "Heterogeneity and parsimony in intertemporal choice.," *Decision*, 2018, *5* (2), 63.

Rubinstein, Ariel, ""Economics and psychology"? The case of hyperbolic discounting," *International Economic Review*, 2003, *44* (4), 1207–1216.

Simon, Herbert A, "A behavioral model of rational choice," *The quarterly journal of economics*, 1955, pp. 99–118.

Stango, Victor and Jonathan Zinman, "Exponential growth bias and household finance," *The Journal of Finance*, 2009, *64* (6), 2807–2849.

Sunde, Uwe, Thomas Dohmen, Benjamin Enke, Armin Falk, David Huffman, and Gerrit Meyerheim, "Patience and comparative development," *a*, 2022, *89* (5), 2806–2840.

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

Thaler, Richard, "Some empirical evidence on dynamic inconsistency," *Economics letters*, 1981, *8* (3), 201–207.

Vieider, F, "Decisions under uncertainty as bayesian inference," Technical Report, Working Paper WP 21/1022 of Faculty of Economics and Business Administration 2021.

Vieider, Ferdinand M, "Noisy coding of time and reward discounting," Technical Report, Ghent University, Faculty of Economics and Business Administration 2021.

Weitzman, Martin L, "Gamma discounting," *American Economic Review*, 2001, *91* (1), 260–271.

Woodford, Michael, "Modeling imprecision in perception, valuation, and choice," *Annual Review of Economics*, 2020, *12*, 579–601.

Zauberman, Gal, B Kyu Kim, Selin A Malkoc, and James R Bettman, "Discounting time and time discounting: Subjective time perception and intertemporal preferences," *Journal of Marketing Research*, 2009, *46* (4), 543–556.

Online Appendix

A Additional Figures

	Option A	Option B
1	$42.00 in 12 months	$2.00 now
2	$42.00 in 12 months	$4.00 now
3	$42.00 in 12 months	$6.00 now
4	$42.00 in 12 months	$8.00 now
5	$42.00 in 12 months	$10.00 now
6	$42.00 in 12 months	$12.00 now
7	$42.00 in 12 months	$14.00 now
8	$42.00 in 12 months	$16.00 now
9	$42.00 in 12 months	$18.00 now
10	$42.00 in 12 months	$20.00 now
11	$42.00 in 12 months	$22.00 now
12	$42.00 in 12 months	$24.00 now
13	$42.00 in 12 months	$26.00 now
14	$42.00 in 12 months	$28.00 now
15	$42.00 in 12 months	$30.00 now
16	$42.00 in 12 months	$32.00 now
17	$42.00 in 12 months	$34.00 now
18	$42.00 in 12 months	$36.00 now
19	$42.00 in 12 months	$38.00 now
20	$42.00 in 12 months	$40.00 now
21	$42.00 in 12 months	$42.00 now

a) *Delay* treatment

	Option A	Option B
1	$42.00 shrunk 12 times	$2.00
2	$42.00 shrunk 12 times	$4.00
3	$42.00 shrunk 12 times	$6.00
4	$42.00 shrunk 12 times	$8.00
5	$42.00 shrunk 12 times	$10.00
6	$42.00 shrunk 12 times	$12.00
7	$42.00 shrunk 12 times	$14.00
8	$42.00 shrunk 12 times	$16.00
9	$42.00 shrunk 12 times	$18.00
10	$42.00 shrunk 12 times	$20.00
11	$42.00 shrunk 12 times	$22.00
12	$42.00 shrunk 12 times	$24.00
13	$42.00 shrunk 12 times	$26.00
14	$42.00 shrunk 12 times	$28.00
15	$42.00 shrunk 12 times	$30.00
16	$42.00 shrunk 12 times	$32.00
17	$42.00 shrunk 12 times	$34.00
18	$42.00 shrunk 12 times	$36.00
19	$42.00 shrunk 12 times	$38.00
20	$42.00 shrunk 12 times	$40.00
21	$42.00 shrunk 12 times	$42.00

b) *Mirror* treatment

Figure A.1: Screenshots from the experimental software.

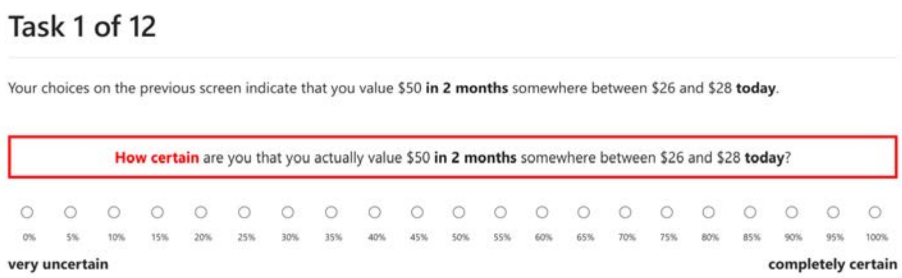

Figure A.2: Screenshot of an example cognitive uncertainty elicitation screen in *Delay-M*

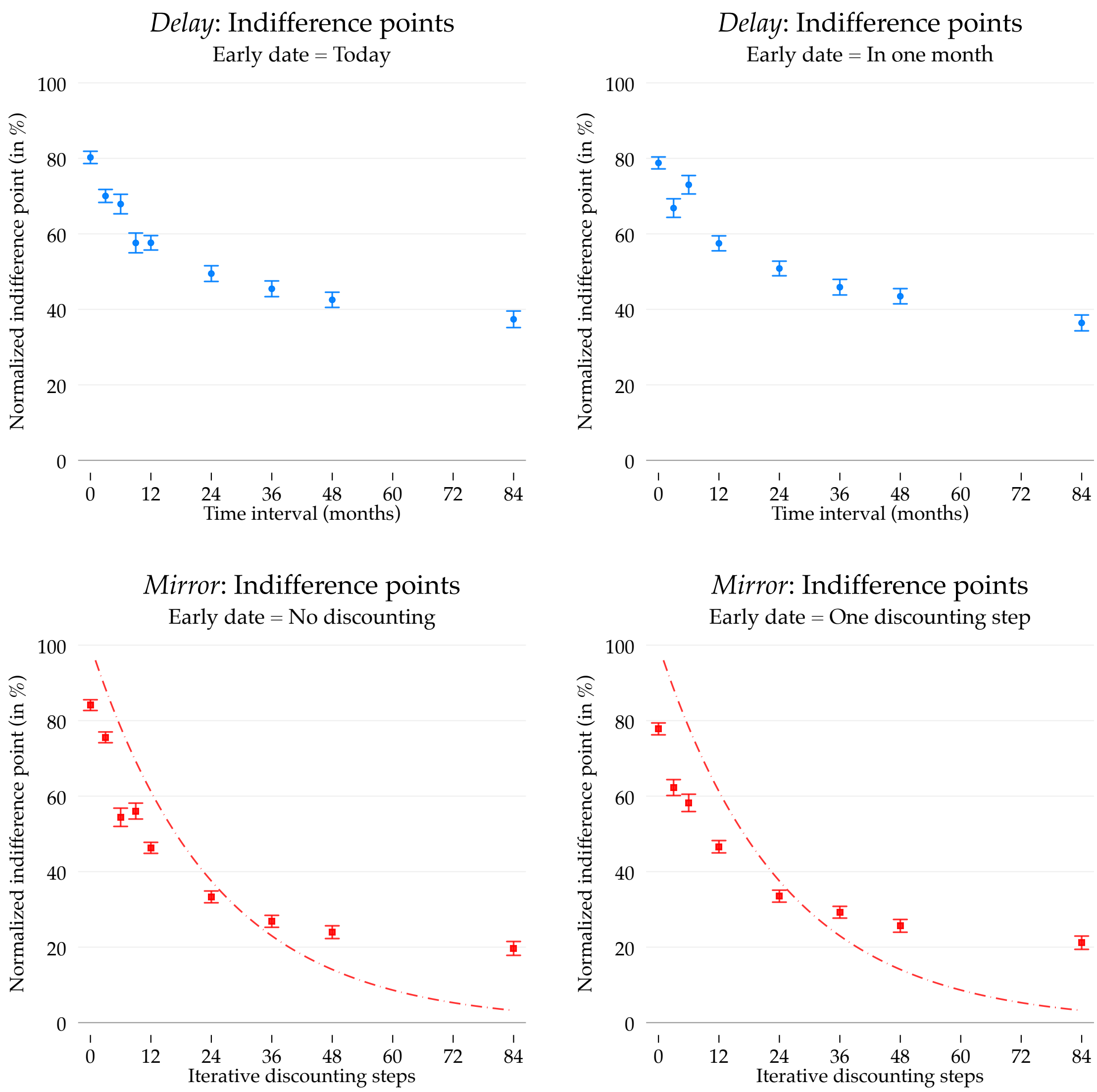

Figure A.3: Average normalized indifference points by time interval (*Delay*) or number of iterations (*Mirror*). Top panels show *Delay* treatment (4,572 decisions from 254 participants). Bottom panels show *Mirror* treatment (4,428 decisions from 246 participants). In the *Mirror* panels, the dashed line represents payoff-maximizing decisions. Sample splits according to whether the earlier payment occurs today/requires no discounting. The time interval in months and the number of iterations are rounded to the nearest multiple of three. Whiskers show standard error bars, computed based on clustering at the subject level.

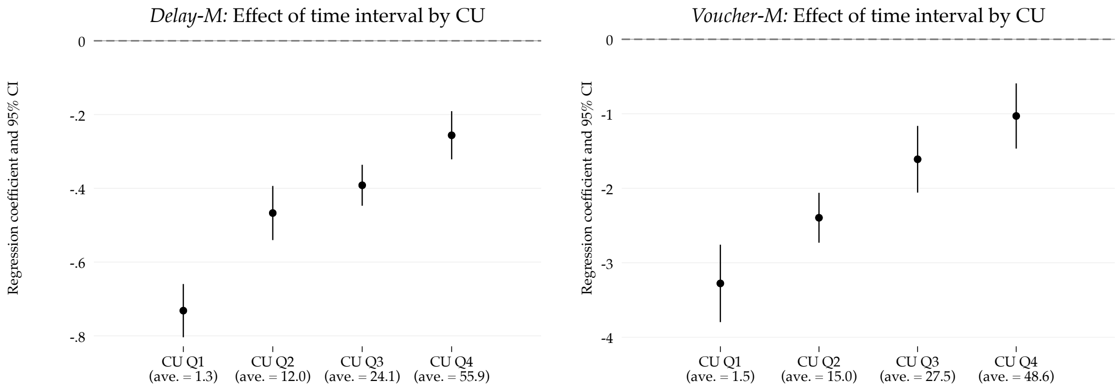

Figure A.4: Coefficients from regressions of normalized indifference points on time interval, split by CU quartiles; left: *Delay-M*; right: *Voucher-M*.

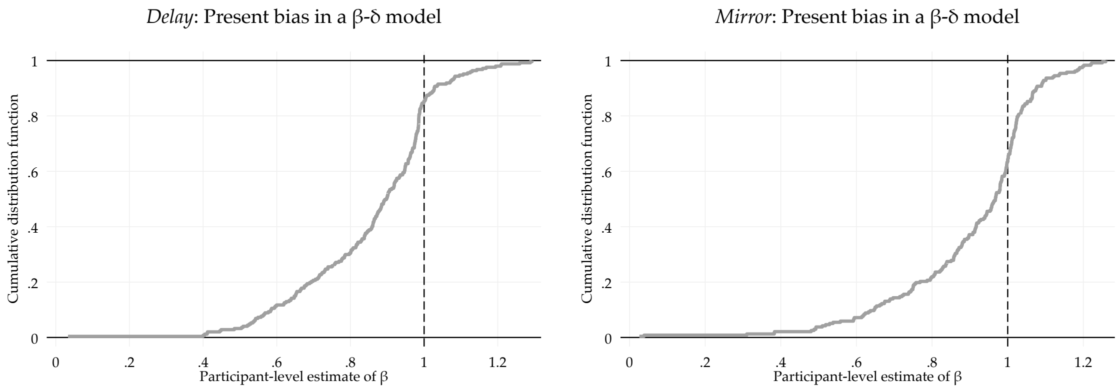

Figure A.5: Empirical CDFs of individual-level estimates of a $\beta - \delta$ model (eq. (2)) in *Delay* ($N = 254$) and *Mirror* ($N = 246$), using first-assigned treatment only. Non-linear least squares estimation based on 18 decisions from each individual. For ease of readability we exclude subjects with $\hat{\beta} > 1.3$.

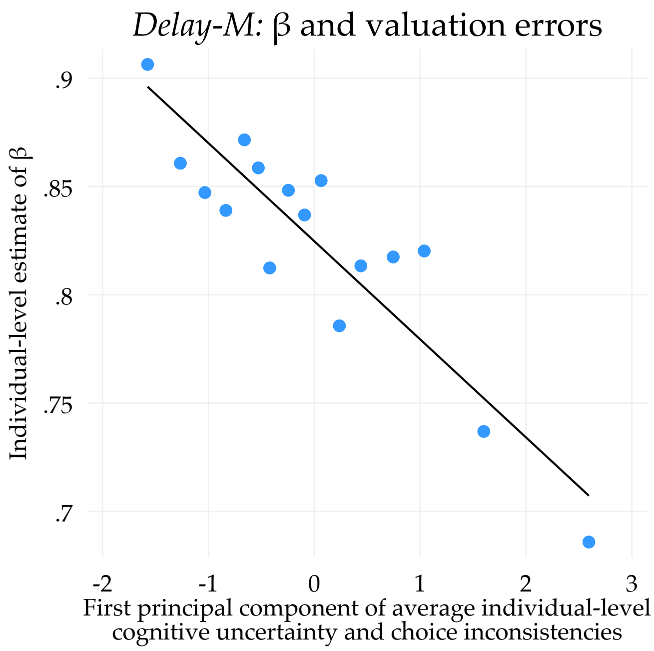

Figure A.6: Binned scatter plot of individual-level $\hat{\beta}_{ST}$ in eq. (2) against first principal component of subject-level average CU and choice inconsistencies. In this figure the individual-level estimate of β is derived taking into account utility curvature, which is separately estimated at the population level based on lottery choice lists. Based on 643 subjects in *Delay-M*. The figure excludes subjects for whom we estimate $\hat{\beta} > 2$.

B Additional Tables

Table A.1: Overview of recent intertemporal choice literature using money as a reward

Title	Authors	Year	Journal	Description
a. Lab-in-the-field studies with patience as an outcome				
Poverty and economic decision-making: Evidence from changes in financial resources at payday	Carvalho et al.	2016	*The American Economic Review*	Studies the effect of payday proximity on intertemporal choices in a survey
Fostering patience in the classroom: Results from randomized educational intervention	Alan and Ertac	2018	*The Journal of Political Economy*	Randomized educational intervention on children decreases impatience
Revising commitments: Field evidence on the adjustment of prior choices	Giné et al.	2018	*The Economic Journal*	Artefactual field experiment on revisions of prior choices regarding future income receipts
Can simple psychological interventions increase preventive health investment?	John and Orkin	2022	*Journal of the European Economic Association*	Two light-touch psychological interventions such as planning prompts affect patience
b. Lab-in-the-field studies with patience as a predictor				
Why do defaults affect behavior? Experimental evidence from Afghanistan	Blumenstock et al.	2018	*The American Economic Review*	Experimental measure of present bias predicts whether default effect impact behavior
Discount rates of children and high school graduation	Castillo et al.	2019	*The Economic Journal*	Experimental measure of patience predicts whether children graduate from high school
Time discounting and wealth inequality	Epper et al.	2020	*The American Economic Review*	Experimental measures of impatience predict wealth
Procrastination in the field: Evidence from tax filing	Martinez et al.	2023	*Journal of the European Economic Association*	Studies present-biased procrastination in tax-filing behavior
Time preferences and food choice	Brownback et al.	2023	*NBER Working Paper*	Incentivized time preference measures predict healthy food choice
c. Laboratory and online studies of patience				
Measuring discounting without measuring utility	Attema et al.	2016	*The American Economic Review*	Introduces a new method to measure temporal discounting of money that does not rely on assumptions about utility
The value of nothing: Asymmetric attention to opportunity costs drives intertemporal decision making	Read et al.	2017	*Management Science*	Studies the role of the salience of opportunity costs for measurement of time preferences
Time matters less when outcomes differ: Unimodal vs. cross-modal comparisons in intertemporal choice	Cubitt et al.	2018	*Management Science*	People are more aversive to delay when trading off delays for the same good (e.g., money earlier versus later) as opposed to delays for different goods
How long is a minute?	Brocas et al.	2018	*Games and Economic Behavior*	People who overestimate objective time intervals are less willing to delay gratification
Arbitrage or narrow bracketing? On using money to measure intertemporal preferences	Andreoni et al.	2018	*NBER Working Paper*	Suggests money is a valid reward; finds evidence for narrow bracketing and against arbitrage reasoning
Intertemporal choices are causally influenced by fluctuations in visual attention	Fisher	2021	*Management Science*	Intertemporal decisions are strongly shaped by allocation of visual attention to different choice elements
Collective intertemporal decisions and heterogeneity in groups	Glätzle-Rützler et al.	2021	*Games and Economic Behavior*	Three-person groups behave more patiently than individuals
Time preferences across language groups: Evidence on intertemporal choices from the Swiss language border	Herz et al.	2021	*The Economic Journal*	Studies differences in discounting behavior between French and German speakers
Concentration bias in intertemporal choice	Dertwinkel-Kalt et al.	2022	*The Review of Economic Studies*	In intertemporal tradeoffs, people overweight advantages that are concentrated in time
d. Large surveys				
Global evidence on economic preferences	Falk et al.	2018	*The Quarterly Journal of Economics*	Documents global variation and correlates of patience and other economic preferences
Patience and comparative development	Sunde et al.	2022	*The Review of Economic Studies*	Studies relationship between patience and comparative development
Patience, risk-taking, and human capital investment across countries	Hanushek et al.	2022	*The Economic Journal*	Patience predict cross-country variation in human capital investment decisions
Econographics	Chapman et al.	2023	*Journal of Political Economy Microeconomics*	Studies relationship between discounting and other behavioral regularities

Notes. This table lists papers reporting measurements of discounting behavior that use money as a reward. We include publications in the Top 5 economics journals and selected field journals as well as working papers. We restrict the list to papers dated 2016 or later so that postdate the seminal contribution of Augenblick et al. (2015) introducing real-effort measures of discounting behavior.

Table A.2: Anomalies in *Delay* and *Mirror*

	Dependent variable: Implied annual impatience (in %)					
Phenomenon:	Decreasing impatience		Subadditivity		Front-end delay	
Treatment:	*Delay*	*Mirror*	*Delay*	*Mirror*	*Delay*	*Mirror*
	(1)	(2)	(3)	(4)	(5)	(6)
Time interval / number of discounting steps (in years)	-5.76***	-5.14***				
	(0.25)	(0.26)				
1 if one long interval			-7.57***	-9.93***		
			(1.38)	(1.16)		
1 if front end delay					-4.24**	3.79**
					(1.85)	(1.69)
Payment amount FE	Yes	Yes	Yes	Yes	Yes	Yes
Subadditivity set FE	No	No	Yes	Yes	Yes	Yes
Observations	4572	4428	508	492	508	492
R^2	0.17	0.19	0.09	0.06	0.07	0.02

Notes. OLS estimates, robust standard errors (in parentheses) are clustered at the subject level. In columns (1) and (2), the sample consists of all decisions in the respective treatment. In columns (3) and (4), the sample consists of two observations per subject: their implied annual discounting over the long (composite) interval and their implied discounting over the two shorter intervals that have a subadditivity structure. In columns (5) and (6), the sample includes those two decisions per subject that have a front-end delay structure. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: Anomalies in *Delay* and *Mirror*, pooling first-assigned and second-assigned treamtents

	Dependent variable: Implied annual impatience (in %)					
Phenomenon:	Decreasing impatience		Subadditivity		Front-end delay	
Treatment:	*Delay*	*Mirror*	*Delay*	*Mirror*	*Delay*	*Mirror*
	(1)	(2)	(3)	(4)	(5)	(6)
Time interval / number of discounting steps (in years)	-6.11***	-4.68***				
	(0.18)	(0.18)				
1 if one long interval			-8.57***	-10.0***		
			(0.89)	(0.81)		
1 if front end delay					-4.59***	5.41***
					(1.28)	(1.21)
Payment amount FE	Yes	Yes	Yes	Yes	Yes	Yes
Subadditivity set FE	No	No	Yes	Yes	Yes	Yes
Observations	9000	8999	1000	1000	1000	1000
R^2	0.19	0.17	0.05	0.06	0.03	0.02

Notes. OLS estimates, robust standard errors (in parentheses) are clustered at the subject level. In columns (1) and (2), the sample consists of all decisions in the respective treatment. In columns (3) and (4), the sample consists of two observations per subject: their implied annual discounting over the long (composite) interval and their implied discounting over the two shorter intervals. In columns (5) and (6), the sample includes those two decisions per subject that have a front-end delay structure. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Short-run and decreasing impatience as functions of CU and choice inconsistency

	Dependent variable: Implied annual impatience (in %)					
Dataset:	*Delay-M*			*Voucher-M*		
Phenomenon:	SR imp. (≤ 1m)	Decreasing impat.		SR imp. (≤ 1m)	Decreasing impat.	
	(1)	(2)	(3)	(4)	(5)	(6)
Time interval		-7.61***	-3.56***		-29.1***	-20.6***
		(0.30)	(0.54)		(2.49)	(6.12)
Cognitive uncertainty	0.42***		0.24***	0.61***		0.57***
	(0.11)		(0.05)	(0.10)		(0.08)
Inconsistent decision	25.2***		11.9***	16.7***		19.6***
	(4.59)		(2.38)	(3.30)		(2.94)
Time interval × Cognitive uncertainty			-0.061***			-0.47***
			(0.01)			(0.13)
Time interval × Inconsistent decision			-4.95***			-13.0**
			(0.61)			(6.44)
Payment amount FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	344	2580	2580	766	2000	2000
R^2	0.24	0.20	0.24	0.17	0.06	0.19

Notes. OLS estimates, robust standard errors (in parentheses) are clustered at the subject level. In columns (1) and (4), the sample is restricted to time intervals of at most one month. To make the samples comparable across columns, we restrict attention to decisions for which the choice inconsistency variable is available. Time interval is in years. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Decreasing impatience in *Delay-M* and within-subject variation of cognitive uncertainty

	Dependent variable: Implied annual impatience (in %)	
Dataset:	*Delay-M*	
Phenomenon:	Decreasing impatience	
	(1)	(2)
Time interval	-6.87***	-5.46***
	(0.18)	(0.26)
Cognitive uncertainty (standard. within subject)		0.28***
		(0.04)
Time interval × Cognitive uncertainty (standard. within subject)		-0.068***
		(0.01)
Payment amount FE	Yes	Yes
Observations	7740	7740
R^2	0.16	0.18

Notes. OLS estimates, robust standard errors (in parentheses) are clustered at the subject level. In these regressions, the measure of cognitive uncertainty was standardized at the subject level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: Choice inconsistencies in *Mirror*

	Dependent variable: Implied annual impatience (in %)	
Phenomenon:	Short-run impatience	Decreasing impatience
	(1)	(2)
Inconsistent decision	17.8***	12.7***
	(3.86)	(3.90)
Number of discounting steps (in years)		-2.64***
		(0.51)
Number of discounting steps (in years) × Inconsistent decision		-3.07***
		(0.53)
Payment amount FE	Yes	Yes
Observations	417	3408
R^2	0.09	0.21

Notes. OLS estimates, robust standard errors (in parentheses) are clustered at the subject level. Regressions include sets of repeated decisions shown to a subject. Column (1) includes decisions with one discounting iteration only, column (2) includes decisions involving any number of iterations. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: Anomalies in the *Delay-M* vs. the *Complex* treatments

Phenomenon:	Manipulation check		Decr. imp.	Subadd.	Front-end
	Dependent variable:				
	CU	Inconsistent	Implied annual impatience (in %)		
	(1)	(2)	(3)	(4)	(5)
Complex treatments	13.6***	0.065**	0.99	3.07	-0.59
	(1.40)	(0.03)	(1.89)	(2.20)	(2.35)
Time interval			-6.88***		
			(0.18)		
Time interval × Complex treatments			-1.92***		
			(0.34)		
1 if one long interval				-8.58***	
				(0.63)	
1 if one long interval × Complex treatments				-7.99***	
				(1.30)	
1 if front-end delay					-3.06***
					(0.99)
1 if front-end delay × Complex treatments					5.32***
					(1.86)
Payment amount FE	Yes	Yes	Yes	Yes	Yes
Subadditivity set FE	No	No	No	Yes	Yes
Observations	11364	3788	11364	2818	3465
R^2	0.06	0.01	0.18	0.04	0.01

Notes. OLS estimates, robust standard errors (in parentheses) are clustered at the subject level. In columns (1) and (3), the sample consists of all decisions in the *Delay-M* and *Complex* treatments. In column (2), the sample includes all sets of repeated decisions shown to a subject. In column (4), the sample consists of two observations per subject: their implied annual discounting over the long (composite) interval and their implied discounting over the two shorter intervals. In column (5), the sample includes those two decisions per subject that have a front-end delay structure. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8: Population-level estimates of $\beta - \delta$ model

	Delay & Mirror		*Delay-M*					*Complex*	*Voucher-M*				
	Delay (1)	Mirror (2)	All (3)	CU=0 (4)	CU>0 (5)	Incons.=0 (6)	Incons.>0 (7)	All (8)	All (9)	CU=0 (10)	CU>0 (11)	Incons.=0 (12)	Incons.>0 (13)
$\hat{\beta}$	.774	.846	.76	.872	.721	.822	.75	.72	.882	.953	.854	.957	.865
$\hat{\delta}$	.982	.96	.978	.973	.98	.983	.977	.989	.942	.955	.941	.968	.936

Notes. Population-level estimates of a $\beta - \delta$ model (eq. (2)). Columns (1) and (2) use the first-assigned treatment only, based on $N = 254$ subjects in *Delay* and $N = 246$ subjects in *Mirror*. Columns (3), (8) and (9) include all subjects in the respective treatments: $N = 645$ in *Delay-M*, $N = 302$ in *Complex* and $N = 500$ in *Voucher-M*. All other columns are based on sample splits of the corresponding treatments. Non-linear least squares estimates.

C What Makes Intertemporal Choice Difficult?

This Appendix tentatively investigates *what* it is about intertemporal choice that makes it complex, and therefore vulnerable to noisy or heuristic decision-making. One possibility, ex ante, is that complexity is a consequence of the fact that it is difficult to introspectively evaluate or calculate one's own time preferences (e.g., one's discount factor). Similarly, another ex ante possibility is that complexity is an outgrowth of the difficulty of integrating one's risk and time preferences to inform choice.

Results from our *Mirror* treatment (in which time preferences are clearly induced and risk and time preferences needn't be integrated), suggest an alternative possibility: that the complexity of intertemporal choice is instead a direct outgrowth of the costs and difficulties of iteratively discounting rewards, which requires an intensive type of recursive reasoning. If true, we would expect the number of required steps of discounting / a longer time delay to be associated with more pronounced valuation errors.[16]

To examine this, re-reconsider equation (2). Rearranging, taking logs and adding a mean-zero noise term yields that a subject's observed indifference point in our experiments can be expressed as

$$ln\left(\frac{x_1}{x_2}\right) = ln(\beta_{t_1=0}) + \Delta t \cdot ln(\delta) + \varepsilon. \tag{3}$$

where the first term on the right-hand side collapses to zero if $\beta = 1$. Importantly, our hypothesis that valuation errors increase in the delay implies that $Var(\varepsilon)$ should not be constant but, instead, heteroscedastic and increasing in the delay. Because in equation (3) a subject's log normalized indifference point is a linear function of the delay, the equation can be estimated using simple OLS. We run this regression and then inspect the variance of the regression residuals.

The top left panel of Figure A.7 shows the results for treatment *Delay*. We find that the variance of the regression residuals indeed strongly increases in the length of the delay. A different way of saying this is that the variance of subjects' normalized indifference points strongly increases in the delay.

The top right panel shows an analogous plot for treatment *Mirror*, where the x-axis now represents the required number of steps of discounting. Again, we see strong evidence of heteroscedasticity, in line with the hypothesis that valuation errors become more pronounced as the number of discounting steps increases.

In a standard exponential discounting model with preference heterogeneity, the regression residuals or the variance of log indifference points *should* increase in the delay[17]

[16]Some models of complexity and intertemporal discounting directly consider this possibility: Gabaix and Laibson (2022) model a decision maker whose degree of cognitive noisiness increases in the delay.

[17]With exponential discounting and linear utility, $Var[ln(x_1/x_2)] = (\Delta t)^2 Var[ln(\delta)] + Var(\epsilon)$.

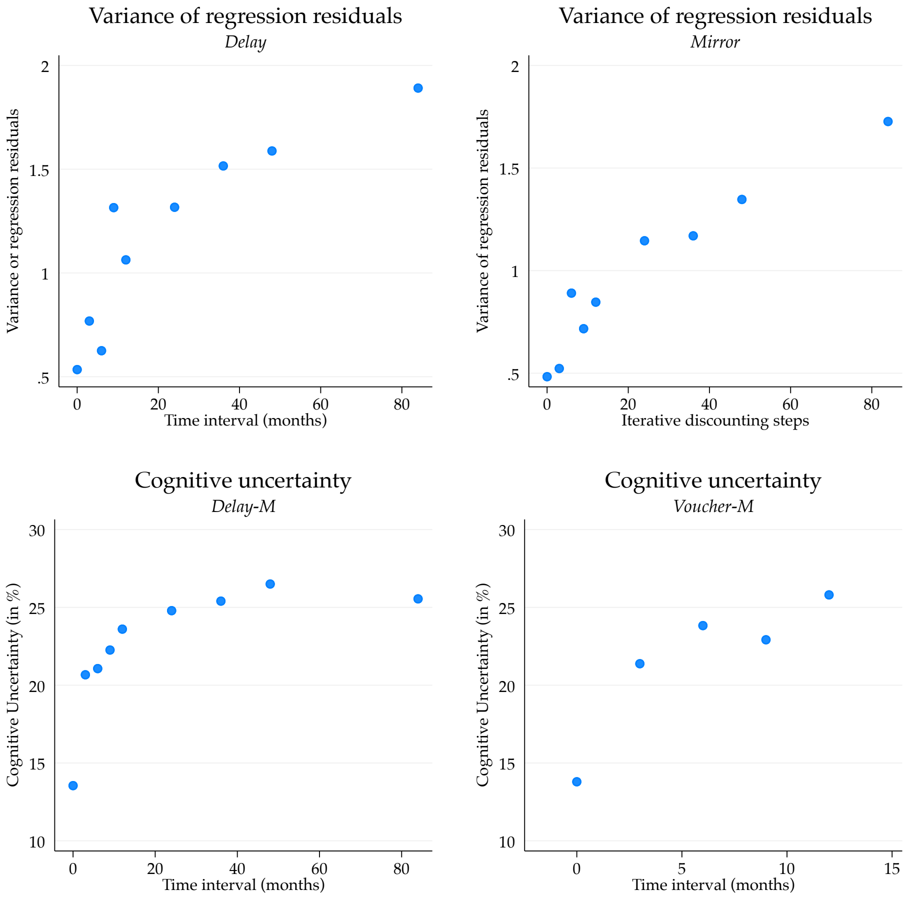

Figure A.7: Noisiness as a function of the delay. Top panels show the variance of the regression residuals of eq. (3) in *Delay* and *Mirror*. Bottom panels show average cognitive uncertainty in *Delay Noise* and *Voucher Noise*. In all panels, delays are rounded to the nearest multiple of three.

However, in treatment *Mirror*, where the increase is almost equally strong, there is no preference heterogeneity available to rationalize the pattern because we experimentally induced the same discount factor for all subjects. In *Mirror*, this pattern must be driven by increasingly idiosyncratic responses to complexity as the number of steps of discounting increases. The fact that that the pattern (including magnitudes) is almost identical in *Delay* suggests the same complexity-based explanation likely applies there as well. Moreover, recall that decisions in *Delay* and *Mirror* are highly correlated within subject, providing further suggestive evidence that the increase in the variance of decisions has the same origin, which cannot be heterogeneity in discount factors.

The bottom panels of Figure A.7 provide additional evidence in suport of this claim. We plot subjects' cognitive uncertainty as a function of the delay in treatments *Delay Noise* and *Voucher Noise*.[18] In both treatments, people report being much more uncertain

[18]Analyzing how choice inconsistencies vary with the delay is confounded by the relationship between choice inconsistency and the "extremity" of the intertemporal decision problem. In all treatments, we find that subjects exhibit less inconsistency when the delay is either very short or very long, in large part

about which decision to take as the delay gets longer. Going from very short delays of less than one month to delays of seven years, CU more than doubles. This increase is concave, with CU barely increasing for delays longer than 1–2 years (recall that in *Voucher Noise* the longest delay is one year).

Taken together, multiple streams of evidence suggest that the difficulty of decision-making increases in the length of the delay / the number of discounting steps required. This, when combined with the appearance of anomalies in the atemporal mirrors (where there is little to drive complexity except the difficulty of iterative discounting), is indicative that an important source of complexity in intertemporal decision-making is the cognitive act of iteratively discounting future rewards.

Of course, the insight that complexity increases in the number of cognitive steps required to discount does not imply that complexity is zero for very short delays. For instance, as Figure A.7 shows, there is substantial CU even for delays of one month and less, consistent with people exhibiting noise-driven extreme short-run impatience.

because in these decision problems a large share of subjects make boundary choices that artifically make them look perfectly consistent.

D Experimental Instructions

D.1 Instructions for *Delay & Mirror* Experiment

D.1.1 First-assigned treatment: *Delay*

Delayed Choices

In this part of the study you will **choose between various hypothetical payments, which pay different amounts at different points in time**. An example decision is between the following two hypothetical payments.

Option A	Option B
$100.00 in 3 months	$90.00 now

In this example we are asking you (hypothetically) would you rather be paid $100 in three months (Option A) or $90 right now (Option B).

For all hypothetical payments in this study, please treat them as if you know you will receive them with certainty, even if they are delayed. That is, please assume there is no risk that you wouldn't actually get paid. Further, assume all payments were made by leaving a check in your mailbox which you can cash at the specified date.

For this part of the experiment, there are no right wrong answers, because how much you like an option depends on your personal taste. Just try your best to think hard about what you'd really prefer.

The Choice List

On your decision screen, you will be asked to choose which of two payment options you prefer. You will see choice lists such as the one below, where **each row is a separate choice**.

In every list, the left-hand option (Option A) is a delayed payment that is identical in all rows. The right-hand side option (Option B) is a payment with an *earlier payment date than Option A*. The earlier, right-hand side payment increases as you go down the list. To make a choice just click on the option you prefer for each choice (i.e. for each row), highlighting your choice yellow. **An effective way to complete these choice lists is to determine in which row you like to switch from preferring Option A to preferring Option B.**

	Option A	Option B
1	$40.00 in 3 months	$2.00 now
2	$40.00 in 3 months	$4.00 now
3	$40.00 in 3 months	$6.00 now
4	$40.00 in 3 months	$8.00 now
5	$40.00 in 3 months	$10.00 now
6	$40.00 in 3 months	$12.00 now
7	$40.00 in 3 months	$14.00 now
8	$40.00 in 3 months	$16.00 now
9	$40.00 in 3 months	$18.00 now
10	$40.00 in 3 months	$20.00 now
11	$40.00 in 3 months	$22.00 now
12	$40.00 in 3 months	$24.00 now
13	$40.00 in 3 months	$26.00 now
14	$40.00 in 3 months	$28.00 now
15	$40.00 in 3 months	$30.00 now
16	$40.00 in 3 months	$32.00 now
17	$40.00 in 3 months	$34.00 now
18	$40.00 in 3 months	$36.00 now
19	$40.00 in 3 months	$38.00 now
20	$40.00 in 3 months	$40.00 now

Auto-completion: The table auto-completes your choices so you don't have to click through all of the rows. You do not have to start at the top of the table. If you select Option A in any one row, we assume you will also prefer Option A in all rows *above* that row. If you select Option B in any one row, we assume that you will also prefer Option B in all rows *below* that row.

Next Part

In the next part of the experiment, we are going to have you make a **very different kind of decision**, also using choice lists.

Instead of making hypothetical decision about money paid out at various points in time (as in Part 1) **we will have you make real money decisions paid as a bonus today.** Specifically, we will ask you to choose between monetary amounts that are shrunk to varying degrees, using a choice list like the one you used in Part 1. The difference is, we will really pay some of you these amounts today!

D.1.2 First-assigned treatment: *Mirror*

Shrunk Choices

In this part of the study you will **choose between various payments (actually paid to you today), which will first be shrunk (reduced in value) some number of times.** An example decision is between the following two payments.

Option A	Option B
\$100.00 shrunk 3 times	\$90.00

Each time a payment is shrunk (as in Option A), its dollar value falls by 4% meaning it shrinks to only 96% of the dollar value from the previous step. For example

- If \$100 is shrunk only 1 time, we would pay you 96% of \$100 or \$96.
- If \$100 is shrunk in only 2 time, we would pay you 96% of 96% of \$100 or \$92.16
- If \$100 is shrunk in only 3 time, we would pay you 96% of 96% of 96% of \$100 or \$88.47

And so on. So, in the example, if you chose Option A (\$100 shrunk 3 times), you would earn \$88.47. On the other hand, Option B isn't shrunk at all so it just pays the \$90 shown (any time we don't mention shrinking for a payment, that means the payment is not shrunk at all).

At the end of the experiment we will randomly select 20% of participants to actually be paid their earnings as a bonus today from a randomly selected choice.

The Choice List

On your decision screen, you will be asked to choose which of two payment options you prefer. You will see choice lists such as the one below, where **each row is a separate choice**.

In every list, the left-hand option (Option A) is a delayed payment that is identical in all rows. The right-hand side option (Option B) is a payment with an *earlier payment date than Option A*. The earlier, right-hand side payment increases as you go down the list. To make a choice just click on the option you prefer for each choice (i.e. for each row), highlighting your choice yellow. **An effective way to complete these choice lists is to determine in which row you like to switch from preferring Option A to preferring Option B.**

	Option A	Option B
1	$40.00 in 3 months	$2.00 now
2	$40.00 in 3 months	$4.00 now
3	$40.00 in 3 months	$6.00 now
4	$40.00 in 3 months	$8.00 now
5	$40.00 in 3 months	$10.00 now
6	$40.00 in 3 months	$12.00 now
7	$40.00 in 3 months	$14.00 now
8	$40.00 in 3 months	$16.00 now
9	$40.00 in 3 months	$18.00 now
10	$40.00 in 3 months	$20.00 now
11	$40.00 in 3 months	$22.00 now
12	$40.00 in 3 months	$24.00 now
13	$40.00 in 3 months	$26.00 now
14	$40.00 in 3 months	$28.00 now
18	$40.00 in 3 months	$36.00 now
19	$40.00 in 3 months	$38.00 now
20	$40.00 in 3 months	$40.00 now

Auto-completion: The table auto-completes your choices so you don't have to click through all of the rows. You do not have to start at the top of the table. If you select Option A in any one row, we assume you will also prefer Option A in all rows *above* that row. If you select Option B in any one row, we assume that you will also prefer Option B in all rows *below* that row.

Next Part

In the next part of the experiment, we are going to have you make a **very different kind of decision**, also using choice lists.

Instead of making real money decisions about money shrunk to various degrees (as in Part 1) **we will have you make hypothetical money decisions about money amounts paid to you at various points in time.** Specifically, we will ask you to choose between monetary amounts paid sooner versus later, using a choice list like the one you used in Part 1. We won't actually pay you based on your choices in this part, but just want to understand when you'd hypothetically rather be paid various combinations of money.

D.2 Instructions for *Delay Noise*

Part 1 of this study: Instructions (1/3)

Please read these instructions carefully. There will be comprehension checks. If you fail these checks, you will immediately be excluded from the study and you will not receive the completion payment.

In this part of the study, you will **choose between various hypothetical payments, which pay different amounts of money at different points in time**. An example decision is between the following two hypothetical payments:

In 30 days: $ 40	OR	**Today: $ 12**

For all hypothetical payments in this study, please treat them as if you knew that you would receive them with certainty, even if they are delayed. That is, please assume that there is no risk that you wouldn't actually get paid. Further assume that all payments were made by leaving a check in your mailbox.

Throughout the experiment, there are no right or wrong answers, because how much you like an option depends on your personal taste. There will be two types of decision screens.

Decision screen 1

On decision screen 1, you will be asked to choose which of two payment options you prefer. You will see choice lists such as the one below, where each row is a separate choice. In every list, the left-hand side option (Option A) is a delayed payment that is identical in all rows. The right-hand side option (Option B) is a payment *with an earlier payment date than Option A*. The earlier, right-hand side payment increases as you go down the list. **An effective way to complete these choice lists is to determine in which row you would like to switch from preferring Option A to preferring Option B.**

Based on where you switch from Option A to Option B in this list, we assess which amount at the early payment date (Option B) you value as much as the amount specified at the later payment date (Option A). For example, in the choice list below, you would value $40 in 30 days somewhere between $12 and $14 today, because this is where switching occurs.

Option A			Option B
In 30 days: $40	●	○	Today: $2
	●	○	Today: $4
	●	○	Today: $6
	●	○	Today: $8
	●	○	Today: $10
	●	○	Today: $12
	○	●	Today: $14
	○	●	Today: $16
	○	●	Today: $18
	○	●	Today: $20
	○	●	Today: $22
	○	●	Today: $24
	○	●	Today: $26
	○	●	Today: $28
	○	●	Today: $30
	○	●	Today: $32
	○	●	Today: $34
	○	●	Today: $36
	○	●	Today: $38
	○	●	Today: $40

On the next page, you will see an example choice list, and you can practice making your selections.

Click "Next" to proceed to the example page.

Part 1 of this study: Instructions (2/3)

Auto-completion: The table auto-completes your choices so you don't have to click through all of the rows. You do not have to start at the top of the table. If you select Option A in any one row, we assume that you will also prefer Option A in all rows *above* that row. If you select Option B in any one row, we assume that you will also prefer Option B in all rows *below* that row.

Option A			Option B
	○	○	**Today: $2**
	○	○	**Today: $4**
	○	○	**Today: $6**
	○	○	**Today: $8**
	○	○	**Today: $10**
	○	○	**Today: $12**
	○	○	**Today: $14**
	○	○	**Today: $16**
	○	○	**Today: $18**
In 30 days: $40	○	○	**Today: $20**
	○	○	**Today: $22**
	○	○	**Today: $24**
	○	○	**Today: $26**
	○	○	**Today: $28**
	○	○	**Today: $30**
	○	○	**Today: $32**
	○	○	**Today: $34**
	○	○	**Today: $36**
	○	○	**Today: $38**
	○	○	**Today: $40**

Part 1 of this study: Instructions (3/3)

Decision screen 2

When you fill out a choice list, you may feel **uncertain about whether you prefer the left or right payment option**. On decision screen 2, we will ask you to select a button to indicate **how certain** you are how much money the larger later payment is worth to you in terms of dollars at the earlier payment date.

In answering this question, we ask you to assume that you would receive both payment options with certainty. We are interested in **your uncertainty about your own preferences regarding these payments**, not in your potential uncertainty about whether you would actually receive the money.

Example

Suppose that on the first decision screen you indicated that you valued $40 in 30 days somewhere between $12 and $14 today. Your second decision screen would look like this.

How certain are you that you actually value $40 **in 30 days** somewhere between $12 and $14 **today**?

○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
0%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%

very uncertain **completely certain**

Comprehension questions

The questions below test your understanding of the instructions.

Important: If you fail to answer any one of these questions correctly, you will not be allowed to participate in the study.

1. Which of the following statements is true?

   ○ In making my decisions, I am asked to assume that I will actually receive all payments as indicated, regardless of whether they take place now or in the future.

   ○ In making my decisions, I am asked to assume that it is less likely that I will actually receive payments that are meant to take place in the future.

   ○ In making my decisions, I am asked to assume that it is less likely that I will actually receive payments that are meant to take place now.

2. Suppose you are 80% certain that your decisions actually correspond to how much the different choice options are worth to you. Which button should you click in this case?

○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
0%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%

very uncertain **completely certain**

3. When we ask you how certain you are about how much different payments are worth to you at different points in time, then which type of uncertainty are we interested in?

   ○ Uncertainty about whether I would actually receive the payments.

   ○ Uncertainty about how much I value the payments, assuming that I know I would receive them with certainty.

D.3 Instructions for *Voucher Noise*

Part 1 of this study: Instructions (1/4)

Please read these instructions carefully. There will be comprehension checks. If you fail these checks, we will have to exclude you from the study and you will not receive the completion payment.

In this part of the study, you will **choose between different UberEats food delivery vouchers. These vouchers will vary along two dimensions**:

- **The vouchers will have different values**
- **The vouchers will be valid at different points in time**

How do the vouchers work?

Each voucher is valid for food delivery during a period of only seven days. A voucher can be used starting **from the indicated date**, and **it remains valid for exactly 7 days after** that date. Specifically, the vouchers work as follows:

- If you win a voucher, you will be informed about the voucher amount and the validity period on the last page of this study. You will then be asked to provide an email address associated with an UberEats account. The voucher will directly be credited to the corresponding UberEats account within the next 10 hours.
- However, the voucher amount **can only be spent during the validity period** of the voucher.
- Vouchers can be used to order from the entire range of restaurants, cafes, and bars that partner with UberEats in your area.
- You do not need to worry about forgetting the validity period: **UberEats will automatically send reminders** about your voucher 24 hours before the validity period starts and 24 hours before it ends.

What decisions will you be asked to make?

An example decision is between the following two vouchers:

Valid in 30 days: $40 Voucher	OR	**Valid today: $20 Voucher**

The left-hand side voucher carries an amount of $40 and can be spent in the 7-day period starting in 30 days from now. The right-hand side voucher is for an amount of only $20, but can be spent in the 7-day period starting immediately.

Throughout the experiment, there are no right or wrong answers because how much you like a voucher depends on your personal taste.

Part 1 of this study: Instructions (2/4)

Decision screen 1

On decision screen 1, you will be asked to choose which of two vouchers you prefer. You will see choice lists such as the one below, where each row is a separate choice. In every list, the left-hand side option (Voucher A) is a voucher that is identical in all rows. The right-hand side option (Voucher B) is a voucher *with an earlier validity period than Voucher A*. The amount associated with the earlier, right-hand side voucher increases as you go down the list. **An effective way to complete these choice lists is to determine in which row you would like to switch from preferring Voucher A to preferring Voucher B.**

Based on where you switch from Voucher A to Voucher B in this list, we assess which voucher amount in the early validity period (Voucher B) you value as much as the voucher amount specified in the later validity period (Voucher A). For example, in the choice list below, you would value a $40 voucher that is valid in 30 days somewhere between a $12 and a $14 voucher that is valid today, because this is where switching occurs.

Voucher A			Voucher B
Valid In 30 days: $40 Voucher	●	○	Valid Today: $2 Voucher
	●	○	Valid Today: $4 Voucher
	●	○	Valid Today: $6 Voucher
	●	○	Valid Today: $8 Voucher
	●	○	Valid Today: $10 Voucher
	●	○	Valid Today: $12 Voucher
	○	●	Valid Today: $14 Voucher
	○	●	Valid Today: $16 Voucher
	○	●	Valid Today: $18 Voucher
	○	●	Valid Today: $20 Voucher
	○	●	Valid Today: $22 Voucher
	○	●	Valid Today: $24 Voucher
	○	●	Valid Today: $26 Voucher
	○	●	Valid Today: $28 Voucher
	○	●	Valid Today: $30 Voucher
	○	●	Valid Today: $32 Voucher
	○	●	Valid Today: $34 Voucher
	○	●	Valid Today: $36 Voucher
	○	●	Valid Today: $38 Voucher
	○	●	Valid Today: $40 Voucher

If you are selected to receive an additional reward from part 1 of the study, your reward will be determined as follows:
Your choice in a randomly selected row of a randomly selected choice list determines the amount of your personal voucher. Each choice list and each row are equally likely to get selected.

Important:

- Your choices may matter for real money! If you are selected to receive a bonus, one of your choices will actually be implemented, and your decision will determine which type of voucher you receive.
- Since only one of your decisions will be randomly selected to count, you should consider each choice list independently of the others. There is no point in strategizing across decisions.

On the next page, you will see an example choice list, and you can practice making your selections.

Click "Next" to proceed to the example page.

Part 1 of this study: Instructions (3/4)

Auto-completion: The table auto-completes your choices so you don't have to click through all of the rows. You do not have to start at the top of the table. If you select Voucher A in any one row, we assume that you will also prefer Voucher A in all *above* that row. If you select Voucher B in any one row, we assume that you will also prefer Voucher B in all rows *below* that row.

Reminder: both vouchers are valid for 7 days starting on the day indicated for each voucher.

Voucher A		Voucher B
	○ ○	**Valid today: $2 Voucher**
	○ ○	**Valid today: $4 Voucher**
	○ ○	**Valid today: $6 Voucher**
	○ ○	**Valid today: $8 Voucher**
	○ ○	**Valid today: $10 Voucher**
	○ ○	**Valid today: $12 Voucher**
	○ ○	**Valid today: $14 Voucher**
	○ ○	**Valid today: $16 Voucher**
	○ ○	**Valid today: $18 Voucher**
Valid in 30 days: $40 Voucher	○ ○	**Valid today: $20 Voucher**
	○ ○	**Valid today: $22 Voucher**
	○ ○	**Valid today: $24 Voucher**
	○ ○	**Valid today: $26 Voucher**
	○ ○	**Valid today: $28 Voucher**
	○ ○	**Valid today: $30 Voucher**
	○ ○	**Valid today: $32 Voucher**
	○ ○	**Valid today: $34 Voucher**
	○ ○	**Valid today: $36 Voucher**
	○ ○	**Valid today: $38 Voucher**
	○ ○	**Valid today: $40 Voucher**

Part 1 of this study: Instructions (4/4)

Decision screen 2

When you fill out a choice list, you may feel **uncertain about whether you prefer the left or right voucher**. On decision screen 2, we will ask you to select a button to indicate **how certain** you are about how much the larger voucher amount with the later validity period is worth to you in terms of voucher credit that can be spent in the earlier validity period.

Example

Suppose that on the first decision screen you indicated that you value a $40 voucher that is valid in 30 days somewhere between a $12 and a $14 voucher that is valid today. Your second decision screen would look like this.

How certain are you that you actually value a $40 voucher that is valid **in 30 days** somewhere between a $12 and a $14 voucher that is valid **today**?

○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
0%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%

very uncertain **completely certain**

Comprehension questions

The questions below test your understanding of the instructions.

Important: If you fail to answer any one of these questions correctly, you will not be allowed to participate in the study, and you will not receive the completion payment.

1. Which of the following statements about the voucher below is true?

   Valid in 1 month: $30 Voucher

   ○ This voucher can be used to order food starting from today until no later than 1 month.

   ○ This voucher can be used to order food any time after 1 month. The validity period has no end date.

   ○ This voucher can be used to order food in the 7-day period starting in 1 month.

2. Suppose you are 80% certain that your decisions actually correspond to how much the different voucher options are worth to you.

   Which button should you click in this case?

○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
0%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%

   very uncertain **completely certain**

3. Which of the following statements is true?

   ○ Even if the validity period starts in the future, my voucher will be credited to my UberEats account shortly after the experiment. I do not have to remember the validity period because UberEats will send me reminders.

   ○ If the validity period of the voucher starts in the future, I should expect to get my voucher credited to my UberEats account only shortly before the validity period starts. I have to memorize the validity period, otherwise I may forget to use the voucher amount. There is also some risk that I will not actually receive the voucher.