# Times Are Changing:

# Projective Misperceptions and Misinferred Time Preferences*

Sili Zhang†

November 8, 2021

[See the most recent version]

**Abstract**

Intertemporal tradeoffs exist in most economic decisions and are usually interpreted as being guided by time preferences. Using an experiment and surveys, this paper shows that projective misperceptions—the tendency to project *either* one's current valuation *or* informational state onto the future—generate behavior that can be systematically misattributed to time preferences. The experiment randomizes the state in which individuals make intertemporal choice, while holding constant the time horizons in a real-effort setting. Results show that the decision state, rather than time discounting, is the primary driver of individuals' choices, and that the time preference parameter inferred from experimental choices over the same time horizon can differ by up to 44 percent, depending on the state. The observed state dependence is largely driven by the projection of informational states and can be mitigated by experience-based learning. I also develop a survey-based measure of projective misperceptions and find that it can meaningfully predict choices in the experiment and self-reported daily behavior, sometimes even better than measures of time preferences, suggesting that projective misperceptions can be seen as a potentially stable trait relevant for intertemporal decisions across domains.

*Keywords:* misperception, projection, state dependence, time preference, experience

*JEL Codes:* C91, D15, D84, D91

*The experiment was pre-registered at AEA RCT registry (AEARCTR-0007153). Financial support from the Forschungskredit grant of the University of Zurich (FK-19-018) is gratefully acknowledged.

†University of Zurich, Department of Economics, Blümlisalpstrasse 10, 8006 Zurich, CH. Email: sili.zhang@econ.uzh.ch. I am deeply indebted to my advisor Roberto Weber for his continuous support and invaluable guidance throughout the course of this project. I am very grateful to George Loewenstein for his hospitality during my visit to the Department of Social and Decision Sciences at Carnegie Mellon University, where the early versions of this project were developed and benefited from his insightful remarks. I thank Nick Netzer for his feedback and helpful input on the theory side, and Sandro Ambühl, Björn Bartling, Lorenzo Casaburi, Gretchen Chapman, Alexander Coutts, Ernst Fehr, Russell Golman, Kareem Haggag, David Huffman, Alex Imas, Anett John, Botond Kőszegi, Michel Maréchal, Wladislaw Mill, Florian Schneider, Julien Senn, Jakub Steiner, Séverine Toussaert, Stefan Trautmann, Stephanie Wang, participants in 2021 ESA Global online meeting, 2021 ESA job market webinar, as well as seminars at the University of Zurich, Carnegie Mellon University and the University of Pittsburgh for useful comments and discussions.

# 1 Introduction

Almost all important economic decisions involve intertemporal tradeoffs. Under standard economic approaches, individuals resolve these tradeoffs based on their time preferences, the extent to which they care about the future relative to the present.[1] Such individual time preferences are also important inputs into the public discount rate that informs a wide range of policies that rely on cost-benefit analysis.[2] However, it has been documented that estimates of time preferences vary substantially across individuals and sometimes even within individuals, and that there is a lack of consensus on reasons underlying this variance (Frederick et al., 2002; Havránek et al., 2021; Imai et al., 2020; Meier and Sprenger, 2015). While existing work sheds light from many perspectives, it generally shares the presumption that people *correctly perceive their own future valuations*.[3] This way, time preferences—as the weights attached to *objectively experienced* utilities at different points in time—can be revealed from intertemporal tradeoffs between *perceived* valuations over time.

However, this presumption may not hold because people may misperceive their own valuations when the environment varies substantially (Kahneman and Thaler, 2006). In fact, many intertemporal settings feature *inherent* state variation over time, because the state in which people make investment decisions usually differs from the state in which most benefits realize. For example, the decision to invest in preventive health (e.g., exercise) is usually made in a current "good" state (e.g., a healthy and/or young state) when the investment might seem unnecessary, and the benefits of investing usually realize in a future "bad" state (e.g., a sick and/or old state) when the investment produces substantial benefits. Prior research has shown that, in such environments with state variation, individuals may *project their current utility* onto the future when anticipating their future valuations.[4] For instance, healthy individuals may overly infer from their current healthy state that they will feel similarly healthy in the future, despite ample experience with illness in the past. Moreover, in settings where directly experiencing multiple states is rare or infeasible, the lack of learning opportunities itself can lead to a tendency to *project the current informational state*. For instance, many health-related decisions made in one's youth with impacts in old age are made on the basis of currently being young, and most likely without an adequate understanding of what it feels like to be old. To sum up, both types of projections can lead to misperceptions that drastically deviate from the objective valuation in intertemporal settings.

This paper investigates whether individuals' *projective misperceptions* about their own future valuations affect their intertemporal choice, and if so, how these misperceptions impact the inference of time preferences.

[1] Examples include retirement saving (Brown and Previtero, 2020; Laibson et al., 1998), human capital accumulation (Cadena and Keys, 2015), education (Carrillo, 2020), job search (DellaVigna and Paserman, 2005), adoption of technology (Duflo et al., 2011; Harstad, 2020), preventive health (Bai et al., 2020), and insurance (Baicker et al., 2015; Casaburi and Willis, 2018), etc.

[2] Examples include policies addressing climate change, the evaluation of long-term infrastructure projects, etc. See Millner (2020) and Millner and Heal (2021).

[3] Misperceptions about the time preference parameter *per se* are sometimes allowed and modeled as naive beliefs (O'Donoghue and Rabin, 1999), which is orthogonal to misperceptions about valuations. I will discuss their connection in Section 6.

[4] See Loewenstein et al. (2003) for the seminal work on intrapersonal projection of utility, a.k.a., projection bias, Gagnon-Bartsch et al. (2021) for taste projection in auctions, and Madarász (2012) for informational projection.

Answers to these questions would be of great value in evaluating both existing policies and measurement practices in the domain of time preferences. Yet, these questions are hard to tackle using observational data because states are largely endogenous and not easily observable. To causally document the empirical relevance of projective misperceptions, this paper uses a stylized experiment that is tightly linked to a theoretical framework and combines it with survey measures. The key hypothesis derived from the framework is that, in environments with state variation, projective misperceptions generate behavior that can be *systematically* misattributed to time preferences, depending on the exact state to be projected.[5] Importantly, projective misperceptions could reflect both (i) *non-informational projection of utility* that can appear despite ample experience; or (ii) *informational projection* due to the lack of experience or learning opportunities. In the preventive health example discussed earlier, by projecting the current health state (likely non-informational) or youth state (likely informational) onto the future, a misperceiving decision maker who thinks she will be equally healthy and youthful in the future may undervalue investments in preventive health and behave as if she discounts the future. Such behavior can be reversed in case the decision is made in a bad (unhealthy) state, e.g., when the individual experiences adverse health shocks or disasters.[6]

To guide the empirical analysis, I first formalize the relationship between projective misperceptions and time preferences in an intertemporal setting with state variation. The framework features decision makers with projective misperceptions à la simple projection bias (Loewenstein et al., 2003), the tendency to project current utility onto the future despite perfect foresight of states, as well as the tendency to project current informational states (Madarász, 2012). The model predicts that decision makers with projective misperceptions in a current "good" state will underestimate how badly they will suffer in a future "bad" state and will thus invest less in activities that would benefit them in such a future "bad" state, hence behaving *as if* they discount their future utilities. In contrast, if the decision is made in a current "bad" state, the behavioral pattern driven by projective misperceptions will be reversed. This means that decision makers will take more costly investment actions because they overestimate how badly they will suffer in a future "bad" state. As a result, an identification problem arises when inferring time preference parameters from observed choices: an analyst who does not take into account decision makers' misperceptions may overestimate time preference parameters from choices made in a good state but underestimate these parameters from choices made in a bad state. In terms of comparative statics, the framework predicts that the identification problem will be more severe when individuals have a higher degree of misperceptions or less experience within a context and when the magnitude of state variation is larger across contexts. Importantly, the framework helps guide the experimental design in order to empirically distinguish between *projection of utility*, which can appear despite familiarity

[5]This paper does not distinguish between different time preference models nor speak to potential preference reversals. Instead, it demonstrates a general and orthogonal point that can be applied to either classic exponential discounting (Ramsey, 1928; Samuelson, 1937) or behavioral models such as quasi-hyperbolic discounting (Laibson, 1997).

[6]For example, see Slade (2012) Callen (2015).

with the relevant state, and *projection of informational states* due to a lack of learning from experience with the alternative state, because the two can be differentiated by examining the differential responsiveness to experience with multiple states. This empirical distinction has immediate implications for using experience as an instrument to mitigate the identification problem.[7]

I then evaluate the empirical relevance of the identification problem using a longitudinal experiment (n = 430) with real-effort, which is similar to settings commonly used to structurally estimate time preferences. The experiment consists of two sessions that are two days apart and involves online participants deciding on reservation wages for future work in different sessions. In each of the two sessions, participants are required to complete ten aversive effort tasks as mandatory work. A mandatory rest is then implemented, after which participants have the opportunity to work on one additional task, either at the end of Session 1 or Session 2, in exchange for potential bonus payments. Specifically, I elicit in Session 1 participants' reservation wages for the one additional task that may take place either later in the same session ($\text{WTA}_{\text{today}}$) or in two days in Session 2 ($\text{WTA}_{\text{future}}$), using an incentive-compatible mechanism. These two wages are associated with the bonus payment to be delivered at the same time, right after Session 1, and only differ, within subject, in the time horizons for the additional work.[8]

The key design feature varies the current decision states while holding constant the time horizons. In Session 1, I vary whether participants report their reservation wages in a good state or a bad state by asking participants to report their wages at different stages during the mandatory work. In a *good state* treatment, participants report their reservation wages for the additional task at the beginning of the mandatory work, after having completed one out of ten tasks, and thus are expected to be relatively fresh. In a *bad state* treatment, participants report their reservation wages at the end of the mandatory work, after having completed nine out of ten tasks, and thus are expected to be tired. In both treatments, participants have experience with at least one task to make an informed decision about one additional task later. Because of the mandatory rest implemented before the additional task, participants' decision states do not carry over to the additional task. The variation in tiredness states in the experiment corresponds to the theoretical specification where the marginal disutility is higher in a bad state than in a good state, which is confirmed by a manipulation check.

Session 1 choices with variation in both time horizons and decision states are the primary focus of the experiment. The impact of decision states on choices and inferred time preferences directly speaks to the theoretical prediction on state dependence. Based on the presumption that people correctly perceive the disutility from additional work in the future, regardless of their current state, the null hypotheses are that

[7]The empirical distinction also helps evaluate existing micro-foundations for projection, such as the one based on recall (Bordalo et al., 2020b), which can predict an experience effect that would not arise from the original framework of projection bias (Loewenstein et al., 2003).

[8]The time delay of two days is relatively short compared to the conventional time frame used to study time discounting. Nevertheless, Augenblick (2018) shows that short-term discounting does exist in effort task setting within days.

there is no effect of current decision states on the willingness-to-accept for future work and that time horizons should matter the most in evaluating the additional work. Neither is the case in my experiment. Instead, I find strong evidence that the variation in current decision states dominates in participants' reservation wages for future work. Participants currently in a bad state (tired) ask for higher reservation wages for additional future work and also predict themselves to be more tired when working in the future than those currently in a good state (fresh), irrespective of whether the additional work takes place later in the same day or in two days. The decision state, rather than time discounting, is the primary driver of participants' wage requests in the experiment. As a result, the time preference parameter (two-day discount factor in this case) inferred from these wages over the same time horizon can differ by up to 44 percent when the analyst ignores decision makers' misperceptions.

The experiment is further designed to investigate mechanisms underlying these state-dependent work choices by focusing on the impact of experience. Recall that the theoretical framework features differential responsiveness to experience between two types of misperceptions, *projection of utility* à la projection bias that is insensitive to experience and *projection of informational states* that can be reduced by learning from experiencing relevant states. I thus elicit again in Session 2 one additional reservation wage for additional work performed later in the session, and randomize the variation in decision states in Session 2 in a similar fashion as in Session 1. Notably, the randomization in decision states in Session 2 is stratified. By stratifying the Session 2 randomization on the decision state in Session 1, I ensure that participants in Session 2 are balanced with respect to Session 1 states. This minimizes potential carryover effects of Session 1 on Session 2 such that the choices in the two sessions are as comparable as possible except for the experience with the relevant tiredness state. By directly comparing the good-bad state wage gap between the two sessions, I find that participants in Session 2 in both states do not exhibit state dependence in their work choices anymore after having experience with relevant states. This is consistent with the interpretation based on informational state projection resulting from a lack of experience-based learning (Malmendier, 2021), instead of non-informational projection of utility. The mitigating impact of experience is also consistent with a memory-based microfoundation as in Bordalo et al. (2020b). I rule out several competing explanations with further analyses.

In addition, exploratory analyses based on a post-experiment survey show that projective misperceptions can be seen as an individual trait and are potentially relevant for predicting intertemporal choice beyond the experimental context. I first develop a new and pre-registered survey measure of projective misperceptions that can account for heterogeneous treatment effects in the experiment. Specifically, I provide suggestive evidence that those who are more prone to misperceptions according to this survey measure ask for more state-dependent reservation wages in the experiment than those who are less prone to misperceptions. I then include several survey-based measures of time preferences that have been validated previously and are widely used in empirical work, including self-reported patience (Falk et al., 2016), discounting measured from the hypothetical "money-earlier-or-later" paradigm (Chapman et al., 2019) in two time frames, and perceived self-control problems as a proxy for sophistication (Ameriks et al., 2007; John, 2020). I also collect three self-reported measures of widely studied daily behavior often associated with time preferences: saving behavior, preventive health investment (take-up of COVID-19 vaccine), and procrastination at work. A comparison between the survey measure of misperceptions and all time preference measures shows that: i) preventive health investment can only be predicted by misperceptions and cannot be explained by time preference measures; ii) procrastination at work can be predicted by both misperceptions and time preference measures; iii) saving can only be predicted by time preference measures but not by misperceptions. These results are consistent with the theoretical prediction that the identification problem is more severe in contexts where the magnitude of state variation in valuations is larger, presumably in non-monetary domains as opposed to in monetary domains.

Taken together, this paper demonstrates an identification issue for the inference of time preferences caused by projective misperceptions, which is arguably *inherent* in many intertemporal investment decisions as they tend to involve systematic state variation over time.[9] For example, people may have systematic state-dependent valuations due to aging, worsening health conditions, habit formation, etc. If time preferences are misinferred as the cause of suboptimal choices in these settings, existing policy recommendations based on time preferences may be mistargeted and fail to deliver the intended impacts. Instead, interventions that target misperceptions—such as visualization-based programs (e.g., Ashraf et al., 2021; John and Orkin, 2021)—may be desirable. One other broad implication of my results is that the connection between time preferences and observed intertemporal choice might not be as strong as previously thought, at least in settings with large state variation. This may matter in both observational studies that focus on the explanatory power of time preferences as well as experimental studies that focus on the measurement of time preferences. For example, my results imply that the estimated time preferences will *appear to be* heterogeneous and context-dependent even if the "inherent" underlying preferences are homogeneous and context-general. The degree of state variation has the potential to serve as a useful input in order to evaluate the magnitude of the identification issue across different economic environments. I discuss these implications and potential solutions in more detail in Section 6.

This paper contributes to four broad strands of research. First, it provides concrete empirical evidence for one psychological mechanism of intertemporal choice, projective misperceptions, and the extent to which these misperceptions create an identification challenge for the inference of time preferences. While the conceptual insight of this paper is a direct implication of the seminal work on projection bias (Loewenstein et al., 2003), the empirical evidence on the joint analysis of time preferences and misperceptions has been lacking to date. Most empirical evidence on projection bias—e.g., early seminal experiments on food choices (Read and Van Leeuwen,

[9]The existence of state variation violates time invariance of intertemporal choices (Halevy, 2015).

1998) and heroin substitutes (Badger et al., 2007), and more observational evidence that I will discuss later—concludes by documenting the role of projection in each setting without further examining the relevance of projection relative to time preferences or the conditions under which projection is likely to dominate. In contrast, I provide concrete empirical evidence on the identification challenge within my experimental setting and suggestive evidence on how prevalent the challenge can be beyond my setting. Notably, in contrast to previous methodological work,[10] projective misperceptions can be directly incorporated to reconcile several open puzzles in the time preference literature (see Ericson and Laibson, 2019), such as the large heterogeneity, context specificity, and weak correlations between measured preferences and daily behavior. I provide more discussion regarding these puzzles in Section 6.

Second, by explicitly connecting projective misperceptions and time preferences, this paper belongs to a growing literature which demonstrates that "revealed" time preferences can reflect a range of other motives[11]: e.g., risk (Epper and Fehr-Duda, 2021); focusing (Dertwinkel-Kalt et al., 2021; Kőszegi and Szeidl, 2013); myopic forecasting errors (Gabaix and Laibson, 2017; Gershman and Bhui, 2020); cognitive uncertainty (Enke and Graeber, 2021); limited memory (Ericson, 2017); optimism (Breig et al., 2021); conservatively dealing with uncertainty about future tastes (Chakraborty, 2021); costly empathizing (Noor and Takeoka, 2021); etc. The intuition behind projective misperceptions is connected to myopic or noisy information processing (Gabaix and Laibson, 2017; Gershman and Bhui, 2020), by which agents mentally simulate the value of future events based on unbiased priors and noisy signals about own values. The projective misperceptions in this paper can be micro-founded by this forecasting process combined with either biased simulation noises or biased priors drawn from the economic environment. This paper is also connected to recent theoretical work showing that ignoring shocks on task completion (Heidhues and Strack, 2021) or personal tastes (Strack and Taubinsky, 2021) can cause non-identifiability of time inconsistency.[12] I build on the state variation over time that is inherent in many intertemporal contexts, and show that projective misperceptions create a different identification challenge even without the dynamic element.

Third, this paper provides a better understanding of mechanisms underlying projective misperceptions by distinguishing between informational projection and utility projection à la projection bias (Loewenstein et al., 2003). This distinction has immediate implications for potential interventions to mitigate misperceptions as well as normative judgements about misperceptions. Prior work shows that people are bad at predicting the impact of an incentive intervention on gym-attendance (Acland and Levy, 2015) or subjective happiness

[10] There are several canonical reasons why the elicitation of time preferences can be confounded, e.g. by utility curvature or arbitrage opportunities, etc. See Cohen et al. (2020) for a thorough overview of previous methodological work.

[11] See Loewenstein and Elster (1992) for a review in psychology. Intertemporal choice can be influenced by psychological distance with future self (Frederick, 2003), time sensitivity (Ebert and Prelec, 2007), tangibility of outcomes (Rick and Loewenstein, 2008), time perceptions (Chen and Zhao, 2020; Zauberman et al., 2009), and forecasting beliefs (McGuire and Kable, 2013) etc.

[12] The identification issue caused by state-dependent utilities in general has also been studied in decision under uncertainty (e.g., Karni et al., 1983) and the dynamic discrete choice literature (e.g., Magnac and Thesmar, 2002; Rust, 1994).

in commuting choices (Frey and Stutzer, 2014) and major life-changing events (Levitt, 2021; Odermatt and Stutzer, 2019). The empirical evidence consistent with projection has mostly been found in observational studies using weather-related data, because the natural variation in weather state is arguably random. For instance, it has been shown that current weather or air pollution situations influence consumers' orders of winter-related products (O'Donoghue and Vogelsang, 2007), college choices (Simonsohn, 2009), car purchases (Busse et al., 2015), outdoor movie tickets (Buchheim and Kolaska, 2017), or health insurance (Chang et al., 2018). Yet, the nature of observational data limits the ability of previous work to pin down the exact mechanism or shut down alternative explanations such as mistaken beliefs or salience (Bordalo et al., 2013), which leaves both policy and welfare implications unclear. In this paper, I provide a better understanding of these channels using a very stylized, simple and transparent experiment while minimizing alternative competing explanations.

Finally, methodologically, this paper is connected with the recent literature that does structural estimation on present bias using intertemporal allocation of real-effort tasks (e.g. Augenblick et al., 2015; Augenblick and Rabin, 2018; Fedyk, 2021; Imas et al., 2021). In the experiment most closely related to mine, Augenblick and Rabin (2018) structurally estimate present bias and naiveté about present bias by repeatedly eliciting both predictions and allocation of tasks within subjects across weeks in the effort setting. In contrast, using a well-powered experiment without structural assumptions, this paper introduces an innovation to their task allocation paradigm by focusing on the average treatment effects of decision states on reservation wages.[13] This allows for a demonstration of an identification issue *inherent* in many investment decisions, beyond the effort setting. While Augenblick and Rabin (2018) also acknowledge the existence of fluctuating tastes over effort and separately consider the role of projection bias, their results appear to be inconclusive since their study is not tailored to examine the question concerned in this paper. Le Yaouanq and Schwardmann (2020) focus on explicitly eliciting beliefs in an effort allocation setting and comparing them with a Bayesian benchmark, while intentionally shutting down the potential influence of projection bias. They show that people can learn self-control but persistently underestimate their future learning, which they interpret as the reason for persistent naiveté. Although I do not study naiveté in particular and stick to a general notion of time preferences, my results are potentially consistent with theirs, since underestimation can be seen as misperceptions. I discuss the connection between misperceptions and persistent naiveté in more detail in Section 6.

The remainder of the paper proceeds as follows. Section 2 presents a theoretical model of projective misperceptions and derives the main hypotheses. Section 3 describes the design of the experiment that is tightly tied to the theoretical framework. Section 4 presents main results from the experiment and Section 5 presents correlational results from the post-experiment survey. Section 6 discusses connections with several puzzles in time preferences and potential implications for applied work. Finally, Section 7 concludes.

[13] A similar setting has been used in Bushong and Gagnon-Bartsch (2020) to study interpersonal projection. Relatedly, Kaufmann (2021) theoretically examines projection bias over effort and its implications on work management.

## 2 Theoretical Framework

In this section, I formalize a simplified intertemporal choice setting where a decision maker may have projective misperceptions that are responsive to experience. The decision maker decides on how much to invest in an activity with immediate cost but delayed state-dependent returns. She may misperceive these returns either because (i) she incorrectly projects her current utility despite the perfect foresight of states; or because (ii) she incorrectly projects her current informational state due to a lack of experience. I first show that either of these projective misperceptions leads to state-dependent choices that can be misattributed to time preferences. That is, the decision maker currently in a good state will behave as if she discounts her future utility and that the decision maker currently in a bad state will exhibit "anti-discounting". An identification problem can therefore arise when inferring the cause of the observed intertemporal choice, because different combinations of the decision maker's time preferences and the degree of projecive misperceptions can lead to the same patterns of actions depending on the current state to be projected. Further, I show that ignoring these projective misperceptions in an environment with state variation can lead an analyst to mistakenly estimate time preference parameters (e.g., discount rates $\delta$ and/or present bias $\beta$). It is because choices driven by projective misperceptions can be incorrectly classified as arising from time preferences depending on the exact state in which observed choices are made. This leads to a prediction that the time preference parameter will be overestimated using choices made in a good state and underestimated using choices made in a bad state. Finally, I demonstrate that this identification issue will be more severe among individuals with a higher degree of misperceptions or less experience with relevant future states within a given context, and when the state variation is larger across contexts.

### 2.1 Decision Problem in Environments with State Variation

Consider a decision maker who decides in Period 1 on how much to invest in an activity $a \in \mathbb{R}_+$ that would benefit her in Period 2 in a two-period setting. The decision maker's instantaneous consumption utility in each period is state-dependent, $u(a, s) : \mathbb{R}_+ \times \mathcal{S} \to \mathbb{R}$, which is assumed to be increasing, concave and continuously differentiable in the activity $a$ for each state $s$.

There are two different states of the world, a good state ($s = g$) and a bad state ($s = b$). In each period, the state of the world is independently drawn from a binary set $s \in \mathcal{S} \equiv \{b, g\}$. The state realizes with probability $Pr(s = b) = \pi \in (0, 1)$ and $Pr(s = g) = 1 - \pi$, respectively. The decision maker holds correct beliefs about the probability distribution. Importantly, what distinguishes between the two states is the following condition:

- **State-dependent return condition**: $u'_a(a, b) > u'_a(a, g)$ for all $a$, where $\nu(a) = u'_a(a, b)/u'_a(a, g) > 1$ captures the magnitude of state variation at the activity $a$.

The above state-dependent return condition imposed on the economic environment requires the costly activity $a$ to be more beneficial to the decision maker in a bad state than in a good state. The ratio between the marginal return in a bad state and the marginal return in a good state, $\nu(a)$, thus captures the magnitude of state variation present in an economic environment. This condition reflects a systematic wedge between the investing state and the benefiting state, which is *inherent* in a wide range of intertemporal investment decisions. Take preventive health investment as an example; this condition means that individuals who are horribly sick ($s = b$) would benefit more from daily investment in health maintenance, and that those who are absolutely healthy ($s = g$) would benefit to a lesser extent.

The beneficial activity $a \in \mathbb{R}_+$ incurs an immediate cost of $C(a)$, which is assumed to be increasing, convex and continuously differentiable in the activity $a$. While the cost of the activity is immediate in Period 1, the return of the activity only realizes with a delay and depends on the exact state realization in Period 2. The decision maker discounts her future utility by $\delta$, which captures her time preference. Note that this paper does not aim to distinguish between different types of discounting models and I restrict attention to a two-period setting for simplicity. Nevertheless, all the results can in principle be applied to the present-bias parameter $\beta$ as in quasi-hyperbolic discounting (Laibson, 1997; O'Donoghue and Rabin, 1999) as well.

When deciding on the activity, the decision maker is aware of the current state $s$ and her current valuation in the current state $u(a, s)$. Before the realization of the state in the next period, $s'$, the decision maker makes her decision based on her perceptions about her future valuation, $u(a, s')$. In particular, she perceives her own valuation in a good state to be $\tilde{u}(a, g|s)$ and her own valuation in a bad state to be $\tilde{u}(a, b|s)$. Taken together, the decision maker's action in the current state $s$ is characterized by the following optimization problem to maximize her perceived expected utility:[14]

$$\max_{a \in \mathbb{R}_+} \delta[(1-\pi)\tilde{u}(a, g|s) + \pi \tilde{u}(a, b|s)] - C(a), \text{ where } s \in \mathcal{S} \equiv \{b, g\}$$

## 2.2 Projective Misperceptions Lead to State-Dependent Actions

The decision problem above leaves open how the decision maker forms perceptions about her own valuations in different states, $\tilde{u}(a, s'|s)$. The following Definition 1 fills the gap by characterizing *projective misperceptions*, with which she tends to project her current utility or informational state onto a different state in the future.[15]

**Definition 1 (Projective Misperceptions).** *A decision maker has projective misperceptions if there exists* $\alpha(e) \in [0, 1]$ *such that for all activity* $a \in \mathbb{R}_+$*, current and future states* $s, s' \in \mathcal{S}$*, her predicted future utility is:* $\tilde{u}(a, s'|s) = \alpha(e)u(a, s) + [1 - \alpha(e)]u(a, s')$*.*

[14] Misperceptions can be interpreted as the separation of ex ante decision utility from ex post experienced utility (Kahneman and Thaler, 2006; Kahneman et al., 1997).

[15] See Buckner and Carroll (2007) for a neuro-foundation for self-projection.

The formulation above of projective misperceptions is akin to simple projection bias, the tendency to project current preferences onto the future, as in Loewenstein et al. (2003). It is motivated by empirical evidence that people understand qualitatively how their tastes will change but usually underestimate the magnitude of such change. One prominent example is that people underestimate their own adaptation to life changes. If $\alpha(e) = 0$, the decision maker has no misperception when predicting her future utility and her prediction equals her true future utility, $\tilde{u}(a, s'|s) = u(a, s')$; if $\alpha(e) = 1$, the decision maker considers her future utility to be exactly the same as her current utility in the current state, $\tilde{u}(a, s'|s) = u(a, s)$.

In addition, the formulation is also augmented with responsiveness to experience, $e \in \mathbb{R}$, with the relevant alternative state $s'$. Projective misperceptions may be mitigated by having more experience with the relevant future state, such that $\alpha(0) \geq 0$, $\frac{\partial \alpha(e)}{\partial e} \leq 0$ and $\lim_{e \to \infty} \alpha(e) = 0$. This potential responsiveness to experience stems from two slightly different interpretations of projective misperceptions[16]:

- **Projection of utilities**: the first interpretation is that the decision maker correctly predicts changes in states but incorrectly perceives how changes in states can translate into changes in utility, reflecting the original intuition of projection bias in Loewenstein et al. (2003). Under this interpretation, projective misperceptions can occur despite ample experience and thus are *not responsive to experience* with relevant unfamiliar states: $\frac{\partial \alpha(e)}{\partial e} = 0$.

- **Projection of informational states**: the second interpretation is that the decision maker incorrectly anticipates changes in states, per se, reflecting the intuition of information projection (Madarász, 2012) or a lack of learning. Under this interpretation, projective misperceptions occur because the decision maker does not have much information or experience with relevant future states. Thus, this type of projection *can be mitigated by having more state-related experience*: $\frac{\partial \alpha(e)}{\partial e} < 0$.

Although Loewenstein et al. (2003) adopt the first interpretation, they are largely agnostic about the two interpretations and deem the responsiveness of projection to experience as an open question. A similar reduced-form perspective is also adopted by most empirical follow-up work (e.g., Busse et al., 2015; Chang et al., 2018; O'Donoghue and Vogelsang, 2007). However, the distinction between the two types of misperceptions, thus the exact role of experience, is potentially important because it yields immediate policy implications on whether it is possible to mitigate misperceptions and any associated consequences. Conceptually, the responsiveness to the experience is also a distinct feature of memory-based choice models that can micro-found projection (Bordalo et al., 2020b), which treats current decision states as contextual cues that may interact with remembered past experience. I provide more discussion in the empirical part, where I tease apart the two types of misperceptions empirically by design.

[16]I provide a different version of the model that explicitly separates between the two types of misperceptions in Appendix A. The two versions are equivalent when the instantaneous utility is assumed to be linear in the state $s$.

One implicit assumption behind the formulation above is that the decision maker is unaware of her current or even future misperceptions such that there is no role for sophistication. Otherwise, the decision maker would have not misperceived were she aware of this tendency. Nevertheless, people may still have some meta-awareness in retrospection when concluding from past occurrences of misperceptions. This self-knowledge makes it possible to have a general measure of misperceptions. I provide more discussion in the empirical part, where I develop a survey-based measure of misperceptions, $\alpha(e)$.

The following Lemma 1 characterizes the optimal activity of a decision maker with time preferences denoted as $\delta$, and projective misperceptions captured by $\alpha(e)$, making decisions in different states.[17] In contrast, the benchmark activity in the absence of misperceptions and time discounting (i.e. $\alpha(e) = 0$ and $\delta = 1$) can be characterized by $C'(a^*) = (1-\pi)u'_a(a^*, g) + \pi u'_a(a^*, b)$. I focus on the non-trivial case where the activity is beneficial, $C'(0) < \delta[1-\alpha(e)][(1-\pi)u'_a(0, g) + \pi u'_a(0, b)]$. Proofs for the lemma and all subsequent results are provided in Appendix B.

**Lemma 1.** *The optimal activity, $a_g^*$, of a decision maker currently in a good state is characterized by:*

$$C'(a_g^*) = \delta\left([1-\alpha(e) + \frac{\alpha(e)}{1-\pi}](1-\pi)u'_a(a_g^*, g) + [1-\alpha(e)]\pi u'_a(a_g^*, b)\right).$$

*The optimal activity, $a_b^*$, of a decision maker currently in a bad state is characterized by:*

$$C'(a_b^*) = \delta\left([1-\alpha(e)](1-\pi)u'_a(a_b^*, g) + [1-\alpha(e) + \frac{\alpha(e)}{\pi}]\pi u'_a(a_b^*, b)\right).$$

The optimal activity is pinned down by the first order condition of the optimization problem.[18] Intuitively, one would always prefer to take up the costly activity as long as the marginal cost is smaller than the perceived marginal benefit. Importantly, the perceived marginal benefit and the actual marginal benefit of a misperceiving decision maker do not necessarily coincide because a misperceiving decision maker decides on the activity based on her predicted future utility as captured by Definition 1. The (perceived) optimal activity will thus be state-dependent since the perceived marginal benefit of the activity will be biased towards the perception in the current state, which is summarized in Proposition 1.

**Proposition 1** (**State Dependence**). *Deciding in a good state, a decision maker with projective misperceptions, $\alpha(e) > 0$, will act less on the activity with delayed benefits than the benchmark level in the absence of misperception and discounting, $a_g^* \leq a^*$ for any $\delta \in (0, 1]$.*

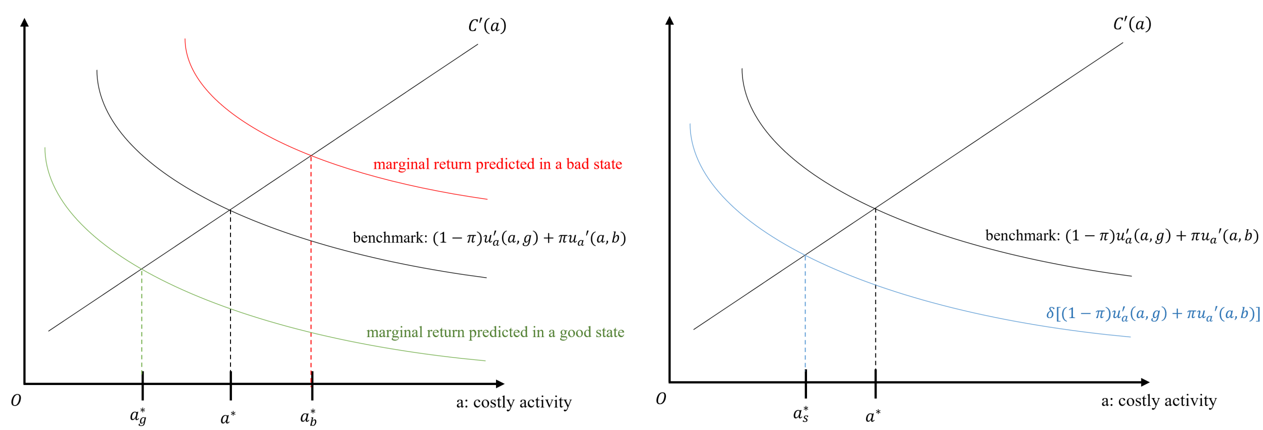
[17] The framework considers a simplified two-period setting. Interesting dynamics may appear when the next new state is determined by previous state-dependent actions, which is beyond the scope of the paper but interesting for future research.

[18] Note that throughout the analysis the use of the terms "optimality" or "benchmark" does not convey any normative judgement for two reasons. First, as mentioned earlier, different time preference models will not be distinguished in this paper. Thus, the behavior of discounting the future can reflect either classic exponential discounting or present bias as a behavioral "mistake". Second, when there are multiple different states, it becomes ambiguous which state is the welfare-relevant one. See Bernheim (2009) and Manzini and Mariotti (2014) for discussions on welfare analysis in behavioral economics.

*Deciding in a bad state, a decision maker with projective misperceptions, $\alpha(e) > 0$, will act more on the activity with delayed benefits than the benchmark level in the absence of misperception and discounting, $a_b^* \geq a^*$ for any $\delta \geq \bar{\delta}$.*

This proposition indicates that the costly activity with delayed benefits will be jointly determined by the decision maker's time preferences and misperceptions in a state-dependent manner. That is, the activity will be lower when a misperceiving decision maker is deciding in a good state than in a bad state. When compared with the benchmark level in the absence of misperceptions and time discounting, with projective misperceptions, the observed intertemporal activity can be either lower or higher depending on the current decision state. Figure 1 provides a graphical illustration.

Figure 1: Decisions with and without Projective Misperceptions




*Notes*: The left figure shows the optimal activity of a misperceiving decision maker, which is characterized by the tradeoff between the marginal cost and the marginal benefit of the activity perceived in a good state ($s = g$) and a bad state ($s = b$). The right figure shows the same tradeoff of a decision maker who has no misperceptions in both states ($s \in \mathcal{S} \equiv \{b, g\}$) but discounts her future utility by $\delta$.

This result provides a direct test of projective misperceptions. Time preferences alone with $\delta \leq 1$ can only lead to lower immediate activity, independently of the current state. A state-dependent pattern of actions as characterized in Proposition 1 is therefore more likely to be generated by projective misperceptions. Put differently, a decision maker with projective misperceptions in a good state will underestimate how badly she will suffer in a future bad state and will thus act less on the costly activity that would benefit her in a bad state, behaving as if she heavily discounts her future utilities. In contrast, when acting in a bad state, the behavior pattern driven by projective misperceptions can be reversed. The decision maker will act more because she misperceives by overestimating how badly she will suffer in a future bad state. She may even exhibit *anti-discounting* when her "true" underlying time preference $\delta$ is sufficiently large.

## 2.3 Inferring Time Preferences from State-Dependent Actions

Consider an analyst who aims to infer time preferences $\delta$ from the decision maker's actions. Let $\mathcal{A} \subseteq \mathbb{R}_+$ be the permissible set of all possible activities generated by projective misperceptions alone. Suppose the decision maker's instantaneous utility function and cost function satisfy all the assumptions specified earlier, are fixed in the environment and are perfectly observable to the analyst.[19] Proposition 2 formalizes a critical implication of projective misperceptions when the analyst aims to investigate the explanatory power of time preferences using observed data from the set $\mathcal{A}$.

**Proposition 2** (**Non-identifiability**). *Any arbitrary level of activity $a \in \mathcal{A}$ generated only by projective misperceptions, $\alpha(e) \in (0, 1]$ and $\delta = 1$, in a given state $s \in \mathcal{S}$ can be rationalized by time preference alone, $\delta > 0$ and $\alpha(e) = 0$, as the solution of the optimization problem.*

The non-identifiability result suggests caution in attributing intertemporal choice to time preferences without further examination. It shows that projective misperceptions alone in a good state ($s = g$) can generate activities that look like those caused by time preferences $\delta \leq 1$ relative to the benchmark, and that misperceptions alone in a bad state ($s = b$) can generate activities that look like those caused by time preferences $\delta > 1$ relative to the benchmark. Put differently, projective misperceptions generate behavior that can be *systematically* misattributed to time preferences depending on the exact current state to be projected.[20] As a direct consequence, time preferences and misperception parameters cannot be separately identified if an analyst only observes choices made in one state.

The next result, Proposition 3, characterizes the estimation bias when the analyst aims to recover the time preference parameter $\delta$ from observed actions but ignores the existence of the decision maker's misperceptions. That is, the analyst mistakenly believes that the decision maker's action is solely driven by time preferences (i.e. $\delta > 0$ and $\alpha(e) = 0$). Notably, the estimation bias depends on both the magnitude of misperceptions and environment-specific distorting factors, $\eta_g$ and $\eta_b$. The detailed characterization of these distorting factors is provided in Appendix B and I discuss how these distorting factors respond to the magnitude of state variation across contexts as comparative statics in the next subsection.

**Proposition 3** (**Mis-estimation**). *When the activity is decided in a good state, the analyst who ignores the decision maker's misperceptions will overestimate time preferences: $\hat{\delta} = \delta[1 - \eta_g \alpha(e)] \leq \delta$ where $\eta_g > 0$.*

*When the activity is decided in a bad state, the analyst who ignores the decision maker's misperceptions will underestimate time preferences: $\hat{\delta} = \delta[1 + \eta_b \alpha(e)] \geq \delta$ where $\eta_b > 0$.*

[19] Of course, there are other canonical reasons why the measurement of time preferences can be confounded, such as by the utility curvature (Andersen et al., 2008; Andreoni and Sprenger, 2012; Attema et al., 2016). These factors are important by themselves, but are assumed away here for simplicity as they are orthogonal to the focus of this paper.

[20] The converse of the statement holds under additional restrictions on the magnitude of state variation in a given economic environment because misperceptions $\alpha(e)$ is bounded by 1 from above.

This result suggests that in settings that involve state variation, time preference parameters may be overestimated or underestimated when not accounting for projective misperceptions, depending on the exact state in which choices are made. In particular, if choices are made in a good state, time preference parameters will be overestimated because projective misperceptions and time preferences are in line with each other. Instead, if choices are made in a bad state, time preference parameters will be underestimated because the two forces are against each other.

One might hope that the identification issue can be alleviated if the analyst can observe choices in both states repeatedly and treat the variation in states as noises in the data. Corollary 4 in Appendix A.2 shows that the distorting impact in the bad state and the impact in the good state do not always cancel out on average and formalizes the conditions under which they do. I also discuss other potential solutions in practice in Section 6 in more detail.

### 2.4 Comparative Statics Within and Across Contexts

The last part of the theoretical analysis considers comparative statics, in particular, how the identification issue caused by state-dependent misperceptions responds to individual heterogeneity within a given context and the magnitude of state variation across different contexts.

**Corollary 1.** *The degree of state dependence is increasing in the magnitude of projective misperceptions:* $\frac{\partial a_g^*}{\partial \alpha(e)} < 0$ *and* $\frac{\partial a_b^*}{\partial \alpha(e)} > 0$ *for any* $\delta > 0$.

**Misperceptions.** Corollary 1 shows how state-dependent actions respond to the heterogeneity in projective misperceptions at the individual level. It suggests that individuals more prone to misperceptions will act less in a good state and act more in a bad state, relative to those less prone to misperceptions. As a result, the state dependence captured by the good-bad state activity gap is larger among individuals who are more prone to projective misperceptions. This result thus provides a basis for heterogeneity analysis of state dependence with respect to a measure of projective misperceptions at the individual level.

**Corollary 2.** *The degree of state dependence is decreasing in experience with relevant states:* $\frac{\partial a_g^*}{\partial e} \geq 0$ *and* $\frac{\partial a_b^*}{\partial e} \leq 0$ *for any* $\delta > 0$.

**Experience.** Recall that projective misperceptions are potentially responsive to experience and can be mitigated by having more experience with relevant states if the misperceptions are due to projection of informational states, $\frac{\partial \alpha(e)}{\partial e} \leq 0$. Corollary 2 suggests that the state dependence would then be reduced by experience within a given context. This result can also be applied across contexts, meaning that the identification issue might be less severe in settings where people make choices repeatedly and have richer context-specific experience relative to settings where people make once-in-a-life choices.

**Corollary 3.** *The estimation bias of ignoring misperceptions is increasing in the magnitude of state variation in the context: $\eta_g$ is increasing in $\nu(a_g^*)$ and $\eta_b$ is increasing in $\nu(a_b^*)$.*

**State variation.** Recall that Proposition 3 characterizes the estimation bias $\hat{\delta} = \delta[1 \pm \eta_s \alpha(e)]$ with distorting factors, $\eta_g$ and $\eta_b$. Corollary 3 considers how these distorting factors respond to the magnitude of state variation across environments. It suggests that the identification issue is more severe in contexts with a larger degree of state variation (i.e. $\nu(a) = u'_a(a,b)/u'_a(a,g)$ at the corresponding $a$). For instance, non-monetary settings may be more likely to have larger state variation as opposed to monetary settings, presumably because people are easily susceptible to taste shocks over consumption goods but usually have relatively stable preferences over money. People also have more experience dealing with money in daily life.

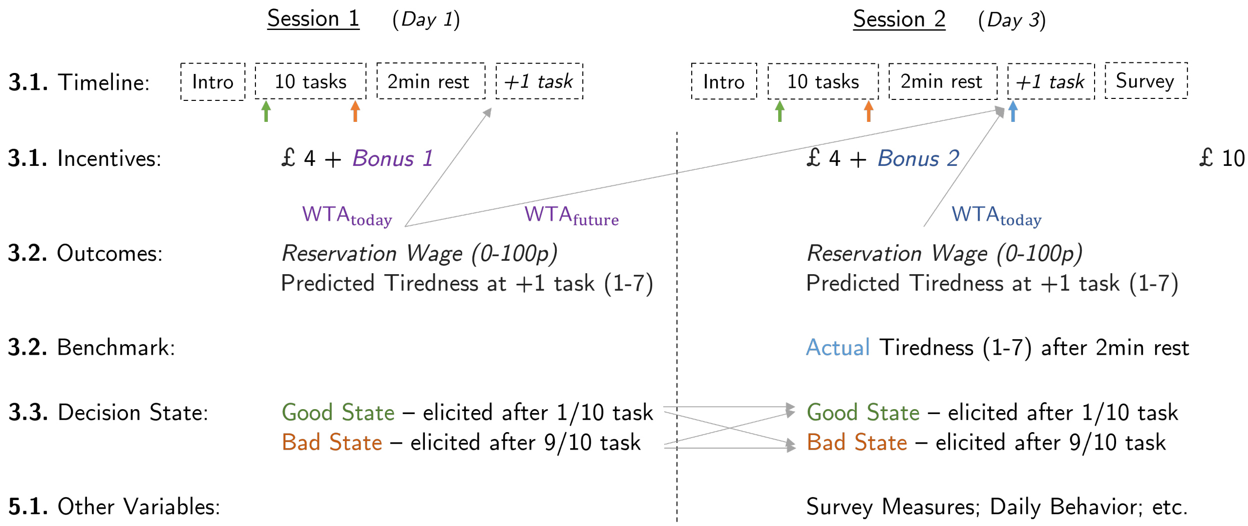
Taken together, these comparative statics suggest that the state dependence generated by misperceptions is highly heterogeneous, can potentially be reduced by having more experience, and highly depends on specific contexts. As a result, when an analyst aims to recover time preferences from observed choices but fails to take these misperceptions into account, the estimated time preference parameters will *appear to be* heterogeneous and domain specific even if the "inherent" underlying time preferences are homogeneous and domain general. I provide more discussion in more detail in Section 6.

# 3 Experimental Design

In this section, I present the design of the online work experiment, in which participants in different *decision states* report reservation wages for future work at different *time horizons*. The experiment consists of two sessions that are two days apart. In each of the two sessions, participants are required to complete the mandatory work of ten aversive tasks with real effort. The pre-registered outcome variables are participants' reservation wages for one additional task that takes place either later in the same session ($\text{WTA}_{\text{today}}$) or in the next session ($\text{WTA}_{\text{future}}$), which reflects the variation in *time horizons*. These wage choices are elicited in different decision states, which is varied by the accumulated tiredness from the mandatory work at different stages in each session. In a *good state*, wages are elicited at the beginning after one out of ten mandatory tasks, when participants are not tired; whereas in a *bad state*, wages are elicited at the end after nine out of ten mandatory tasks, when participants are tired. Figure 2 shows an overview of the experimental design and I will explain each element in more detail in the corresponding section.

The experiment is tightly tied to the theoretical predictions on state-dependence and mis-estimation, as in Proposition 1 - 3. The main objective is to test whether the variation in current tiredness states (Good State vs. Bad State) affects wage choices ($\text{WTA}_{\text{today}}$ and $\text{WTA}_{\text{future}}$) about future work at different points in time, and how this state variation can lead to misinferences regarding time preferences over a fixed time horizon.

Figure 2: An Overview of the Experimental Design




As with any experiment, both the randomized assignment of artificial decision states and the stylized setting come at a cost of generalizability.[21] Nevertheless, I aim to show that even in this highly simplified and transparent context where the scope of misperceptions is limited, the state-dependent pattern can still emerge and confound the inference of time preferences. To further alleviate the concern, the experiment is also designed to link behavior within my experimental context with self-reported daily behavior outside of the experimental context. This part of the analyses based on survey variables is presented in Section 5. Moreover, in contrast to a naturalistic environment, the controlled setting allows me to rule out several explanations that are otherwise hard to exclude, such as mistaken probabilistic beliefs.

## 3.1 Experimental Setup

**Real-Effort Tasks.** The experiment builds on the recent literature that studies present-biased time preferences using the allocation of real-effort tasks over time (e.g., Augenblick et al., 2015; Augenblick and Rabin, 2018). The particular task is to count the number of zeros in a randomly generated matrix that contains different numbers of ones and zeros. To complete each task, participants need to count the number of zeros in one matrix correctly. There is no time limit for each task but participants cannot proceed to the next task until the previous one is completed. The experimental instructions including a sample matrix are provided in Appendix C.

The matrix used for the task is four times as large as the usual matrix used in earlier studies (e.g., Abeler et al., 2011) and thus has a high degree of aversiveness. This feature helps induce the treatment variation in

[21]The original plan for an in-person experiment across weeks with more hedonic states unfortunately became infeasible due to COVID-19. The current setting is inspired by the effort task paradigm in Augenblick and Rabin (2018), which is later used in Bushong and Gagnon-Bartsch (2020) to study interpersonal projection.

tiredness states while keeping the experimental structure simple, which I will explain in detail later. Moreover, the task itself is also quite straightforward and involves minimal skills. This limits the role of learning from task skills and overconfidence about task skills. I also keep track of the response time and the number of errors during each task in the experiment to fully control for these factors. Nevertheless, learning is of many different kinds and the lack of learning itself can be a potentially important reason for projection of information states, which I will discuss in more detail when presenting the results.

**Timeline and Incentives.** The experiment consists of two sessions that are two days apart, Session 1 and Session 2. In each session, participants need to complete a mandatory work block of ten counting tasks to receive a fixed payment of £4 to be delivered after that session. After these mandatory tasks in each of the two sessions, participants have a two-minute rest. They can do one additional task *after having rested* and receive an additional bonus for doing so. The bonus payment for each session is determined by participants' *reservation wage for the one additional task* that takes place at different points in time, as described in more detail in the next subsection.

Note that the time of two days elapsed between two sessions in the experiment is relatively short compared to the conventional time frame used to study time discounting. Despite the conventional belief that time preferences usually operate over longer time periods, existing evidence suggests that short-term discounting exists in an effort task setting within the span of days (Augenblick, 2018) and that discounting over primary rewards can appear in brain activities within minutes (McClure et al., 2007).[22] This design choice was made to reduce attrition, a fairly common issue in longitudinal studies online. I also offer participants a completion payment of £10 on top of their session-specific payments if they complete both sessions. This further reduces attrition across days, minimizing any potential impact of selection on the empirical inference.

## 3.2 Elicitation of Main Outcomes

**Reservation Wages and Variation in Time Horizons.** I elicit three different reservation wages—as participants' willingness-to-accept (WTA)—for one additional task that takes place at different points in time as primary outcomes. In Session 1, each participant reports two reservation wages, one for the additional task later at the end of Session 1 in the same day and the other for the additional task at the end of Session 2 in two days. In Session 2, most participants report another reservation wage for the additional task later at the end of Session 2 in the same day.[23] As shown in Figure 2, the first bonus payment in Session 1 is determined by either the wage for the additional task later in Session 1 ($\text{WTA}_{\text{today1}}$) or the wage for the additional task in two days in Session 2 ($\text{WTA}_{\text{future}}$), and the second bonus payment in Session 2 is determined by the wage

[22] "How soon is now" itself remains an important and unsettled open question (Ericson and Laibson, 2019).

[23] A few participants do not make Session 2 choice as part of the incentive-compatible elicitation procedure, as explained later.

for the additional task later in Session 2 ($\text{WTA}_{\text{today2}}$). The specific procedure to ensure incentive-compatible elicitation will be explained in detail later. All three wage choices are summarized as follows.

- $\text{WTA}_{\text{today1}}$ to be received in Session 1 for one additional task at the end of Session 1
- $\text{WTA}_{\text{future}}$ to be received in Session 1 for one additional task at the end of Session 2
- $\text{WTA}_{\text{today2}}$ to be received in Session 2 for one additional task at the end of Session 2

The only difference in two reservation wages reported in Session 1, $\text{WTA}_{\text{today1}}$ and $\text{WTA}_{\text{future}}$, is whether the additional task takes place later in the same day or in two days, since either of the two wages may determine the Session 1 bonus payment that will be received at the same time. Therefore, as a result of discounting future work disutility, participants who discount future disutilities because of time preferences are likely to ask for lower wages if additional work takes place in two days instead of later today.

**Incentive-Compatible Elicitation.** For each wage choice, participants face an incentivized multiple price list that trades off one additional task at a certain point in time against a bonus payment (up to £1, which would be paid out right after the corresponding session). Depending on their choices, participants either have to complete one additional task and receive the implemented bonus payment after the corresponding session or skip both the task and the corresponding bonus payment, based on the outcome of a BDM mechanism (Becker et al., 1964). The screenshot of multiple price list choices is provided in Appendix C.[24]

To ensure incentive compatibility of all three wages, participants are told in each session that the computer will determine *whether and when* an additional task takes place by chance after they make all the choices in that session. If the computer determines that the additional task takes place, one of their choices for the corresponding additional task in the multiple price list will be randomly selected to count (each choice is equally likely to be selected). I do not guarantee that the additional task will take place or that one of their choices will be implemented. Instead, I randomize 90% of the participants in Session 1 to have *no* additional task taking place at all, independently of their wage choices in Session 1. This randomization ensures that the wage choice made in Session 2 is meaningful because the status of the Session 2 additional task has not yet been determined by wage choices made in Session 1 for the majority of participants.[25] This procedure also

[24]The participant who switches in their multiple price list choices at most once is considered as making consistent choices that can be used to infer reservation wages. I did not enforce single switching as it may induce extra bias on participants' choices. Instead, I included explanatory text at the bottom of the choice list explaining that the switching point indicates their required minimum compensation for the task and that it would be sensible to switch once at most. Yet, participants can make choices however they want without getting any further notifications. Empirically, less than 5 percent of participants switch more than once for each choice. These inconsistent choices will be excluded from the analysis.

[25]This procedure involves no deception because participants are explicitly told that whether the additional task takes place at all is determined by chance. An alternative design that allows participants to override their earlier choices as a surprise may be considered as deceptive. As a result of this randomization, 90% of the participants will make choices that are incentive-compatible in both sessions. The remaining 10% of the participants in Session 1 have the additional task taking place (based on their Session 1 choices) either at the end of Session 1 or at the end of Session 2 with equal probabilities. This means that around 5% of the participants will have their Session 2 task status determined in Session 1, thus, they do not make another set of wage choices in Session 2.

helps avoid potential impact of selection on Session 2 choices that may arise if participants' Session 1 wage choices correlate with their task experiences when deciding again on Session 2 wage choices.

To ensure that participants understand the setting and the elicitation procedures, all the choices to be made in each session together with the relevant screenshots are previewed at the beginning of the corresponding session, before the start of the mandatory work. Participants must also pass a comprehension quiz on how their bonus payment will be determined. They can proceed to the main experiment only if they answer all the quiz questions correctly.

**Benchmark Tiredness and Predictions.** State in this experiment is implemented as the perceived tiredness (or disutility). The actual state in which the additional task takes place is thus the experienced tiredness after a 2-minute rest following the mandatory work at the end of each session. This compulsory rest is included to reset the task effort, and participants can spend the 2-minute rest period however they want. I elicit actually experienced tiredness (on a scale of 1 to 7, from "not tired at all" to "extremely tired") reported after the 2-minutes rest and right before the realization of the additional task in Session 2 as an objective benchmark. The treatment variation in decision states is described in the next subsection.

Recall from the theoretical framework that projective misperceptions can stem from both projection of utilities and projection of (informational) states. Participants with utility misperceptions can misperceive their work preferences under a given tiredness level while accurately perceiving the tiredness level, whereas those with state misperceptions can accurately predict their work preferences under a given tiredness level but misperceive the tiredness level. The experiment allows me to distinguish these two interpretations. To this end, I ask participants to predict how tired they would be (on a scale of 1 to 7, from "not tired at all" to "extremely tired") if they had to perform one additional task at a certain point in time whenever a wage choice is elicited.[26] The comparison between participants' predicted tiredness and the benchmark tiredness thus reveals whether participants can correctly predict their tiredness state.

## 3.3 Treatment Variation in Decision States

**Variation in Decision States.** Similar experimental settings with time variation are commonly used for structural estimation of time preferences. I introduce an innovation to this experimental paradigm by randomizing states in which participants report their reservation wages on top of the variation in the time dimension. Specifically, I elicit participants' wages at different stages of mandatory work in each of the two sessions, varying in the degree of accumulated tiredness from mandatory tasks between subjects. As summarized in the following, in the *good state* treatment, the wage choices are elicited at the beginning of the mandatory work, after one out of ten tasks, when participants have experienced one task but are still quite

[26]To avoid consistency considerations, the self-reported predictions are not incentivized based on accuracy.

fresh. In the *bad state* treatment, the wage choices are elicited at the end of the mandatory work, after nine out of ten tasks, when participants are tired. Importantly, these tiredness states induced by mandatory tasks are supposed to be transient and should not carry over after having rested.[27]

- *Good state*: wage choices are elicited after 1/10 mandatory tasks, when participants are fresh
- *Bad state*: wage choices are elicited after 9/10 mandatory tasks, when participants are tired

This experimental manipulation keeps constant most features other than the tiredness state. For instance, participants in both treatments are in the same process of completing mandatory work and they all have the experience with at least one task to make an informed decision about one task.[28] To validate the experimental manipulation, I elicit self-reported tiredness (on a scale of 1 to 7, from "not tired at all" to "extremely tired") whenever participants are about to make wage choices. This allows a manipulation check of whether the state-dependent return condition in the theoretical framework, $u'_a(a, b) > u'_a(a, g)$, holds in the experimental setting, i.e. whether participants in a bad state have higher marginal disutilities from additional work.

**Stratification in Session 2.** Session 2 is designed to pin down the exact mechanism underlying projective misperceptions, by examining whether misperceptions are sensitive to experience with relevant tiredness states. Recall that both types of projective misperceptions can be reflected in Definition 1, *the projection of informational states* due to a lack of experience with unfamiliar states and *the projection of utilities* despite accurate perceptions of states.

I randomize decision states in Session 2 in a stratified manner conditional on the decision state in Session 1. By stratifying the Session 2 randomization on the decision state in Session 1, I ensure that participants in Session 2 are balanced in respect to Session 1 states. This ensures that choices in two sessions are as comparable as possible except for the experience with relevant tiredness states. For instance, it minimizes potential carryover effects of Session 1 on Session 2, such as the anchoring to previous choices, etc.[29]

This way, the impact of experience can be directly inferred from the comparison between the good-bad state wage gap in Session 1 and the good-bad state wage gap in Session 2. If individuals' misperceptions result mainly from projecting their current utilities, the good-bad state wage gap is likely to persist in Session 2. Whereas if individuals have misperceptions mainly from projecting their current (limited) informational states, experience with relevant tiredness states can potentially mitigate the good-bad state wage gap, which provides

[27]In the experiment, the "true" benchmark state when the additional task actually takes place is implicitly in the middle between a good state and a bad state by design. This reflects the specification in the theoretical framework, where the ex ante "true" benchmark is always a linear combination between a good state and a bad state induced by probabilities.

[28]The exact timing in two states also differs because of 8 tasks in the mandatory work block, which is unlikely to be an salient feature when participants are overwhelmed by these aversive tasks. The difference in exact timing is also unlikely to matter for WTA_future because 8 tasks take relatively little time compared to the delay of two days.

[29]The treatment variation in tiredness states and the overall structure in Session 2 is kept nearly identical to Session 1, thus, alternative explanations such as mechanical effects or salience (Bordalo et al., 2013) can also be evaluated. I provide more discussion on all competing explanations in Section 4.3.

immediate implications on how the identification issue can be addressed.

### 3.4 Other Variables of Interests

At the end of Session 2, I elicit several important variables with incentives to shed light on mechanisms, including incentivized beliefs about own performance (i.e. the average time spent on each task) and incentivized recall of wage choices made in Session 1 as a measure of memory.[30] I also ask participants which decision state they would have preferred, in order to shed light on the normative discussion on which state is welfare-relevant.

In the post-experiment survey, I additionally include a rich set of survey measures and self-reported daily behavior usually associated with time preference to evaluate the extent to which the identification issue also matters beyond my experiment. I will introduce these measures and corresponding results in more detail in Section 5, after presenting the main results on work choices.

### 3.5 Hypotheses and Empirical Roadmap

The empirical analysis is closely tied to the theoretical framework in Section 2. The primary focus of the experiment is to evaluate the empirical relevance of the identification issue outlined in Proposition 2 by testing (i) whether the variation in current decision states can produce state-dependent reservation wages about future work as described in Proposition 1; and (ii) how the state dependence in turn can translate into problematic inferences about time preferences when ignoring decision makers' misperceptions. For instance, the analyst can come up with very different estimates of time preference parameters over the same time horizon depending on the exact state in which choices are made, as described in Proposition 3. The pre-registered primary hypotheses are summarized as follows:

- **Hypothesis 1.1 (Baseline state dependence):** *participants in Session 1 currently in a good state predict themselves to be less tired and ask for lower wages for additional work than those currently in a bad state, both when the additional task takes place later today ($WTA^{good}_{today1} < WTA^{bad}_{today1}$) and in two days ($WTA^{good}_{future} < WTA^{bad}_{future}$).*

- **Hypothesis 1.2 (Non-identifiability of time preferences):** *an analyst who infers time preferences over the same time horizon from the ratio $WTA_{future}/WTA_{today}$ will estimate a higher (lower) intensity of time discounting if the advance choice, $WTA_{future}$, is made in a good state (bad state).*

While above main hypotheses hold independently of the specific source of misperceptions, the experiment is further designed to disentangle mechanisms underlying potential state-dependent wage choices in order to inform policy interventions. As described in Section 2.2, while projection of informational states can be

[30] Participants receive £0.5 for each approximately correct guess or recall.

mitigated by having more state-relevant experience, projection of utility is not sensitive to experience with relevant states. Because participants in both states in Session 2 will all have more state-relevant experience compared to in Session 1, as long as informational projection plays a role in participants' misperceptions, state dependence captured by the good-bad wage gap is predicted to be smaller in Session 2 than in Session 1, as described in Corollary 2. The pre-registered secondary hypothesis is summarized as follows:

- **Hypothesis 2 (Mitigation of informational projection by experience):** *the magnitude of state dependence captured by the good-bad wage gap in Session 2 is smaller than the gap in Session 1 if projective misperceptions are informational* ($WTA^{bad}_{today1} - WTA^{good}_{today1} \geq WTA^{bad}_{today2} - WTA^{good}_{today2}$).

### 3.6 Sample and Procedures

The two-part longitudinal experiment was conducted in late March 2021 on Prolific, a UK-based online recruitment platform. Those with a past approval rate above 95% are allowed to take part. The study was advertised with generic information about task completion. To minimize the impact of different devices on the task completion, the experiment is advertised as desktop-only. Those who try to participate via mobile devices are detected and screened out automatically.

Potential participants are informed that the experiment contains two sessions with similar length and payment structure (a fixed payment of £4 for each session that takes around 45 minutes, plus the possibility to earn additional bonus depending on their choices) and that they will receive an additional completion payment of £10 if they complete both sessions. Both sessions are posted at the same time on the corresponding day. Participants who have completed Session 1 will be invited to take part in Session 2 in two days and they have a ten-hour window to complete Session 2 after it is posted.

Among a total of 452 participants who completed Session 1, 430 of them completed both sessions and thus constitute the final sample size.[31] The pre-registered sample size is based on the power calculation with an effect size of $d = 0.2$ and statistical power of 0.8 when using a one-sided test at the 0.05 significance level. Among the full sample of $n = 430$, 217 participants are randomized into the bad state in Session 1, whereas 213 participants are randomized into the good state. Recall that around 5% of the participants in Session 1 have the additional task taking place at the end of Session 2 as determined by their reservation wage provided in Session 1, they thus do not report another wage in Session 2. As a result, there are 408 wage choices reported in Session 2, 204 of which are reported in a good state and the other 204 are reported in a bad state.[32] Participants are balanced across characteristics recorded by Prolific, as in Table D.1 in Appendix D.

[31] The attrition rate is 5% across sessions in the data collection, which is quite low relative to typical studies implemented online, presumably thanks to all the design efforts to reduce dropout. The number of those who drop out is balanced across treatments.

[32] Among the full sample of $n = 430$, less than 5% of observations (4.4% for WTA1, 4.9% for WTA2, 4.7% for WTA3) are excluded from the analysis for making inconsistent choices by switching more than once in the choice list for each wage.

# 4 Main Results On Work Choices

In this section, I first present the result that the experimental manipulation successfully generates variations in marginal disutilities between the good state treatment and the bad state treatment. I then proceed with empirical analysis following the roadmap described in Section 3.5. Focusing on reservation wages reported in Session 1, I first show that participants ask for state-dependent wages for future additional work at different points in time and that this state dependence leads to misinference of time preference parameters, in support of Hypothesis 1.1 and Hypothesis 1.2. I then move to reservation wages and predictions reported in Session 2 and show that participants' state-dependent wages are most likely driven by incorrect projection of informational states due to a lack of learning from the relevant tiredness state, consistent with Hypothesis 2. Finally, I conduct additional analyses to exclude other explanations.

## 4.1 First Stage in State Variation

I first present the first stage result as a manipulation check. Table 1 shows the average marginal disutility as captured by self-reported tiredness at the prospect of completing one additional task on a scale of 1 to 7 in different treatment states in both sessions.

Table 1: Marginal Disutility on Average

| | Treatment State | | | |
|---|---|---|---|---|
| | Good State | Bad State | Difference | $p$-value |
| *Session 1* | | | | |
| Self-reported tiredness (from 1 to 7) | 3.31 | 5.06 | -1.74 | $< 0.001$ |
| | (0.11) | (0.08) | (0.14) | |
| N of observations | 213 | 217 | | |
| *Session 2* | | | | |
| Self-reported tiredness (from 1 to 7) | 2.58 | 4.19 | -1.61 | $< 0.001$ |
| | (0.10) | (0.12) | (0.15) | |
| N of observations | 204 | 204 | | |

*Notes*: Standard errors are in parentheses. $p$-value is based on one-sided t-test.

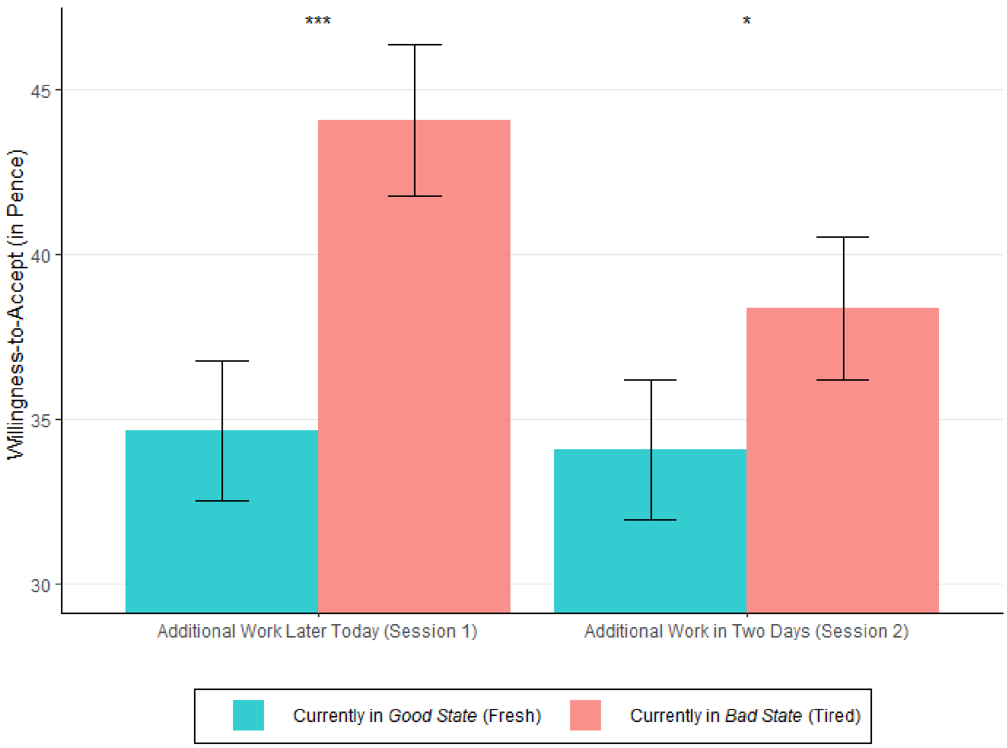
Participants in the bad state treatment indeed anticipate greater tiredness when completing one additional task than those in the good state treatment ($p < 0.01$; one-sided t-test used throughout the analysis given the pre-registered directional predictions unless otherwise noted). This indicates that the experimental manipulation in whether choices are made at the beginning or at the end of mandatory work succeeds at generating state variation in marginal disutility as specified in the theoretical framework.[33]

[33]The manipulation does not impact participants' decision quality (as measured by the number of errors or the completion time), which I also discuss later when evaluating alternative explanations.

## 4.2 State-Dependent Wages and Seemingly Time Preferences

Turning to the substantive results, I first examine whether, consistent with Hypothesis 1.1, participants in Session 1 in the good state treatment ask for lower reservation wages for future work than those in the bad state treatment. Figure 3 shows the average reservation wages for future additional work either later in the same session or in two days in the next session, requested in either the good or the bad state treatment in Session 1.

Figure 3: Participants' Wage Choices Made in Session 1




*Notes*: The left two bars show participants' WTAs reported in Session 1 for one additional task at the end of Session 1 in different treatment conditions. The right two bars show participants' WTAs reported in Session 1 for one additional task at the end of Session 2 in different treatment conditions. Error bars represent standard errors.

Holding constant the time horizon, reservation wages for future additional work highly depend on the current state in which participants make wage choices. When the additional work takes place later today, participants in the good state treatment ask for 21 percent lower wages than those in the bad state treatment ($\text{WTA}^{\text{good}}_{\text{today}} < \text{WTA}^{\text{bad}}_{\text{today}} : p < 0.01$); whereas when the additional work takes place in two days, participants in the good state treatment ask for 11 percent lower wages than those in the bad state treatment ($\text{WTA}^{\text{good}}_{\text{future}} < \text{WTA}^{\text{bad}}_{\text{future}} : p = 0.08$).[34]

As a result, this state dependence can lead to misinference of time preference parameters, consistent with

[34] The result is not driven by participants at the extreme end as it is robust to the exclusion of top-censored participants who choose not to do the additional task for each level of wages in the choice list thus potentially indicate a WTA higher than £1.

Proposition 2. To illustrate, suppose an analyst wants to infer time preferences from the observed WTA choices in the experiment. He does not know the data structure of the experiment and only knows that some choices are for work later today and some choices are for work in two days. Other important factors, such as utility curvature, are assumed away for simplicity. What the analyst can to do is to select two wage choices for work at different points in time, one for work later today and one for work in two days. When the analyst does not take into account decision makers' state-dependent misperceptions, if he happens to pick an immediate choice in the bad state ($\text{WTA}^{\text{bad}}_{\text{today1}} = 44$) and an advance choice in the good state ($\text{WTA}^{\text{good}}_{\text{future}} = 34$), he will infer the time preference parameter (in this case, two-day discount factor) to be 0.77, which is the smallest possible number he can get from this setting. Alternatively, if he happens to pick an immediate choice in the good state ($\text{WTA}^{\text{good}}_{\text{today1}} = 34$) and an advance choice in the bad state ($\text{WTA}^{\text{bad}}_{\text{future}} = 38$), he will infer the time preference parameter to be 1.11, which is the largest possible estimate he can get from this setting. That is, the parameter estimates inferred from wage choices over the same time horizon from the same dataset can differ substantially, in this case, by up to 44 percent,[35] depending on the exact states in which choices are made, consistent with Hypothesis 1.2.[36]

I formalize the results from Figure 3 in Table 2, combining wage choices for additional work at both points in time made in Session 1 to estimate how average willingness-to-accept changes with the decision state and the time horizon. I estimate the average treatment effect using the following regression,

$$WTA_{it} = \delta_0 + \delta_1 \mathbb{1}[\text{Bad State}]_i + \delta_2 \mathbb{1}[\text{In Two Days}]_t + \delta_3 \mathbb{1}[\text{Bad State}]_i \mathbb{1}[\text{In Two Days}]_t + \epsilon_{it}$$

where $WTA_{it}$ is participant $i$'s willingness-to-accept for the additional task that takes place at time $t$. Regressions in Column (1)-(3) are from OLS estimations whereas regressions in Column (4)-(6) are from Tobit estimations that account for the censoring at 0 and 100.

Regressions show that the reservation wages in Session 1 are mainly driven by state variation, instead of time variation. Column (1) shows the overall state dependence–participants ask for higher wages in the bad state treatment. Column (2) shows that participants ask for a lower wage for an additional task in two days as opposed to an additional task later today, seemingly consistent with time discounting. However, Column (3) suggests that once the state variation is controlled, there appears to be no state-*in*dependent discounting in my setting.[37] All these patterns are qualitatively similar when censoring is accounted for in Tobit models.

[35] While the analyst in reality may have more information about the data structure, such that he does not have to use data in this specific way, 44 percent can be seen as an upper bound of possible variations of the two-day discount factor. The difference can only be bigger once compounding is taken into consideration.

[36] A similar pattern at the individual level is shown in Figure E.1 in Appendix E.1, which displays the cumulative distribution of the two-day discount factor inferred from $\text{WTA}_{\text{future}}/\text{WTA}_{\text{today2}}$. This individual-level illustration is not possible using $\text{WTA}_{\text{today1}}$ because $\text{WTA}_{\text{future}}$ and $\text{WTA}_{\text{today1}}$ are made in the same state within subject.

[37] The self-reported predicted tiredness is also highly state-dependent. Participants in the good state predict themselves to be less tired from the additional task than those in the bad state, regardless of whether the additional task takes place in two days or later today ($p < 0.01$). Similar figures and regressions are provided in Appendix E.

Table 2: Participants' Wage Choices Made in Session 1

| | Dep Var: WTA, 0 to 100 Pence | | | | | |
|---|---|---|---|---|---|---|
| | (1) OLS | (2) OLS | (3) OLS | (4) Tobit | (5) Tobit | (6) Tobit |
| Bad State | 6.868** | | 9.420*** | 9.088** | | 12.52*** |
| | (2.939) | | (3.124) | (3.597) | | (3.849) |
| In Two Days | | -3.206*** | -0.555 | | -3.658*** | -0.0483 |
| | | (0.891) | (1.197) | | (1.118) | (1.490) |
| Bad State × In Two Days | | | -5.147*** | | | -6.933*** |
| | | | (1.768) | | | (2.223) |
| Constant | 34.36*** | 39.45*** | 34.64*** | 30.22*** | 36.68*** | 30.26*** |
| | (2.023) | (1.580) | (2.117) | (2.596) | (1.943) | (2.730) |
| Observations | 820 | 820 | 820 | 820 | 820 | 820 |
| R-squared | 0.012 | 0.003 | 0.016 | - | - | - |

*Notes*: Standard errors clustered at the individual level. "Bad State" is a binary indicator for the bad state treatment. "In Two Days" is a dummy variable indicating that the wage is for additional task that tasks place in two days instead of later today. In Column (4)-(6), 138 WTA choices are left-censored at 0 and 22 WTA choices are right-censored at 100. *** p<0.01, ** p<0.05, * p<0.1

Notably, in Columns (3) and (6) there is a significantly negative interaction term between the bad state and the time delay, consistent with the patterns in Figure 3. This can either be interpreted as state-dependent discounting or evidence that misperceptions about states may differ across different time horizons. The data is consistent with the latter interpretation and supports a horizon-dependent effect of misperceptions.[38] I provide more discussion of this point in Appendix E.2.

## 4.3 Mechanisms Driving the State Dependence

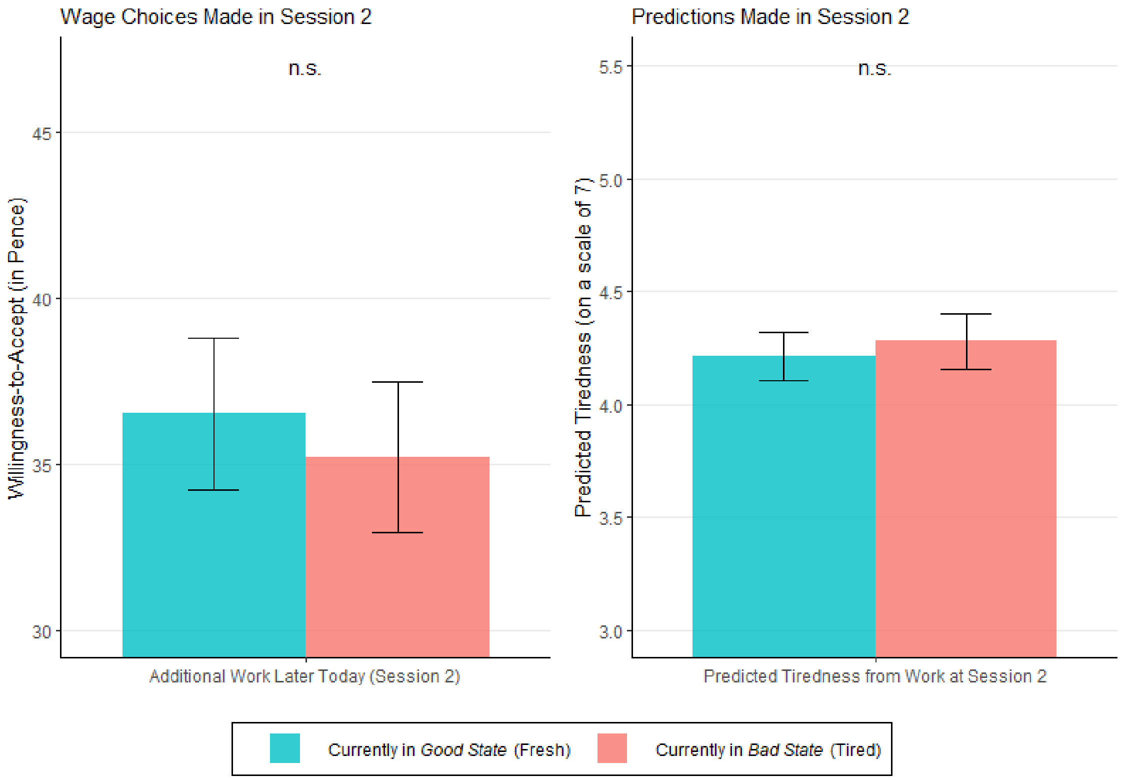
### 4.3.1 Projection of Informational States vs. Projection of Disutility

I further investigate mechanisms driving the state dependence by examining choices in Session 2. Recall that participants in Session 2 make a similar decision ($\text{WTA}_{\text{today2}}$) about the additional work in the same day as in Session 1 ($\text{WTA}_{\text{today1}}$), but with experience of all tiredness states. Figure 4 shows the average reservation wages (left) for and predicted tiredness (right) from the additional task at the end of Session 2 in either the good or the bad state treatment in Session 2. The figure on the left shows that when deciding again about an additional task later today in Session 2, participants in the good state treatment ask for similar wages ($p = 0.66$) as those in the bad state treatment. That is, the state dependence in Session 1 is completely mitigated by experiencing relevant tiredness states on aggregate, in line with Hypothesis 2.

Recall that according to the theoretical framework of projective misperceptions, projection of current

[38]This horizon-dependent effect of misperceptions is potentially consistent with subadditivity in intertemporal choice (Dohmen et al., 2017; Read, 2001).

Figure 4: Wage Choices and Predicted Tiredness Made in Session 2 with Experience




*Notes*: The left figure shows participants' WTAs for one additional task at the end of Session 2 reported in Session 2 in different treatment conditions. The right figure shows participants' predicted tiredness from additional work at the end of Session 2 reported in Session 2 in different treatments. Error bars represent standard errors.

tiredness states is likely to be informational, thus responsive to experience with relevant tiredness states; whereas projection of current utilities can persist regardless of experience. The dramatic impact of experience documented in Figure 4 thus suggests that state-dependent wage choices observed in Session 1 are mostly driven by the fact that participants in Session 1 mistakenly project their tiredness states due to a lack of experience with relevant tiredness states, rather than the non-informational projection of utilities. That is, projective misperceptions in my setting are mainly informational as opposed to non-informational.

In addition, the comparison between participants' self-reported predictions of the actual tiredness state and the actually perceived tiredness in the neutral benchmark suggests that their state-dependent predictions in *both* states are likely to be *mistakes*. As shown in Figure E.2 in Appendix E.3, participants in the good state in Session 1 predict themselves to be less tired than those in the bad state ($p < 0.01$), and their predicted tiredness in both states differ from the actually perceived tiredness in the neutral benchmark. This indicates that their predictions in Session 1 about the actual tiredness states are indeed misperceptions. Moreover, as shown in Figure E.3 in Appendix E.4, in Session 2, those in the good state revise their predictions upwards by predicting themselves to be more tired than those who predict in Session 1 (two-sided t-test: $p < 0.01$), whereas those in the bad state treatment revise their predictions downwards by predicting themselves to be less tired than those who predict in Session 1 (two-sided t-test: $p = 0.08$). As a result, participants in Session 2 predict themselves to be equally tired ($p = 0.34$) and their predictions are getting closer to the neutral benchmark state. This means that not only the state dependence in wages disappears in Session 2, but prediction mistakes are also corrected by having more experience, in line with informational state projection.

#### 4.3.2 Interpreting Informational State Projection

**Lack of experience-based learning** (Malmendier, 2021) in Session 1 can result in the projection of informational states, where the learning is about one's own internal tiredness state.[39] In particular, the experience-based learning about the tiredness state in Session 2 needs to happen in both states. Under this interpretation, participants in a good state in Session 2 learn from the bad state experience that the one additional task can be more unpleasant than they think, whereas those in a bad state in Session 2 learn from the good state experience that the one additional task can be more pleasant than they think.[40]

**Cue-based recall** as in Bordalo et al. (2020b) can also potentially account for both state dependence in Session 1 and the decrease of state dependence in Session 2. Specifically, the state variation in both sessions can be seen as contextual cues that may impact valuations through interacting with retrieval from associative memory[41] using their language. Their framework can allow the mitigating impact of experience in Session 2 because contextual cues can interact with remembered conflicting experience in the past. They predict that the effect of contextual cues (i.e. tiredness state in the current session) becomes less pronounced among experienced than among inexperienced, because the contextual cue and the retrieved conflicting experience (i.e. tiredness state experienced in the previous session) interfere with each other in the recalling process. This interpretation can coexist with the mechanism through projective misperceptions, because their framework offers a memory-based foundation that unifies a wide range of behavioral patterns, including projection bias (Loewenstein et al., 2003).

[39]Experience-based learning differs from the traditional notion of learning that is purely based on information about the external environment. For instance, Simonsohn et al. (2008) show that information from direct experience is weighted more heavily than observed information. Recall that participants in both treatments in Session 1 have completed at least one task when reporting reservation wages for one additional task later, and that the experiment has an extreme transparent and simplistic structure, both of which limit the scope of learning from the objective environment.

[40]In contrast, learning from the accumulated task disutility is not consistent with the data. Recall that participants in the bad state treatment have completed more tasks than those in the good state treatment (nine vs. one), and the state variation in tiredness is of a cumulative nature. This means that differential learning could also account for the state-dependent work choices in Session 1, if participants in the bad state treatment can learn about the true aversiveness of the tasks through completing more tasks while those in the good state treatment do not learn. However, this interpretation implies that the bad state wage in Session 1 reflects the actual aversiveness of the tasks. Therefore, participants in both treatments would have asked for similarly high wages in Session 2 since everyone has learned about the actual aversiveness after Session 1. That is not the case in the data. As shown in Figure E.4 in Appendix E.5, having experience with relevant tiredness states mitigates the state dependence detected in Session 1, but mostly by reducing bad state wages, not by increasing good state wages.

[41]See Kahana (2012) for a textbook description and Enke et al. (2020) for empirical evidence.

#### 4.3.3 Other Explanations

**Learning from task performance** is inconsistent with observed state dependence. Recall that participants are paid a fixed amount for the mandatory work instead of a piece rate, thus there is not much to learn except for the average time spent on each task. Participants can potentially use their task completion time as a reference to calculate the opportunity cost for one additional task. While there is an improvement of the completion time over the course of the experiment, I find no evidence of participants' task performance predicting their work choices or predicted tiredness, as shown in Table E.2 in Appendix E.6. If anything, participants in the bad state treatment are, on average, faster at completing one task than those in the good state, presumably due to more task experience at the time when making decisions. If participants in the bad state treatment use the faster completion as a reference for their reservation wages, learning about task performance would predict that those in the bad state ask for lower wages than those in the good state, which is a counter hypothesis.

**Poor decision quality due to fatigue** can also be ruled out as an explanation. First, there is limited evidence that participants in the bad state treatment are making worse decisions than those in the good state treatment. Participants in both states on average spend similar amounts of time on making wage decisions (two-sided t-test: $p = 0.77$ in Session 1; $p = 0.42$ in Session 2) and make similar numbers of errors when performing the task (two-sided t-test: $p = 0.93$ in Session 1; $p = 0.70$ in Session 2). Participants in the bad state are even better at performing tasks than those in the good state treatment in terms of completion time (two-sided t-test: $p < 0.01$ in Session 1; $p = 0.04$ in Session 2). Second, even if decision quality were to have an impact and partly drives the state dependence in Session 1, a similar effect would be expected to arise in Session 2; this is not the case.

**Cognitive noisiness** as captured in the noisy value representation framework (Gabaix and Laibson, 2017; Gershman and Bhui, 2020) can predict state dependence in different directions but neither is entirely consistent with the data. In their framework, agents mentally simulate the value of future events based on unbiased priors and noisy signals about own values. This framework thus predicts that higher noises can lead to "as-if" impatient behavior, i.e., lower reservation wages in my experiment. However, it is not clear a priori whether one should expect higher noise in a bad state or in a good state. Higher noise can be associated with a bad state because participants are tired and have depleted cognitive resources; on the other hand, higher noises can also be associated with a good state because participants in a bad state can learn about the task disutility more precisely from completing more tasks. Tests on the equality of variances of elicited wages do not support any of these relationships (two-sided sd-test: $p = 0.14$ and $p = 0.59$ in Session 1; $p = 0.93$ in Session 2). Moreover, even if the cognitive noisiness were to partly drive the state dependence in Session 1, the similar effect would have also appeared in Session 2.

**Mechanical effects** such as a convex cost function or inherent state-dependent discounting are unlikely to explain the data in both sessions for two reasons. First, if the state dependence in Session 1 is driven by mechanical effects, the same pattern would have also appeared in Session 2 given that the experimental manipulation is identical. That is not the case in Session 2. Second, the 2-minute rest between mandatory work and the potential additional work in each of the two sessions is designed to reset the effort cost. Even if this 2-minute rest may not be sufficient, the cost of effort should be identical for participants in both treatments, because all of them have completed the same amount of mandatory work before the additional task.

**Short-term discounting** is unlikely to explain state dependence in the wage choices made in Session 1. Admittedly, the exact timing in two states differs because of eight tasks in the mandatory work block given that choices are elicited after one task in the good state and after nine tasks in the bad state. This means that short-term discounting would in principle lead to higher wages in the bad state treatment because the time span between the choice and the task is slightly shorter (by about thirty minutes). Nevertheless, the same pattern would have also appeared in Session 2 given that the short-term discounting remains identical. In addition, the state dependence in the wage choice made in Session 1 about the additional task in Session 2, $\text{WTA}_{\text{future}}$, is unlikely to be influenced by this difference since eight tasks take relatively little time compared to the delay of two days.

**Consistency motives** can in theory explain the decrease of state dependence in Session 2 but are not supported by the data. If participants in Session 2 tend to remain consistent with their earlier choices about the same additional task made in Session 1, because of stratified randomization, half of participants in the good state treatment in Session 2 would increase their wages to be consistent with their bad state wages reported in Session 1. The same reasoning would also apply to those who are in the bad state treatment in Session 2 but were in the good state treatment in Session 1. If this interpretation were the case, the wages reported in Session 2 would have depended on the treatment states in Session 1. I do not find statistically meaningful heterogeneous effects supporting this interpretation, as wages reported in both states in Session 2 do not differ by states in Session 1 (two-sided t-test: $p = 0.41$ in the good state and $p = 0.20$ in the bad state). In addition, participants need to remember their early wage choices in order to be consistent. Recall that I elicit their recall of wage choices made in Session 1 during the post-experiment survey with incentives. Only 17% of participants recall exactly correctly and the percentage increases to 57% only if I allow approximate recalls (within $\pm 5$). I find no evidence that those who remember early choices exhibit less state dependence than those who do not remember (two-sided t-test: $p = 0.33$ for those who remember and $p = 0.52$ for those who do not remember).[42]

**Salience** as captured in Bordalo et al. (2013) is one of the mechanisms that can account for data in previous observational studies on the impact of weather on particular consumption goods (e.g., Busse et al.,

[42] The memory measured by incentivized recall has a caveat because it is endogenously determined. For instance, those who are able to remember the exact wage choices may find those tasks not hard to do thus have lower reservation wages.

2015; Chang et al., 2018; O'Donoghue and Vogelsang, 2007). It can explain state dependence observed in Session 1 if participants consider the tiring attribute of the additional work more salient in the bad state treatment than in the good state treatment. However, as salience does not differ between Session 1 and Session 2 across treatments, state dependence driven by salience would have also appeared in Session 2.

Finally, none of above alternative explanations predicts heterogeneous state dependence with respect to the magnitude of projective misperceptions, which I discuss in more detail in the next section.

# 5 Projective Misperceptions Beyond The Experiment

In this section, after demonstrating the empirical relevance of the identification issue in my experiment, I investigate the heterogeneity of projective misperceptions based on the post-experiment survey. At the end of the experiment, participants complete a survey eliciting several measures, including a new survey-based measure of projective misperceptions, several widely used measures of time preferences that have been previously validated, and several self-reported measures of daily behavior usually associated with time preferences: health, procrastination, and saving. Exploratory analyses on these variables show that projective misperceptions can be seen as an individual trait and are potentially relevant for predicting intertemporal choice both within and beyond the experimental context.

I first describe all the variables in detail and then examine the extent to which the survey-based measure of projective mispercetions can predict work choices in the experiment. This constitutes a direct test of the comparative statics prediction that the state-dependent behavior is increasing in the magnitude of misperceptions, as described in Corollary 1. Although the survey measure is arguably noisy and might inevitably pick up something beyond its intended scope, as long as it captures some misperceptions as a meaningful individual trait beyond noise, it allows me to detect potential heterogeneous effects of state dependence with respect to misperceptions. I continue the analysis by exploring the link between this measure and self-reported daily behavior usually associated with time preferences in different contexts. As implied by Corollary 3, projective misperceptions are more likely to play a role in contexts where the magnitude of state variation is larger, presumably in non-monetary contexts, relative to in monetary contexts.

## 5.1 Variables Included in the Post-Experiment Survey

**Survey-based measure of projective misperceptions.** I first develop a new and scalable survey-based measure of projective misperceptions. The measure is designed with an intention to reflect the projection of utility à la projection bias.[43] As reviewed in the original seminal work (Loewenstein et al., 2003), the most

[43] I pre-registered this measure as a measure of projection bias. Nevertheless, the item may reflect both types of projective misperceptions, as projection of utility and projection of informational states are not distinguished in Loewenstein et al. (2003).

prominent evidence in support of the modeling of projection bias is that people tend to *underestimate their adaptation* to life changes. I thus design the measure as the extent to which adaptation is underestimated based on this intuition. Participants are asked to answer the following question, "how good are you at estimating your own ability to adapt to changes or challenges in your life", on a seven-point scale from "always overestimate" to "always underestimate". Those who are more likely to underestimate their adaptation are classified as more prone to projective misperceptions based on this measure.[44]

**Measures of time preferences.** I then include several measures of time preferences that have been previously validated and are widely used in empirical work. The first one is a self-reported measure of patience (Falk et al., 2016) on a seven-point scale. Participants are asked how willing they are to give up something that is beneficial today in order to benefit more from that in the future. The second and the third items are continuous discounting measured from the hypothetical "money-earlier-or-later" paradigm (Chapman et al., 2019) in two time frames. Participants are asked about how much money the experimenter would have to pay them in thirty days so that they would be willing to forego a payment of £100 today and how much money the experimenter would have to pay them in sixty days so that they would be willing to forego a payment of £100 in thirty days. The ratio of 100 and their reported amount in both questions approximate measures of time preference, $\beta\delta$ and $\delta$. The ratio between the two measures then gives a measure of present bias, $\beta$. The last item is a hypothetical survey measure of perceived self-control problems as a proxy for sophistication (Ameriks et al., 2007; John, 2020). Participants are presented with a hypothetical scenario of winning ten restaurant certificates, each of which can be used (once) to receive a "dream restaurant night" in a world without the COVID-19 pandemic. Participants are then asked to come up with the ideal allocation of the ten certificates over two years, the allocation were they to give in their temptation (i.e. the tempted allocation), and the expected allocation. The difference between the tempted allocation and the ideal allocation gives the measure of sophistication.[45]

**Daily behavior associated with time preferences.** I collect a few self-reported measures of daily behaviors often associated with time preferences, in particular, saving behavior, procrastination at work, and preventive health investment. Each particular measure is listed as follows.

- Saving: participants are asked about how much savings they currently have set aside as emergency funds, and select their answers from five levels ranging from "none" to "more than a year's income".
- Procrastination: participants select one out of four statements that best describe the way they plan for work or study, from "start too early" to "wait until the last minute".

[44] I included the full spectrum from overestimation to underestimation for the survey question, instead of using a double-barreled question that directly asks the extent to which one underestimate adaptation. Yet, I have no ex ante expectation on whether many people would overestimate their adaptation or whether those who overestimate would exhibit no projection (i.e. $\alpha(e) = 0$) or anti-projection (i.e. $\alpha(e) < 0$).

[45] As noted in John (2020), this measure may also capture "self-control types" who face a cost of temptation (Toussaert, 2018).

- Preventive health: participants are asked about their intention to get a COVID-19 vaccine when it is mass-produced and open to register for their age group, and select their answers from five levels ranging from "won't get it for sure" to "will get it for sure".[46]

**Other variables.** Other variables for exploratory purposes include self-reported confidence, self-reported risk attitudes, frequency of regretting life choices, and interpersonal empathy (Spreng et al., 2009).

## 5.2 Projective Misperceptions Predict Work Choices In the Experiment

Figure E.5 in Appendix E.7 presents the full empirical distribution of the survey-based measure of projective misperceptions, which does not differ across treatments (two-sided KS test: $p = 0.48$). While participants' self-awareness reflected in this survey measure might seems at odds with the theoretical construct of misperceptions, because the latter presumes that decision makers have little self-awareness, I briefly discuss one interpretation that can reconcile this based on self-knowledge (e.g., Falk et al., 2021) in Appendix E.7.

I first examine whether this survey-based measure of projective misperceptions can predict state-dependent wage choices in the experiment. Table 3 presents results from linear regressions using wage choices in the experiment, where *If_Misperception* is a binary variable indicating that the measured projective misperceptions are above the median response provided by all participants. The positive coefficients of the interaction term between the misperception indicator and the bad state treatment indicator (*Bad State* × *If_Misperception*) in Column (2) suggest that the survey measure has explanatory power for the reaction to states in the experiment. Those who are more prone to projective misperceptions according to the survey measure ask for more state-dependent wages in the experiment than those who are less prone to projective misperceptions according to the survey measure. This is consistent with the comparative statics captured in Corollary 1 of the theoretical framework, which predicts that the magnitude of state dependence is increasing in the degree of projective misperceptions. I conduct similar exercises with other control variables as placebo tests and find no similar relationships, as shown in Table E.3 in Appendix E.8.

This relationship is not strong in terms of statistical significance, potentially because ex ante the the measure is designed with an intention to capture the projection of utility à la projection bias, whereas ex post most of the state dependence (at least in Session 1) is driven by the projection of informational states. This interpretation is consistent with the fact that the relationship is stronger in Session 2 as in Column (6) than in Session 1 as in Column (4).

This validation exercise provides suggestive evidence that projective misperceptions may be a meaningful

[46]The experiment was conducted in March 2021, before the take-up of COVID-19 vaccine became a somewhat political issue. As a sanity check, I include a binary variable indicating whether each participant has got flu shot in the past and find that this binary indicator strongly correlates with COVID-19 vaccine take-up ($p < 0.01$). This means that the COVID-19 vaccine take-up indeed involved health investment consideration at least when the experiment was conducted.

Table 3: Wage Choices Predicted by Projective Misperceptions

| | (1) Pooled | (2) Pooled | (3) Session 1 | (4) Session 1 | (5) Session 2 | (6) Session 2 |
|---|---|---|---|---|---|---|
| Bad State | 4.209* | -0.129 | 6.868** | 3.518 | -1.294 | -7.873 |
| | (2.190) | (3.227) | (2.939) | (4.487) | (3.216) | (4.905) |
| If_Misperception | | -3.902 | | -3.176 | | -5.666 |
| | | (3.516) | | (4.088) | | (4.672) |
| *Bad State × If_Misperception* | | 8.462* | | 6.529 | | 12.78** |
| | | (4.395) | | (5.889) | | (6.502) |
| In Two Days | -3.218** | -3.184** | | | | |
| | (1.559) | (1.557) | | | | |
| Bad State × In Two Days | 0.0640 | 0.0555 | | | | |
| | (2.506) | (2.510) | | | | |
| Experience | -3.533** | -3.477** | | | | |
| | (1.382) | (1.382) | | | | |
| Constant | 37.30*** | 39.41*** | 34.36*** | 36.09*** | 36.51*** | 39.68*** |
| | (1.858) | (2.772) | (2.023) | (3.115) | (2.278) | (3.733) |
| Observations | 1,209 | 1,209 | 820 | 820 | 389 | 389 |
| R-squared | 0.007 | 0.012 | 0.012 | 0.015 | 0.000 | 0.011 |

*Notes*: Standard errors clustered at the individual level. "Bad State" is a binary indicator for the bad state treatment. "In Two Days" is a dummy variable indicating that the wage is for additional task that tasks place in two days instead of later today. "Experience" is a dummy variable indicating choices in Session 2. "If_Misperception" is a dummy variable indicating that the measured projective misperceptions are above median. *** p<0.01, ** p<0.05, * p<0.1

and general trait at the individual level, which can predict meaningful heterogeneous effects in my experiment. It would be valuable in future research to further validate this measure at scale given the limited statistical strength here. Once validated, this measure has the potential to be included in surveys more broadly to explain a broad range of economically-relevant behaviors in intertemporal settings given its simple and easily scalable nature. Including this survey measure as a control variable would also be one way of tackling the identification issue caused by projective misperceptions.

### 5.3 Projective Misperceptions Predict Daily Behavior Beyond the Experiment

Recall that I also include several measures of time preferences that are widely used in empirical work, as well as several self-reported measures of daily behaviors that are often associated with time preferences. Table 4 presents results from linear regressions where I predict those daily behaviors using both the survey measure of projective misperceptions and measures of time preferences.

Results show that preventive health investment (in the context of COVID-19) is better explained by projective misperceptions than by time preferences as shown in Column (1), that procrastination is explained by both projective misperceptions and time prefererences as shown in Column (2), and that saving is only explained by time preferences rather than by projective misperceptions as shown in Column (3). One potential

Table 4: Predicting Self-reported Daily Behavior: Misperceptions vs. Time Preferences

| | (1)<br>Covid | (2)<br>Procrastination | (3)<br>Saving |
|---|---|---|---|
| If_Misperception | 0.279*** | 0.171** | 0.104 |
| | (0.102) | (0.0843) | (0.101) |
| If_Patient | -0.0220 | 0.00149 | 0.140 |
| | (0.121) | (0.101) | (0.116) |
| BetaDelta_Lower | -0.0650 | 0.0135 | -0.224** |
| | (0.110) | (0.0907) | (0.110) |
| Beta_Lower | 0.110 | -0.0442 | -0.0523 |
| | (0.110) | (0.0921) | (0.108) |
| If_Sophisticated | 0.147 | 0.146* | -0.101 |
| | (0.102) | (0.0840) | (0.101) |
| Constant | 0.841*** | 0.290** | 2.448*** |
| | (0.165) | (0.131) | (0.151) |
| Observations | 430 | 430 | 430 |
| R-squared | 0.028 | 0.018 | 0.017 |

*Notes*: This table reports results from OLS regressions of three self-reported variables on measures of preference misperceptions and time preferences in binary form. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

interpretation of these results is that non-monetary contexts presumably involve more state variation than monetary contexts, therefore, projective misperceptions are more likely to play a role in the former than in the latter. This is consistent with the comparative statics characterized in Corollary 3 of the theoretical framework, which predicts that the severity of the identification issue caused by projective misperceptions is increasing in the magnitude of state variation across contexts.

One may be concerned that these results are mainly due to the fact that some of the above measures are highly correlated. Additional analyses on the predictive power of each of these measures separately indicate that this is not the case. First, most of the estimated coefficients stay qualitatively similar when analysed in isolation. Second, as presented in Table E.4 and Table E.5 in Appendix E.9, only very few measures elicited using similar questions are correlated. Lastly, regressions using these measures in continuous forms also share a similar pattern, as shown in Table E.6 in Appendix E.10.

# 6 Discussion and Implications

In this section, I first discuss how the identification issue caused by projective misperceptions can be used to meaningfully explain several puzzles in time preferences. Then, as the way forward, I discuss potential solutions to the identification issue and policy interventions to address *seemingly* self-control problem.

## 6.1 Related Puzzles in Time Preferences

**Heterogeneity and Domain Specificity of Time Preferences.** It has been documented that estimates of time preference parameters are extremely heterogeneous (Chapman et al., 2019; Falk et al., 2018; Frederick et al., 2002; Havránek et al., 2021; Imai et al., 2020), can be domain specific (Chapman, 1996; Ubfal, 2016), and malleable (Ebert and Prelec, 2007). Sometimes these estimates can even be unstable within individuals (Meier and Sprenger, 2015) and systematically change with age (Huffman et al., 2019). These regularities are consistent with both my theoretical framework and my empirical results, which indicate that the identification issue caused by projective misperceptions is inherently dependent on the magnitude of state variation in the context as in Corollary 3 and on experience with unfamiliar states as in Corollary 2. When the analyst fails to take into account misperceptions, "revealed" time preferences can be highly domain specific and heterogeneous, even if the "true" underlying time preferences are domain general and homogeneous.[47]

**Higher Discounting Estimated in the Non-Monetary Domain.** Previous research also documents that time preference parameters estimated in non-monetary domains (e.g., effort) are generally smaller than those estimated in monetary domains (Cohen et al., 2020).[48] As in Corollary 3, in contexts with state variation, behavior in a usual "good" state may lead to overestimation of time preference parameters. If non-monetary settings involve larger state variation than monetary settings, presumably because people's preferences for money are stable and experiencing utility from spending money is a familiar activity, a more severe identification issue will naturally lead to more discounting estimated in non-monetary settings.[49]

**Correlations with Daily Behavior.** Despite the conceptual importance of time preferences, the direct evidence on the association between time preference measures and intertemporal behavior in real life at the individual level is mixed. A handful of papers examine the correlation between experimental measures of time discounting and saving contracts (Ashraf et al., 2006), smoking (Khwaja et al., 2007), self-reported health behavior (Chabris et al., 2008), credit card borrowing (Meier and Sprenger, 2010, 2012), vaccination (Andreoni et al., 2016), and procrastination in tax-filing behavior (Martinez et al., 2017), etc. Correlations in these studies are not strong, with a few exceptions (e.g., Sutter et al., 2013), and the strongest evidence by far is in the monetary domain (e.g., Ashraf et al., 2006). These combined findings are also in line with my findings on the predictive power of misperceptions and time preferences across domains. Taken together, the evidence is consistent with the interpretation that projective misperceptions are more likely to play an important role in intertemporal choice in non-monetary settings, which presumably involves larger state variation.

[47] Relatedly, Echenique et al. (2020) characterize an axiomatic foundation for discounting-based time preferences and show that less than half of the subjects in existing experiments pass their non-parametric revealed-preference tests, which indicates that intertemporal choice may also be driven by factors beyond time preferences.

[48] See Imai et al. (2020) for a meta-study of present bias estimates, and Havránek et al. (2021) for a meta-study of individual discount rate estimates.

[49] Of course, the misperception-based interpretation is not mutually exclusive from other explanations, e.g. ones that are based on the fungibility of money (Augenblick et al., 2015; Cohen et al., 2020; Cubitt and Read, 2007).

**Persistent Naiveté.** Naive beliefs are characterized by the failure to anticipate future present-focused time preferences, i.e. the lack of self-control (O'Donoghue and Rabin, 1999). While both full naiveté and partial naiveté are documented in a range of empirical settings (e.g., Augenblick and Rabin, 2018; DellaVigna and Malmendier, 2006), it remains an open puzzle why people are persistently naive and do not learn from their past experiences (Ericson and Laibson, 2019).[50] Although this paper does not investigate naiveté specifically, it sheds light on this puzzle as projective misperceptions generate behavior that can be misattributed to time preferences and misperceptions by definition preclude the existence of self-awareness. Relatedly, Le Yaouanq and Schwardmann (2020) argue that persistent naiveté is not a result of a fundamental inferential bias in learning by documenting experimentally that people are fully capable of updating their beliefs about their own self-control. They find that people persistently underestimate their future learning. Notably, this finding is consistent with my interpretation in the sense that underestimation reflects the intuition of the survey measure of projective misperceptions as described in Section 5.1.

**Weak Demand for Commitment Devices.** Commitment devices represent one potential policy instrument to address the underinvestment problem caused by present-focused time preferences (Laibson, 1997; O'Donoghue and Rabin, 1999).[51] If the decision maker worries that she invests too little on the activity with delayed benefits because of her present focus, she may want to restrict her choice set as a self-commitment. However, if the underinvestment is instead driven by projective misperceptions instead of present-focused time preferences, the decision maker will not demand the commitment at the first place.[52] Even if she takes up the commitment for other reasons such as to reduce temptation (Gul and Pesendorfer, 2001; Toussaert, 2018) or by mistakes (Carrera et al., 2021), she may not follow through afterwards because she may start to realize that the commitment is not beneficial when states vary.[53] These combined interpretations are consistent with the real world observations that the use of commitment devices is rare in practice (Laibson, 2015).

## 6.2 Implications for Applied Work

**Inference about the Role of Time Preferences** can be confounded by projective misperceptions depending on how much state variation is involved in the particular domain. This also concerns experimental measurement of time preference parameters ($\delta$ and/or $\beta$). To address the identification challenge, my results shed light on three potential solutions that leverage the impact of experience, additional elicitation in the alternative state, and the measure of projective misperceptions as an individual trait, respectively.

[50]See Ahn et al. (2019) and Ahn et al. (2020) for an axiomatic characterization of naiveté based on the costly self-control in the sense of Gul and Pesendorfer (2001).

[51]See Bryan et al. (2010) for a thorough review of commitment device.

[52]The existence of naive beliefs is often used as a theoretical explanation for the absence of commitment. Yet, a few recent empirical studies document a surprising non or negative correlation between commitment demand and sophistication (Carrera et al., 2021; John, 2020; Sadoff et al., 2020), calling for further investigation of this theoretical relationship.

[53]The intuition is similar to ones that attribute the low commitment demand to uncertainty about the future opportunity cost of time or preferences (John, 2020; Laibson, 2015).

The first solution stemming from the comparison between Session 1 and Session 2 is to provide experience and familiarity with states, not only information, to help mitigate participants' misperceptions before the main elicitation. This solution has a few limits because the impact of experience can be very setting-specific. It is possible that participants in my experiment are able to correct their misperceptions quickly only because the experimental setting is highly transparent and straightforward. While I find most state dependence in my setting can be attributed to state projection that can be mitigated by experience, experience-independent utility projection can exist in other contexts. One other limit of the solution is that providing experience can simply be infeasible in many real life settings, where intertemporal choices are only made once a while or even once-in-a-life. For instance, when deciding on retirement investments and long-term preventive health measures, people at young ages do not have a separate opportunity to experience being old. When it is impossible to obtain state-relevant experience, other efforts will be needed to address the identification issue.

The second solution is to additionally elicit choices in an alternative unusual state on top of the usual state choices across groups of people. This can be achieved by including a separate group of people in a different state. For instance, both Augenblick and Rabin (2018) and Fedyk (2021) randomize choice states (warm up time hours in their cases) before actual effort allocations. This way, they can either balance choices in all states on average so that the effects of different states cancel out each other,[54] or add information about states into a joint estimation to separate out whether choices stem from time preferences or misperceptions.

The last solution is to add the survey measure of projective misperceptions as a control variable, together with other individual traits such as patience, risk attitudes, etc. The validation exercise using experimental behavior in this paper provides the first piece of suggestive evidence supporting misperceptions as an individual trait. It thus opens up avenues for future research on its robustness, stability, heterogeneity, and predictive power in explaining other important economic behavior.

**Time-Preference-Based Policy Interventions** that aim to address underinvestment behavior can be less effective than previously thought. At least in settings involving large state variation, interventions such as the use of commitment devices may need to be revisited because suboptimal behavior in these settings could also be impacted by misperceptions instead of time preferences. This implication is largely consistent with several recent studies documenting that people take up welfare-reducing commitment contracts and do not follow through (Bai et al., 2020; John, 2020) and that people take up commitment devices with conflicting goals at the same time (Carrera et al., 2021).

Inspired by findings in this paper, new interventions targeting projective misperceptions may be potentially desirable. When policy makers are worried about suboptimal behavior in a usual "good" state, such as underinvestment for future benefits or other *apparent* self-control problems, hypothetical or vicarious experience

[54] I show theoretically that a full cancellation only happens under certain condition, as in Corollary 4 in Appendix A.2.

regarding alternative "bad" states have the potential to be part of an effective intervention.[55] More specifically, if projective misperceptions are due to a lack of ability to imagine the concrete future, training programs on such an ability can be another useful type of intervention.[56] These interpretations are consistent with the success of a few visualization-based interventions, such as age-progressed renderings of oneself to encourage saving (Hershfield et al., 2011), visualization of an alternative future to encourage usage of preventive health products (John and Orkin, 2021), and training programs about visualizing future scenarios to promote entrepreneurial success (Ashraf et al., 2021).[57]

# 7 Concluding Remarks

Research on intertemporal choice has an extremely rich history in social science as intertemporal tradeoffs are ubiquitous in almost all important economic decisions. The discounting-based time preference model has been extensively studied and adopted as one of the benchmark elements in economics thanks to its tractability and parsimonious nature. Yet, the accumulated empirical results guided by the time preference benchmark seem to have pointed out a handful of unsolved puzzles (Cohen et al., 2020). This paper identifies an inherent wedge between the decision state and the benefiting state inherent in intertemporal settings and connects the state variation with projective misperceptions,[58] contributing to a recently growing effort that aims to better resolve existing puzzles by exploring alternative psychological mechanisms of intertemporal choice. This paper formalizes an identification problem in distinguishing between "pure" underlying time preferences and projective misperceptions in different states. More importantly, it combines a stylized experiment and surveys that are tightly linked to the theoretical framework to demonstrate the empirical relevance of projective misperceptions for predicting intertemporal choice both within and beyond the experiment.

The results on the role of experience and the mechanisms underlying projective misperceptions suggest that targeting misperceptions may have important policy implications for reducing suboptimal behavior in intertemporal contexts. I find that that these misperceptions are largely driven by incorrectly projecting informational states, rather than projecting current utility, and, therefore, experience with relevant states can largely help mitigate the influence of misperceptions on average. As discussed in Section 6, several potential solutions to deal with the identification challenge and new policy designs based on experience are inspired by these results, opening up avenues for future research.

While one may be tempted to interpret the impact of experience as the evidence indicating that the identi-

[55]See Caplin (2003) for a discussion of fear as an policy instrument.

[56]Borghans and Golsteyn (2005) find a positive correlation between measured discount rate with the ability to imagine the future among Dutch high school and college graduates.

[57]Relatedly, the science fiction disaster film that depicts catastrophic climatic events, *The Day After Tomorrow*, makes viewers more concerned about climate change and more motivated to act on climate change (Lowe et al., 2006).

[58]Projective misperceptions can be seen as deviation from the rational expectation in the language of macroeconomics.

fication issue is not important, there are some important caveats to consider. First, the complete correction of misperceptions may be partly due to the extremely transparent nature of the stylized experiment. This does not necessarily generalize in many daily situations, where the settings are more complex. In more complex settings, an individual is unlikely to encounter the exact state again and thus may never have perfect experience as in my experiment. Therefore, the mitigating impact of experience may be limited. Second, when making many important intertemporal decisions, while people may have ample experience in some settings, they do not always have opportunities to obtain relevant experience with the alternative state that could have happened in the counterfactual in all settings. In fact, to gain the perfect experience is impossible in many important settings that are extremely policy-relevant, such as contexts that involve decision making at a young age for outcomes to be realized at an old age. This paper thus provides a useful framework to evaluate contexts where the identification issue can be concerning based on state variation.

Lastly, although I provide suggestive evidence that projective misperceptions may matter beyond the experimental context, there are several important open questions. While people may have systematic state-dependent valuations due to various reasons such as worsening health conditions, aging, or habit formation, it remains unknown how large these state variations are quantitatively in each specific context. Relatedly, while the comparison of variation in states between monetary domains and non-monetary domains sounds intuitive, the comparison is highly speculative. This paper thus speaks to the need for further research that can help precisely compare settings where state variation is more or less likely to be prevalent. It would be useful to come up with new empirical methods that can help either measure or bound the magnitude of state variation across contexts, so that applied researchers can better evaluate the identification issue in their specific context.

## References


Abeler, J., Falk, A., Goette, L., and Huffman, D. (2011). Reference points and effort provision. *American Economic Review*, 101(2):470–92.

Acland, D. and Levy, M. R. (2015). Naiveté, projection bias, and habit formation in gym attendance. *Management Science*, 61(1):146–160.

Ahn, D. S., Iijima, R., Le Yaouanq, Y., and Sarver, T. (2019). Behavioural characterizations of naivete for time-inconsistent preferences. *The Review of Economic Studies*, 86(6):2319–2355.

Ahn, D. S., Iijima, R., and Sarver, T. (2020). Naivete about temptation and self-control: Foundations for recursive naive quasi-hyperbolic discounting. *Journal of Economic Theory*, 189:105087.

Ameriks, J., Caplin, A., Leahy, J., and Tyler, T. (2007). Measuring self-control problems. *American Economic Review*, 97(3):966–972.

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

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

Andreoni, J. and Sprenger, C. (2012). Estimating time preferences from convex budgets. *American Economic Review*, 102(7):3333–56.

Ashraf, N., Delfino, A., Holmes, E., Iacovone, L., and Pople, A. (2021). Learning to see the world's opportunities: The impact of visualization on entrepreneurial success.

Ashraf, N., Karlan, D., and Yin, W. (2006). Tying odysseus to the mast: Evidence from a commitment savings product in the philippines. *The Quarterly Journal of Economics*, 121(2):635–672.

Attema, A. E., Bleichrodt, H., Gao, Y., Huang, Z., and Wakker, P. P. (2016). Measuring discounting without measuring utility. *American Economic Review*, 106(6):1476–94.

Augenblick, N. (2018). Short-term time discounting of unpleasant tasks. *UC Berkeley Working Paper*.

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

Augenblick, N. and Rabin, M. (2018). An experiment on time preference and misprediction in unpleasant tasks. *Review of Economic Studies*, 86(3):941–975.

Badger, G. J., Bickel, W. K., Giordano, L. A., Jacobs, E. A., Loewenstein, G., and Marsch, L. (2007). Altered states: The impact of immediate craving on the valuation of current and future opioids. *Journal of Health Economics*, 26(5):865–876.

Bai, L., Handel, B., Miguel, E., and Rao, G. (2020). Self-control and demand for preventive health: Evidence from hypertension in india. *Review of Economics and Statistics*, pages 1–55.

Baicker, K., Mullainathan, S., and Schwartzstein, J. (2015). Behavioral hazard in health insurance. *The Quarterly Journal of Economics*, 130(4):1623–1667.

Becker, G. M., DeGroot, M. H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral science*, 9(3):226–232.

Bernheim, B. D. (2009). Behavioral welfare economics. *Journal of the European Economic Association*, 7(2-3):267–319.

Bordalo, P., Gennaioli, N., Ma, Y., and Shleifer, A. (2020a). Overreaction in macroeconomic expectations. *American Economic Review*, 110(9):2748–82.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*, 121(5):803–843.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2020b). Memory, attention, and choice. *The Quarterly Journal of Economics*, 135(3):1399–1442.

Borghans, L. and Golsteyn, B. (2005). Imagination, time discounting and human capital investment decisions. *Unpublished working paper, University of Maastricht, Maastricht.*

Breig, Z., Gibson, M., and Shrader, J. (2021). Why do we procrastinate? present bias and optimism. *Working Paper.*

Brown, J. R. and Previtero, A. (2020). Saving for retirement, annuities and procrastination. Technical report, Working Paper.

Bryan, G., Karlan, D., and Nelson, S. (2010). Commitment devices. *Annual Review of Economics*, 2(1):671–698.

Buchheim, L. and Kolaska, T. (2017). Weather and the psychology of purchasing outdoor movie tickets. *Management Science*, 63(11):3718–3738.

Buckner, R. L. and Carroll, D. C. (2007). Self-projection and the brain. *Trends in cognitive sciences*, 11(2):49–57.

Bushong, B. and Gagnon-Bartsch, T. (2020). An experiment on interpersonal projection bias. *Working Paper*.

Busse, M. R., Pope, D. G., Pope, J. C., and Silva-Risso, J. (2015). The psychological effect of weather on car purchases. *The Quarterly Journal of Economics*, 130(1):371–414.

Cadena, B. C. and Keys, B. J. (2015). Human capital and the lifetime costs of impatience. *American Economic Journal: Economic Policy*, 7(3):126–53.

Callen, M. (2015). Catastrophes and time preference: Evidence from the indian ocean earthquake. *Journal of Economic Behavior & Organization*, 118:199–214.

Caplin, A. (2003). Fear as a policy instrument. *Time and decision: Economic and psychological perspectives on intertemporal choice*, 441:452.

Carrera, M., Royer, H., Stehr, M., Sydnor, J., and Taubinsky, D. (2021). Who chooses commitment? evidence and welfare implications.

Carrillo, B. (2020). Present bias and underinvestment in education? long-run effects of childhood exposure to booms in colombia. *Journal of Labor Economics*, 38(4):1127–1265.

Casaburi, L. and Willis, J. (2018). Time versus state in insurance: Experimental evidence from contract farming in kenya. *American Economic Review*, 108(12):3778–3813.

Chabris, C. F., Laibson, D., Morris, C. L., Schuldt, J. P., and Taubinsky, D. (2008). Individual laboratory-measured discount rates predict field behavior. *Journal of risk and uncertainty*, 37(2):237–269.

Chakraborty, A. (2021). Present bias.

Chang, T. Y., Huang, W., and Wang, Y. (2018). Something in the air: Pollution and the demand for health insurance. *Review of Economic Studies*, 85(3):1609–1634.

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

Chapman, J., Dean, M., Ortoleva, P., Snowberg, E., and Camerer, C. (2019). Econographics.

Chen, X. and Zhao, X. (2020). How time flies! *Available at SSRN 3285834*.

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

Cubitt, R. P. and Read, D. (2007). Can intertemporal choice experiments elicit time preferences for consumption? *Experimental Economics*, 10(4):369–389.

DellaVigna, S. and Malmendier, U. (2006). Paying not to go to the gym. *American Economic Review*, 96(3):694–719.

DellaVigna, S. and Paserman, M. D. (2005). Job search and impatience. *Journal of Labor Economics*, 23(3):527–588.

Dertwinkel-Kalt, M., Gerhardt, H., Riener, G., Schwerter, F., and Strang, L. (2021). Concentration bias in intertemporal choice.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2017). The robustness and pervasiveness of sub-additivity in intertemporal choice. In *Working Paper*.

Duflo, E., Kremer, M., and Robinson, J. (2011). Nudging farmers to use fertilizer: Theory and experimental evidence from kenya. *American economic review*, 101(6):2350–90.

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

Echenique, F., Imai, T., and Saito, K. (2020). Testable implications of models of intertemporal choice: Exponential discounting and its generalizations. *American Economic Journal: Microeconomics*, 12(4):114–43.

Enke, B. and Graeber, T. (2021). Cognitive uncertainty and intertemporal choice.

Enke, B., Schwerter, F., and Zimmermann, F. (2020). Associative memory and belief formation. Technical report, National Bureau of Economic Research.

Epper, T. and Fehr-Duda, H. (2021). Risk in time: The intertwined nature of risk taking and time discounting. Technical report, Mimeo.

Ericson, K. M. (2017). On the interaction of memory and procrastination: Implications for reminders, deadlines, and empirical estimation. *Journal of the European Economic Association*, 15(3):692–719.

Ericson, K. M. and Laibson, D. (2019). Intertemporal choice. In *Handbook of Behavioral Economics: Applications and Foundations 1*, volume 2, pages 1–67. Elsevier.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., and Sunde, U. (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, 133(4):1645–1692.

Falk, A., Becker, A., Dohmen, T. J., Huffman, D., and Sunde, U. (2016). The preference survey module: A validated instrument for measuring risk, time, and social preferences. *Working Paper*.

Falk, A., Neuber, T., and Strack, P. (2021). Limited self-knowledge and survey response behavior.

Fedyk, A. (2021). Asymmetric naivete: Beliefs about self-control. *Available at SSRN 2727499*.

Frederick, S. (2003). Time preference and personal identity. *Time and decision*, pages 89–113.

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

Frey, B. S. and Stutzer, A. (2014). Economic consequences of mispredicting utility. *Journal of Happiness Studies*, 15(4):937–956.

Gabaix, X. and Laibson, D. (2017). Myopia and discounting. *NBER Working Paper*.

Gagnon-Bartsch, T., Pagnozzi, M., Rosato, A., et al. (2021). Projection of private values in auctions. *American economic review*.

Gershman, S. J. and Bhui, R. (2020). Rationally inattentive intertemporal choice. *Nature communications*, 11(1):1–8.

Gul, F. and Pesendorfer, W. (2001). Temptation and self-control. *Econometrica*, 69(6):1403–1435.

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

Harstad, B. (2020). Technology and time inconsistency. *Journal of Political Economy*, 128(7):2653–2689.

Havránek, T., Irsova, Z., and Matousek, J. (2021). Individual discount rates: A meta-analysis of experimental evidence.

Heidhues, P. and Strack, P. (2021). Identifying present-bias from the timing of choices. *American Economic Review*.

Hershfield, H. E., Goldstein, D. G., Sharpe, W. F., Fox, J., Yeykelis, L., Carstensen, L. L., and Bailenson, J. N. (2011). Increasing saving behavior through age-progressed renderings of the future self. *Journal of Marketing Research*, 48(SPL):S23–S37.

Huffman, D., Maurer, R., and Mitchell, O. S. (2019). Time discounting and economic decision-making in the older population. *The journal of the economics of ageing*, 14:100121.

Imai, T., Rutter, T. A., and Camerer, C. F. (2020). Meta-analysis of present-bias estimation using convex time budgets. *Economic Journal*, pages Art–No.

Imas, A., Kuhn, M., and Mironova, V. (2021). Waiting to choose: The role of deliberation in intertemporal choice. *American Economic Journal: Microeconomics*.

John, A. (2020). When commitment fails: evidence from a field experiment. *Management Science*, 66(2):503–529.

John, A. and Orkin, K. (2021). Can simple psychological interventions increase preventive health investment?

Kahana, M. J. (2012). *Foundations of human memory*. OUP USA.

Kahneman, D. and Thaler, R. H. (2006). Anomalies: Utility maximization and experienced utility. *Journal of economic perspectives*, 20(1):221–234.

Kahneman, D., Wakker, P. P., and Sarin, R. (1997). Back to bentham? explorations of experienced utility. *The quarterly journal of economics*, 112(2):375–406.

Karni, E., Schmeidler, D., and Vind, K. (1983). On state dependent preferences and subjective probabilities. *Econometrica*, pages 1021–1031.

Kaufmann, M. (2021). Projection bias in effort choices. *Working Paper arXiv:2104.04327*.

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

Kőszegi, B. and Szeidl, A. (2013). A model of focusing in economic choice. *The Quarterly journal of economics*, 128(1):53–104.

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

Laibson, D. (2015). Why don't present-biased agents make commitments? *American Economic Review*, 105(5):267–72.

Laibson, D. I., Repetto, A., Tobacman, J., Hall, R. E., Gale, W. G., and Akerlof, G. A. (1998). Self-control and saving for retirement. *Brookings papers on economic activity*, 1998(1):91–196.

Le Yaouanq, Y. and Schwardmann, P. (2020). Learning about one's self. *CEPR Discussion Paper No. DP13510*.

Levitt, S. D. (2021). Heads or tails: The impact of a coin toss on major life decisions and subsequent happiness. *The Review of Economic Studies*, 88(1):378–405.

Loewenstein, G. and Elster, J. (1992). *Choice over time.* Russell Sage Foundation.

Loewenstein, G., O'Donoghue, T., and Rabin, M. (2003). Projection bias in predicting future utility. *The Quarterly Journal of Economics*, 118(4):1209–1248.

Lowe, T., Brown, K., Dessai, S., de França Doria, M., Haynes, K., and Vincent, K. (2006). Does tomorrow ever come? disaster narrative and public perceptions of climate change. *Public understanding of science*, 15(4):435–457.

Madarász, K. (2012). Information projection: Model and applications. *The Review of Economic Studies*, 79(3):961–985.

Magnac, T. and Thesmar, D. (2002). Identifying dynamic discrete decision processes. *Econometrica*, 70(2):801–816.

Malmendier, U. (2021). Exposure, experience, and expertise: why personal histories matter in economics. *National Bureau of Economic Research.*

Manzini, P. and Mariotti, M. (2014). Welfare economics and bounded rationality: the case for model-based approaches. *Journal of Economic Methodology*, 21(4):343–360.

Martinez, S.-K., Meier, S., and Sprenger, C. (2017). Procrastination in the field: Evidence from tax filing.

McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G., and Cohen, J. D. (2007). Time discounting for primary rewards. *Journal of neuroscience*, 27(21):5796–5804.

McGuire, J. T. and Kable, J. W. (2013). Rational temporal predictions can underlie apparent failures to delay gratification. *Psychological review*, 120(2):395.

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

Meier, S. and Sprenger, C. D. (2012). Time discounting predicts creditworthiness. *Psychological science*, 23(1):56–58.

Meier, S. and Sprenger, C. D. (2015). Temporal stability of time preferences. *Review of Economics and Statistics*, 97(2):273–286.

Millner, A. (2020). Nondogmatic social discounting. *American Economic Review*, 110(3):760–75.

Millner, A. and Heal, G. (2021). Choosing the future: Markets, ethics, and rapprochement in social discounting.

Noor, J. and Takeoka, N. (2021). Optimal discounting. Technical report.

Odermatt, R. and Stutzer, A. (2019). (mis-) predicted subjective well-being following life events. *Journal of the European Economic Association*, 17(1):245–283.

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

O'Donoghue, T. and Vogelsang, T. J. (2007). Projection bias in catalog orders. *American Economic Review*, 97(4):1217–1249.

Ramsey, F. P. (1928). A mathematical theory of saving. *The economic journal*, 38(152):543–559.

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

Read, D. and Van Leeuwen, B. (1998). Predicting hunger: The effects of appetite and delay on choice. *Organizational Behavior and Human Decision Processes*, 76(2):189–205.

Rick, S. and Loewenstein, G. (2008). Intangibility in intertemporal choice. *Philosophical transactions of the royal Society B: Biological Sciences*, 363(1511):3813–3824.

Rust, J. (1994). Structural estimation of markov decision processes. *Handbook of econometrics*, 4:3081–3143.

Sadoff, S., Samek, A., and Sprenger, C. (2020). Dynamic inconsistency in food choice: Experimental evidence from a food desert. *The Review of Economic Studies*, 87:1954–1988.

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

Simonsohn, U. (2009). Weather to go to college. *The Economic Journal*, 120(543):270–280.

Simonsohn, U., Karlsson, N., Loewenstein, G., and Ariely, D. (2008). The tree of experience in the forest of information: Overweighing experienced relative to observed information. *Games and Economic Behavior*, 62(1):263–286.

Slade, A. N. (2012). Health investment decisions in response to diabetes information in older americans. *Journal of Health Economics*, 31(3):502–520.

Spreng, R. N., McKinnon, M. C., Mar, R. A., and Levine, B. (2009). The toronto empathy questionnaire: Scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of personality assessment*, 91(1):62–71.

Strack, P. and Taubinsky, D. (2021). Dynamic preference "reversals" and time inconsistency. *Working Paper*.

Sutter, M., Kocher, M. G., Glätzle-Rützler, D., and Trautmann, S. T. (2013). Impatience and uncertainty: Experimental decisions predict adolescents' field behavior. *American Economic Review*, 103(1):510–31.

Toussaert, S. (2018). Eliciting temptation and self-control through menu choices: A lab experiment. *Econometrica*, 86(3):859–889.

Ubfal, D. (2016). How general are time preferences? eliciting good-specific discount rates. *Journal of Development Economics*, 118:150–170.

Wilson, T. D. and Gilbert, D. (2003). Affective forecasting. *Advances in Experimental Social Psychology*, 35:345–411.

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

# Appendix A Additional Theoretical Analyses

## Appendix A.1 Modeling Differential Responsiveness to Experience

Here I illustrate a version of the model that explicitly models differential responsiveness to experience by separating between state projection and utility projection. In addition to all the assumptions made in the main text, the instantaneous utility is assumed to be linear in the state $s$, with $u(a, ms+ns') = mu(a,s)+nu(a,s')$ for any $m,n \in \mathbb{R}$. The decision maker may exhibit two types of projective misperceptions when predicting her future preferences $\tilde{u}(a,s'|s)$, as defined in Definition 2.

**Definition 2** (**Projective Misperceptions**). *A decision maker has projective misperceptions if there exists* $\gamma(e) \in [0,1]$ *and* $\alpha \in [-1,1]$ *such that for all activity* $a \in \mathbb{R}$*, current and future states* $s, s' \in \mathcal{S}$*, her predicted future utility is*

$$\tilde{u}(a,s'|s) = \underbrace{\alpha u(a,s) + (1-\alpha)u[a, \overbrace{\gamma(e)s + (1-\gamma(e))s'}^{\text{state projection}}]}_{\text{utility projection}}$$

The first type of misperceptions, **projection of states**, reflects the intuition of information projection (Madarász, 2012; Wilson and Gilbert, 2003), with which the decision maker incorrectly anticipate changes in states. As captured in $\gamma(e)$, where $e \in \mathbb{R}_+$ denotes experience with the relevant future state, projection of current states mainly occurs when the decision maker does not have much information or experience of relevant future states. This type of misperceptions can therefore be mitigated by having more experience such that $\gamma(0) > 0$, $\gamma'(e) < 0$ and $\lim_{e\to\infty} \gamma(e) = 0$. If $\gamma(e) = 0$, the decision maker has no misperception when predicting her future state, and her prediction equals her true future state, $s'|s = s'$; if $\gamma(e) = 1$, the decision maker considers her future state to be exactly the same as her current one, $s'|s = s$.

The second type of misperceptions, **projection of utilities**, reflects the simple projection bias $\alpha$ as in Loewenstein et al. (2003), with which the decision maker correctly predict changes in states but incorrectly perceive how changes in states can lead to changes in utility. In the original framework of Loewenstein et al. (2003), if $\alpha = 0$, the decision maker has no misperception when predicting her future utility and her prediction equals her true future utility, $\tilde{u}(a,s'|s) = u(a,s')$; if $\alpha = 1$, the decision maker considers her future utility to be exactly the same as her current utility in the current state, $\tilde{u}(a,s'|s) = u(a,s)$.

Note that I also allow $\alpha < 0$ to reflect that people may also overestimate changes in tastes, a heterogeneity that has not been discussed in prior research. While Loewenstein et al. (2003) characterize projection bias as underestimating changes in tastes, it is not clear according to their framework whether those who overestimate changes in tastes exhibit no bias or the opposite misperceptions. In the latter case, state projection needs to be sufficiently large, $\gamma(e) > -\frac{\alpha}{1-\alpha}$, for the misperceptions to be projective.

## Appendix A.2 Distorting Impacts in Both States

**Corollary 4.** *The distortions on the estimate, $\hat{\delta}$, of ignoring misperceptions in the good state and the bad state cancel out on average if and only if* $\frac{1+\pi[\nu(a_b^*)-1]}{1+\pi[\nu(a_g^*)-1]}\frac{\nu(a_g^*)-1}{\nu(a_b^*)-1} = \frac{1-\pi}{\pi}$, *where* $\nu(a) = u_a'(a,b)/u_a'(a,g)$.

Above result suggests that the distorting impact of misperceptions on the estimated time preferences is dependent on both the magnitude of state variation and the probability of state realization in the economic environment, which may potentially provide useful information for the analyst to pick the right specification about noises. It shows that the impacts from both states may cancel out only in a very special case. Nevertheless, even though the distortions may not cancel out fully, eliciting choices in different states still helps reduce the estimation bias to some extent in practice. This may be the case when there is little information about the economic environment thus the exact probability of state realization and the magnitude of state variation are unknown.

# Appendix B Proofs

## Proof of Lemma 1

*Proof.* The decision maker solves the following optimization problem:

$$\max_{a\in\mathbb{R}_+} \delta[(1-\pi)\tilde{u}(a,g|s)+\pi\tilde{u}(a,b|s)]-C(a), \text{ where } s\in\mathcal{S}\equiv\{b,g\}$$

The decision maker who is currently in a good state $s=g$ predicts her future utility to be $\tilde{u}(a,g|g)=u(a,g)$ and $\tilde{u}(a,b|g)=\alpha(e)u(a,g)+[1-\alpha(e)]u(a,b)$. This thus gives the following first order condition that characterizes $a_g^*$:

$$C'(a_g^*)=\delta\left([1-\alpha(e)+\frac{\alpha(e)}{1-\pi}](1-\pi)u_a'(a_g^*,g)+[1-\alpha(e)]\pi u_a'(a_g^*,b)\right).$$

Similarly, the decision maker who is currently in a bad state $s=b$ predicts her future utility to be $\tilde{u}(a,b|b)=u(a,b)$ and $\tilde{u}(a,g|b)=\alpha(e)u(a,b)+[1-\alpha(e)]u(a,g)$. This gives the following first order condition that characterizes $a_b^*$:

$$C'(a_b^*)=\delta\left([1-\alpha(e)](1-\pi)u_a'(a_b^*,g)+[1-\alpha(e)+\frac{\alpha(e)}{\pi}]\pi u_a'(a_b^*,b)\right).$$

□

## Proof of Proposition 1

*Proof.* As the cost function is convex, $C''(a)>0$, the marginal cost $C'(a)$ is *strictly increasing* in $a\in\mathcal{A}$.

As the instantaneous utility function is concave in $a$, $u_a''(a,s)<0$, for all $s\in\mathcal{S}$, the instantaneous marginal utility function $u_a'(a,s)$ is *strictly decreasing* in $a\in\mathbb{R}_+$ for all $s\in\mathcal{S}$. Therefore, any linear combination of $u_a'(a,g)$ and $u_a'(a,b)$ is as well *strictly decreasing* in $a\in\mathbb{R}_+$, such as:

- perceived marginal utility in good state $MR_g(a)\equiv\delta\left([1-\alpha(e)+\frac{\alpha(e)}{1-\pi}](1-\pi)u_a'(a,g)+[1-\alpha(e)]\pi u_a'(a,b)\right)$
- perceived marginal utility in bad state $MR_b(a)\equiv\delta\left([1-\alpha(e)](1-\pi)u_a'(a,g)+[1-\alpha(e)+\frac{\alpha(e)}{\pi}]\pi u_a'(a,b)\right)$
- the actual marginal utility $MR(a)=(1-\pi)u_a'(a,g)+\pi u_a'(a,b)$

Since I focus on the non-trivial case where the activity is beneficial, $C'(0)<\delta[1-\alpha(e)][(1-\pi)u_a'(0,g)+\pi u_a'(0,b)]$, there exists unique $a^*$, $a_g^*$, $a_b^*$ as characterized in Lemma 1.

**I prove $a_g^*\leq a^*$ by contradiction**. Suppose $a_g^*>a^*$, where the benchmark $a^*$ is characterized by

$C'(a^*) = MR(a^*) \equiv (1-\pi)u'_a(a^*, g) + \pi u'_a(a^*, b)$. As $MR(a)$ is decreasing and $C'(a)$ is increasing in $a$, the following inequalities hold, $MR(a^*_g) < MR(a^*) = C'(a^*) < C'(a^*_g)$.

Note that $C'(a^*_g) = MR_g(a^*_g) \equiv \delta\left([1-\alpha(e)+\frac{\alpha(e)}{1-\pi}](1-\pi)u'_a(a^*_g, g) + [1-\alpha(e)]\pi u'_a(a^*_g, b)\right)$ according to Lemma 1. The RHS can be rewritten as $\delta\left((1-\pi)u'_a(a^*_g, g) + \pi u'_a(a^*_g, b) + \alpha(e)\pi[u'_a(a^*_g, g) - u'_a(a^*_g, b)]\right)$. Recall the state-dependent return condition, $u'_a(a, b) > u'_a(a, g)$ for all $a$, therefore, $u'_a(a^*_g, g) - u'_a(a^*_g, b) < 0$. This means that $C'(a^*_g) < (1-\pi)u'_a(a^*_g, g) + \pi u'_a(a^*_g, b) = MR(a^*_g)$, contradicting the inequalities above.

**I prove $a^*_b \geq a^*$ when $\delta \geq \bar{\delta} = \frac{(1-\pi)u'_a(a^*_b, g) + \pi u'_a(a^*_b, b)}{\left((1-\pi)u'_a(a^*_b, g) + \pi u'_a(a^*_b, b) + \alpha(e)(1-\pi)[u'_a(a^*_b, b) - u'_a(a^*_b, g)]\right)}$ by contradiction.** Suppose $a^*_b < a^*$, the monotonicity of $u'_a(a, s)$ and $C'(a)$ leads to $MR(a^*_b) > MR(a^*) = C'(a^*) > C'(a^*_b)$.

Note that $C'(a^*_b) = MR_b(a^*_b) \equiv \delta\left([1-\alpha(e)](1-\pi)u'_a(a^*_b, g) + [1-\alpha(e)+\frac{\alpha(e)}{\pi}]\pi u'_a(a^*_b, b)\right)$ according to Lemma 1. The RHS can be rewritten as $\delta\left((1-\pi)u'_a(a^*_b, g) + \pi u'_a(a^*_b, b) + \alpha(e)(1-\pi)[u'_a(a^*_b, b) - u'_a(a^*_b, g)]\right)$. Recall the state-dependent return condition, $u'_a(a, b) > u'_a(a, g)$ for all $a$, therefore, $u'_a(a^*_b, b) - u'_a(a^*_b, g) > 0$. This means that whenever $\delta \geq \bar{\delta}$, $C'(a^*_b) = MR_b(a^*_b) > MR(a^*_b)$, which contradicts the inequalities above. □

## Proof of Proposition 2

*Proof.* Consider an activity $a \in \mathcal{A}$ generated by projective misperceptions alone: $\alpha(e) \in (0, 1]$ and $\delta = 1$.

**If the activity $a$ is determined in a good state**, Lemma 1 characterizes $a^*_g$ as follows:

$$\begin{aligned}
C'(a^*_g) &= [1-\alpha(e)+\frac{\alpha(e)}{1-\pi}](1-\pi)u'_a(a^*_g, g) + [1-\alpha(e)]\pi u'_a(a^*_g, b) \\
&= (1-\pi)u'_a(a^*_g, g) + \pi u'_a(a^*_g, b) + \alpha(e)\pi[u'_a(a^*_g, g) - u'_a(a^*_g, b)] \\
&< (1-\pi)u'_a(a^*_g, g) + \pi u'_a(a^*_g, b), \text{ because } u'_a(a^*_g, g) < u'_a(a^*_g, b)
\end{aligned}$$

Therefore, one can rewrite the above by construction as:

$$C'(a^*_g) = \delta\left((1-\pi)u'_a(a^*_g, g) + \pi u'_a(a^*_g, b)\right), \text{ where } \delta \in (0, 1)$$

This means that any $a^*_g$ can also be generated by $\alpha(e) = 0$ and $\delta \in (0, 1)$.

**If the activity $a$ is determined in a bad state**, Lemma 1 characterizes $a^*_b$ as follows:

$$\begin{aligned}
C'(a^*_b) &= [1-\alpha(e)](1-\pi)u'_a(a^*_b, g) + [1-\alpha(e)+\frac{\alpha(e)}{\pi}]\pi u'_a(a^*_b, b) \\
&= (1-\pi)u'_a(a^*_b, g) + \pi u'_a(a^*_b, b) + \alpha(e)(1-\pi)[u'_a(a^*_b, b) - u'_a(a^*_b, g)] \\
&> (1-\pi)u'_a(a^*_b, g) + \pi u'_a(a^*_b, b), \text{ because } u'_a(a^*_b, b) > u'_a(a^*_b, g)
\end{aligned}$$

Therefore, one can rewrite the above by construction as:

$$C'(a_b^*) = \delta\left((1-\pi)u_a'(a_b^*, g) + \pi u_a'(a_b^*, b)\right), \text{ where } \delta > 1$$

This means that any $a_b^*$ can also be generated by $\alpha(e) = 0$ and $\delta > 1$. $\square$

## Proof of Proposition 3

*Proof.* Recall that $\nu(a) = u_a'(a, b)/u_a'(a, g) > 1$ denotes the magnitude of state variation at the activity $a$.

Suppose the analyst observes the activity $a_s^*$ made in a state $s$. When the analyst ignores the decision maker's misperceptions by mistakenly assuming $\alpha(e) = 0$, he will estimate time preference as follows:

$$\hat{\delta} = \frac{C'(a_s^*)}{(1-\pi)u_a'(a_s^*, g) + \pi u_a'(a_s^*, b)}.$$

**If the activity is made in a good state**, Lemma 1 characterizes the "true" underlying time preference:

$$\delta_g = \frac{C'(a_g^*)}{(1-\pi)u_a'(a_g^*, g) + \pi u_a'(a_g^*, b) + \alpha(e)\pi[u_a'(a_g^*, g) - u_a'(a_g^*, b)]}.$$

Therefore, the analyst's estimated time preference from the activity made in a good state is:

$$\begin{aligned}
\Rightarrow \hat{\delta} &= \left(1 + \alpha(e)\frac{\pi[u_a'(a_g^*, g) - u_a'(a_g^*, b)]}{(1-\pi)u_a'(a_g^*, g) + \pi u_a'(a_g^*, b)}\right)\delta_g \\
&= \left(1 - \alpha(e)\frac{1}{\frac{u_a'(a_g^*, g)}{\pi[u_a'(a_g^*, b) - u_a'(a_g^*, g)]} + 1}\right)\delta_g \\
&= \left(1 - \alpha(e)\frac{1}{\frac{1}{\pi[\nu(a_g^*) - 1]} + 1}\right)\delta_g \\
&= [1 - \eta_g\alpha(e)]\delta_g, \text{ where } \eta_g = \frac{1}{\frac{1}{\pi[\nu(a_g^*) - 1]} + 1}.
\end{aligned}$$

**If the activity is made in a bad state**, Lemma 1 characterizes the "true" underlying time preference:

$$\delta_b = \frac{C'(a_b^*)}{(1-\pi)u_a'(a_b^*, g) + \pi u_a'(a_b^*, b) + \alpha(e)(1-\pi)[u_a'(a_b^*, b) - u_a'(a_b^*, g)]}.$$

Therefore, the analyst's estimated time preference from the activity made in a bad state is:

$$
\begin{aligned}
\Rightarrow \hat{\delta} &= \left(1+\alpha(e)\frac{(1-\pi)[u_a'(a_b^*,b)-u_a'(a_b^*,g)]}{(1-\pi)u_a'(a_b^*,g)+\pi u_a'(a_b^*,b)}\right)\delta_b \\
&= \left(1+\alpha(e)\frac{1}{\frac{u_a'(a_b^*,b)}{(1-\pi)[u_a'(a_b^*,b)-u_a'(a_b^*,g)]}-1}\right)\delta_b \\
&= \left(1+\alpha(e)\frac{1}{\frac{1}{(1-\pi)[1-\frac{1}{\nu(a_b^*)}]}-1}\right)\delta_b \\
&= [1+\eta_b\alpha(e)]\delta_b, \text{ where } \eta_b = \frac{1}{\frac{1}{(1-\pi)[1-\frac{1}{\nu(a_b^*)}]}-1}.
\end{aligned}
$$

□

## Proof of Corollary 1

*Proof.* Recall that the activity in a good state $a_g^*$ is characterized in Lemma 1 by the following:

$$C'(a_g^*) = \delta\left([1-\alpha(e)+\frac{\alpha(e)}{1-\pi}](1-\pi)u_a'(a_g^*,g)+[1-\alpha(e)]\pi u_a'(a_g^*,b)\right)$$

I prove $\frac{\partial a_g^*}{\partial\alpha(e)}<0$ by contradiction. Take arbitrary $\alpha_1(e),\alpha_2(e)\in[0,1]$ such that $\alpha_1(e)<\alpha_2(e)$.

Suppose $a_{g1}^*\leq a_{g2}^*$, then the first inequality holds because $u_a'(a,s)$ is *strictly decreasing* in $a\in\mathbb{R}_+$ for all $s\in\mathcal{S}$, and the second inequality holds because $u_a'(a_{g2}^*,g)-u_a'(a_{g2}^*,b)<0$:

$$
\begin{aligned}
C'(a_{g1}^*) &= \delta\left([1-\alpha_1(e)+\frac{\alpha_1(e)}{1-\pi}](1-\pi)u_a'(a_{g1}^*,g)+[1-\alpha_1(e)]\pi u_a'(a_{g1}^*,b)\right) \\
&> \delta\left([1-\alpha_1(e)+\frac{\alpha_1(e)}{1-\pi}](1-\pi)u_a'(a_{g2}^*,g)+[1-\alpha_1(e)]\pi u_a'(a_{g2}^*,b)\right) \\
&= \delta\left((1-\pi)u_a'(a_{g2}^*,g)+\pi u_a'(a_{g2}^*,b)+\alpha_1(e)\pi[u_a'(a_{g2}^*,g)-u_a'(a_{g2}^*,b)]\right) \\
&> \delta\left((1-\pi)u_a'(a_{g2}^*,g)+\pi u_a'(a_{g2}^*,b)+\alpha_2(e)\pi[u_a'(a_{g2}^*,g)-u_a'(a_{g2}^*,b)]\right) \\
&= C'(a_{g2}^*)
\end{aligned}
$$

However, $C'(a_{g1}^*)>C'(a_{g2}^*)$ contradicts to the fact that $C'(a)$ is *strictly increasing* in $a\in\mathbb{R}_+$. Therefore, it must be that $\alpha_1(e)<\alpha_2(e)$ and $a_{g1}^*>a_{g2}^*$, i.e. $\frac{\partial a_g^*}{\partial\alpha(e)}<0$.

$\frac{\partial a_b^*}{\partial\alpha(e)}>0$ can be proved using similar reasoning. □

## Proof of Corollary 2

*Proof.* Corollary 2 directly follows Corollary 1 because $\alpha'(e)\leq 0$. □

## Proof of Corollary 3

*Proof.* The distorting impact on estimated time preference $\hat{\delta}$ in a good state can be characterized as follows:

$$\eta_g = \frac{1}{\frac{1}{\pi[\nu(a_g^*)-1]}+1} \in (0,1) \text{ is increasing in } \nu(a_g^*).$$

Similarly, the distorting impact in a bad state can be characterized as follows:

$$\eta_b = \frac{1}{\frac{1}{(1-\pi)[1-\frac{1}{\nu(a_b^*)}]}-1} \in (0, \frac{1-\pi}{\pi}) \text{ is increasing in } \nu(a_b^*)$$

□

## Proof of Corollary 4

*Proof.* Recall that $\nu(a) = u_a'(a,b)/u_a'(a,g) > 1$. The following condition, $\eta_g = \eta_b$, must hold so that the distorting impact in a good state and the impact in a bad state can cancel out on average:

$$\frac{1}{\frac{1}{\pi[\nu(a_g^*)-1]}+1} = \frac{1}{\frac{1}{(1-\pi)[1-\frac{1}{\nu(a_b^*)}]}-1}$$

$$\iff \frac{\pi+\frac{1}{\nu(a_b^*)}-\frac{\pi}{\nu(a_b^*)}}{(1-\pi)[1-\frac{1}{\nu(a_b^*)}]} = \frac{1+\pi\nu(a_g^*)-\pi}{\pi[\nu(a_g^*)-1]}$$

$$\iff \frac{1+\pi\nu(a_b^*)-\pi}{\nu(a_b^*)}\pi[\nu(a_g^*)-1] = [1+\pi\nu(a_g^*)-\pi](1-\pi)\frac{\nu(a_b^*)-1}{\nu(a_b^*)}$$

$$\iff \frac{1+\pi[\nu(a_b^*)-1]}{1+\pi[\nu(a_g^*)-1]}\frac{\nu(a_g^*)-1}{\nu(a_b^*)-1} = \frac{1-\pi}{\pi}$$

□

# Appendix C Experimental Instructions

**Introduction**

This is a longitudinal study conducted by researchers from the University of Zurich.

The entire study consists of two parts that will take place today and on Wednesday. In both parts of the study, you will do 10 counting tasks and we will ask you to make a few choices about these tasks. For each counting task, you need to count the number of ”0”s in a matrix. Each of these matrices comes one after another, once the previous task is correctly completed. You can find a sample matrix below and every matrix will be similar to this one except for the exact number of “0”s.

| | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 |
| 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 |
| 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 |
| 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 |
| 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 |
| 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 |

Session 1 today will take around 45-55 minutes in total. You will receive at least a fixed payment of 4 Pounds for completing this study. In addition, you may earn additional bonus payments depending on random chance and your choices during this study. Upon completion of this session, you will be given a code that you can submit to Prolific so that you can receive your payment shortly after your completion.

Session 2 will take place on Wednesday and will be similar to the session today. You will receive a participation bonus of 10 Pounds if you complete both parts of the study, on top of the payment you are to receive from each part. The invitation to Session 2 will be sent to you at 10:30 am (GMT+1) / 11:30 am (CET) on Wednesday, Mar.31, and you can participate in it until 20:00 pm (GMT+1) / 21:00 pm (CET).

**Overview of Session 1 and Session 2**

Before starting the task, here is a preview of all the choices you are going to make.

On top of completing ten mandatory tasks, you may earn a bonus payment today by performing an additional counting task. The computer will determine whether this additional task takes place by chance. If this additional task takes place, it will be similar to all other tasks and will take place after all the mandatory tasks and a 2-minute break, either at the very end of the session today, or at the very end of Session 2 on Wednesday. You can not affect whether or when this additional task takes place because it will be determined by the computer.

At some point later, we will ask you how much we should compensate you for you to accept an additional counting task. You need to choose whether you prefer the option on the left or the option on the right in each row on the following screens, assuming that the additional task takes place either today or in Session 2 on Wednesday. Here are the screenshots of these choices:

After you have made all your choices, if the computer determines that the additional task takes place, one of your choices will be randomly selected by the computer to count. Whatever you have chosen in that row will be implemented. That is:

- If the computer determines the additional task takes place today, and in the selected row you chose the option on the left, you do not need to perform an additional task at the end of the study today. You may complete the study directly and receive your fixed payment today.
- If the computer determines the additional task takes place today, and in the selected row you chose the option on the right, you need to perform an additional task at the end of the study today before completing the study. You will receive the bonus payment listed in the option on the right on top of your fixed payment today.
- If the computer determines the additional task takes place on Wednesday, no additional task is needed today; and if in the selected row you chose the option on the left, you do not need to perform an additional task at the end of Session 2 on Wednesday. You may complete the study directly and receive your fixed payment today.
- If the computer determines the additional task takes place on Wednesday, no additional task is needed today; and if in the selected row you chose the option on the right, you need to perform an additional task at the end of Session 2 on Wednesday before completing the study. You will receive the bonus payment listed in the option on the right on top of your fixed payment today.

Screenshot 1: Choices for the additional task today

| | |
|---|---|
| If the additional task takes place at the end of **today**: I prefer not to do it | I prefer this additional task without any bonus today |
| If the additional task takes place at the end of **today**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.05 today |
| If the additional task takes place at the end of **today**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.10 today |
| If the additional task takes place at the end of **today**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.15 today |
| If the additional task takes place at the end of **today**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.20 today |
| If the additional task takes place at the end of **today**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.25 today |
| If the additional task takes place at the end of **today**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.30 today |
| If the additional task takes place at the end of **today**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.35 today |
| If the additional task takes place at the end of **today**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.40 today |
| If the additional task takes place at the end of **today**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.45 today |
| If the additional task takes place at the end of **today**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.50 today |
| If the additional task takes place at the end of **today**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.60 today |
| If the additional task takes place at the end of **today**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.70 today |
| If the additional task takes place at the end of **today**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.80 today |
| If the additional task takes place at the end of **today**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.90 today |
| If the additional task takes place at the end of **today**: I prefer not to do it | I prefer this additional task for a bonus of GBP 1.00 today |

Screenshot 2: Choices for the additional task on Wednesday

| | |
|---|---|
| If the additional task takes place at the end of **Part 2**: I prefer not to do it | I prefer this additional task without any bonus today |
| If the additional task takes place at the end of **Part 2**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.05 today |
| If the additional task takes place at the end of **Part 2**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.10 today |
| If the additional task takes place at the end of **Part 2**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.15 today |
| If the additional task takes place at the end of **Part 2**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.20 today |
| If the additional task takes place at the end of **Part 2**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.25 today |
| If the additional task takes place at the end of **Part 2**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.30 today |
| If the additional task takes place at the end of **Part 2**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.35 today |
| If the additional task takes place at the end of **Part 2**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.40 today |
| If the additional task takes place at the end of **Part 2**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.45 today |
| If the additional task takes place at the end of **Part 2**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.50 today |
| If the additional task takes place at the end of **Part 2**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.60 today |
| If the additional task takes place at the end of **Part 2**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.70 today |
| If the additional task takes place at the end of **Part 2**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.80 today |
| If the additional task takes place at the end of **Part 2**: I prefer not to do it | I prefer this additional task for a bonus of GBP 0.90 today |
| If the additional task takes place at the end of **Part 2**: I prefer not to do it | I prefer this additional task for a bonus of GBP 1.00 today |

(The amount in the right option to which you first switch your choices from the left indicates your required minimum compensation. You should therefore only switch your choices once at most in both sets of questions.)

As you do not know which row the computer will select before your choices and each row could equally be the choice-that-counts, you should consider each of your choices carefully and treat it as if it will count.

Note that your choices will NOT influence the fixed payment you are to receive today nor your participation payment, regardless of whether you choose to perform an additional task or not.

**Quiz questions**

You must answer this question correctly in order to proceed. Is this statement true or false: "You may submit the study after completing all mandatory counting tasks and a 2-minute break if you chose not to perform an additional counting task"?

You must answer this question correctly in order to proceed. Is this statement true or false: "If the computer determines the additional task takes place on Wednesday, you may receive the associated bonus payment today and do the additional task on Wednesday"?

You must answer this question correctly in order to proceed. Is this statement true or false: "You will receive a bigger participation payment if you choose to perform an additional counting task"?

**Choices during mandatory counting tasks**

On a scale of 7, please evaluate how tired or exhausted you are at the moment when you have to perform the next counting task?

[manipulation check]

If you had to perform an additional task at the very end of the study today, please predict how tired or exhausted you will be?

[predicted tiredness for the additional task later today]

Assuming that the additional task takes place at the end of the study today, please tell us how much we should compensate you today for you to accept this additional task by choosing whether you prefer the option on the left or the option on the right in each row.

[wage choices for the additional task later today]

Recall that Session 2 will be similar to the session today and you will need to perform the same amount of mandatory tasks.

If you had to perform an additional task at the very end of Session 2 on Wednesday, please predict how tired or exhausted you will be?

[predicted tiredness for the additional task in two days]

Assuming that the additional task takes place at the end of Session 2 on Wednesday instead of today, please tell us how much we should compensate you today for you to accept this additional task by choosing whether you prefer the option on the left or the option on the right in each row.

[wage choices for the additional task in two days]

**2-minute rest & realization of additional task**

You have completed all mandatory counting tasks. Please take a 2-minute break and then continue. You can spend your break time however you want.

[realization of additional task]

**Final word**

You have completed Session 1 of the study today. Thank you for your participation!

After you have copied the Prolific completion code, please click the button below to submit your response. You will receive your payment together with the bonus payment by the end of today.

Session 2 of the study will take place in two days and you will receive the invitation at 10:30 am (GMT+1) / 11:30 am (CET) on Wednesday (Mar. 31). You will receive the participation payment of 10 Pounds after completing Session 2.

# Appendix D Overview of Participants in the Experiment

Table D.1: Overview of Experimental Participants

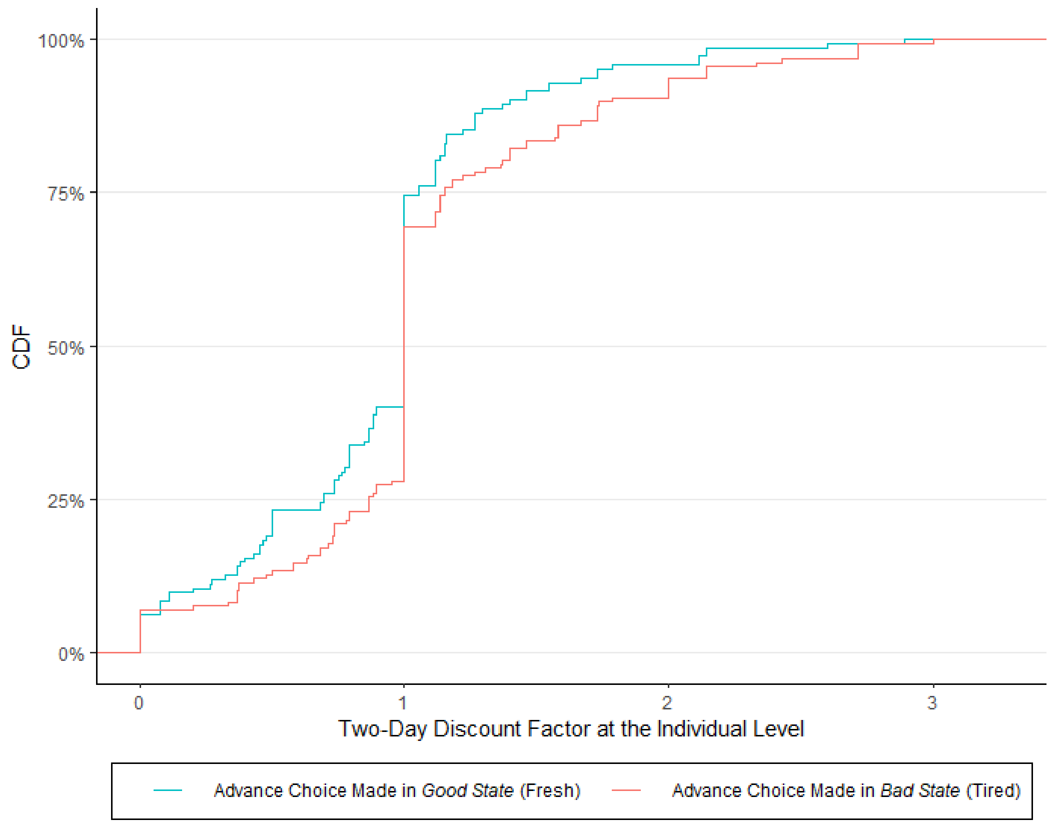
| | Treatment State | | Difference | $p$-value |
|---|---|---|---|---|
| | Good State | Bad State | | |
| *Session 1* | | | | |
| Age | 26.40 | 26.91 | -0.51 | 0.55 |
| | (0.58) | (0.63) | (0.86) | |
| Female | 0.38 | 0.45 | -0.08 | 0.12 |
| | (0.03) | (0.03) | (0.05) | |
| Student | 0.61 | 0.60 | 0.00 | 0.94 |
| | (0.03) | (0.03) | (0.05) | |
| Fully Employed | 0.30 | 0.31 | -0.01 | 0.83 |
| | (0.03) | (0.03) | (0.05) | |
| *Session 2* | | | | |
| Age | 26.63 | 26.60 | 0.03 | 0.9 |
| | (0.61) | (0.64) | (0.89) | |
| Female | 0.41 | 0.41 | 0.00 | 0.94 |
| | (0.03) | (0.04) | (0.05) | |
| Student | 0.59 | 0.62 | -0.03 | 0.59 |
| | (0.04) | (0.03) | (0.5) | |
| Fully Employed | 0.31 | 0.29 | 0.02 | 0.68 |
| | (0.03) | (0.03) | (0.05) | |

*Notes*: Standard errors are in parentheses. $p$-value is based on two-sided t-test.

# Appendix E Additional Empirical Analyses

## Appendix E.1 Two-Day Discount Factor at the Individual Level

Figure E.1: Two-Day Discount Factor at the Individual Level




*Notes*: The figure shows the cumulative distribution of the two-day discount factor at the individual level inferred from $\text{WTA}_{\text{future}}/\text{WTA}_{\text{today2}}$, where $\text{WTA}_{\text{today2}}$ is made in Session 2. The inferred discount factor is quite noisy as many participants in Session 2 report their reservation wages $\text{WTA}_{\text{today2}}$ to be zero or sufficiently small, such that the revealed two-day discount factor is too large relative to a reasonable range. The figure thus restricts attention to the data below the 90th percentile.

## Appendix E.2 Horizon-Dependent Misperceptions

There are two different interpretations of the horizon effects appeared in wages in the bad state treatment in Session 1, as captured in Figure 3. The first interpretation is that participants exhibit state-dependent discounting in the sense that they discount their future disutilities but only in a good state. The second interpretation is that the time horizon in my experimental setting is too short for time discounting to operate, and that the horizon effects reflect another type of mispredictions that may be state-*in*dependent and has not been documented before.

Specifically, the second interpretation reflects the conjecture that people may underappreciate the extent to which the current state can equally affect preferences in the remote future the same way as it affects preferences in the near future, which I refer as the lack of *intertemporal* empathy.[59] This is in contrast with projective misperceptions studied in this paper, which reflect a lack of *atemporal* empathy and appear when predicting preferences in a different context.[60]

Table E.1: Predicted Tiredness in Session 1

| | Dep Var: Predicted Tiredness, 1 to 7 | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Bad State | 0.893*** | | 0.919*** |
| | (0.142) | | (0.151) |
| In Two Days | | -0.495*** | -0.469*** |
| | | (0.0556) | (0.0830) |
| Bad State × In Two Days | | | -0.0513 |
| | | | (0.111) |
| Constant | 3.930*** | 4.628*** | 4.164*** |
| | (0.107) | (0.0785) | (0.116) |
| Observations | 860 | 860 | 860 |
| R-squared | 0.072 | 0.022 | 0.094 |

*Notes*: Standard errors clustered at the individual level. "Bad State" is a binary indicator for the bad state treatment. "In Two Days" is a dummy variable indicating that the wage is for additional task that tasks place in two days instead of later today. *** p<0.01, ** p<0.05, * p<0.1

If the second interpretation is the case, participants may also predict themselves to be less tired on top of being willing to accept lower compensations when additional work takes place in Session 2 as opposed to in Session 1. Whereas time discounting interpretation would predict that people are willing to accept lower compensations when additional work takes place in Part 2 and that the predicted tiredness is similar regardless

[59] The conjecture is partly motivated by the coexistence of overreaction and underreaction in individual and consensus forecasts (Bordalo et al., 2020a). The lack of *intertemporal* empathy may attenuate projective misperceptions when making predictions about remote future (as opposed to the immediate future).

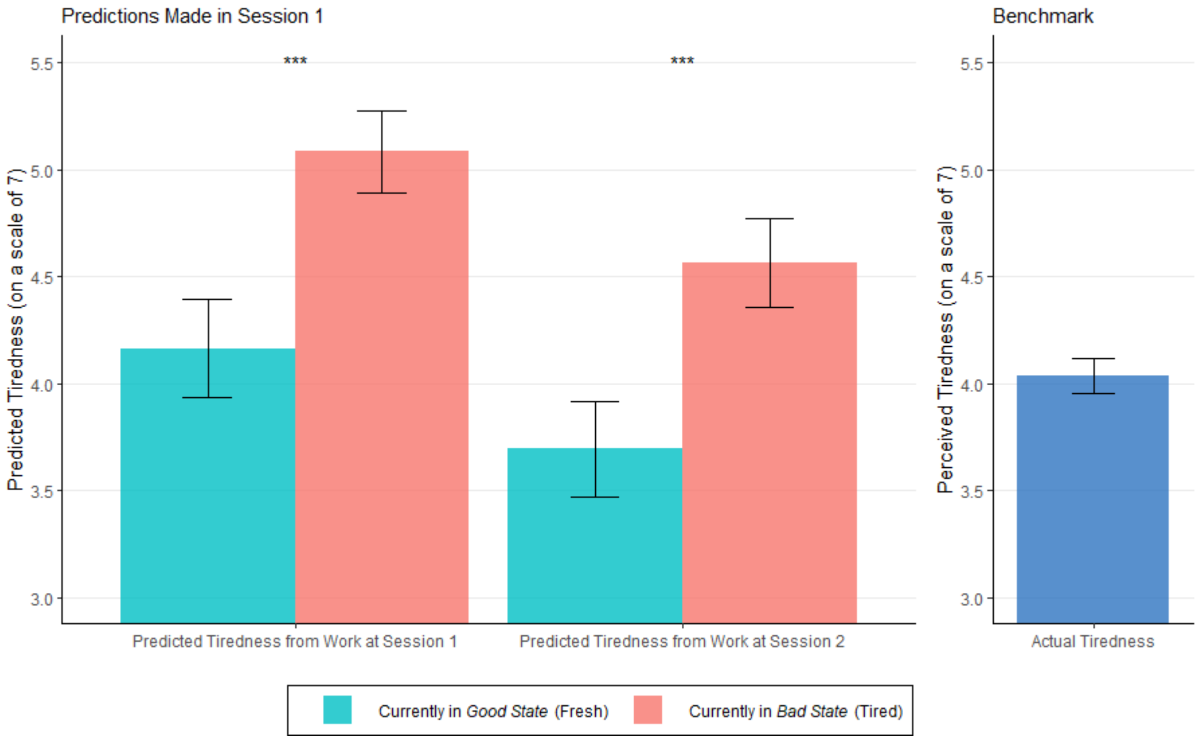
[60] See Noor and Takeoka (2021) for a theoretical analysis of an agent optimally incurring a cognitive cost to empathize with her future selves.

of the time horizon. The data suggests that the discounting interpretation is not likely to be the case because participants predict themselves to be less tired from additional work in two days than from additional work later today, as in Table E.1.

Nevertheless, this paper is not tailored to identify the lack of *intertemporal* empathy and the evidence here is suggestive at best. It would be interesting to better understand it in future research, because it may also impact the identification of time preferences and challenge the welfare analysis but in a different way.

## Appendix E.3 Whether Misperceptions Are Mistakes

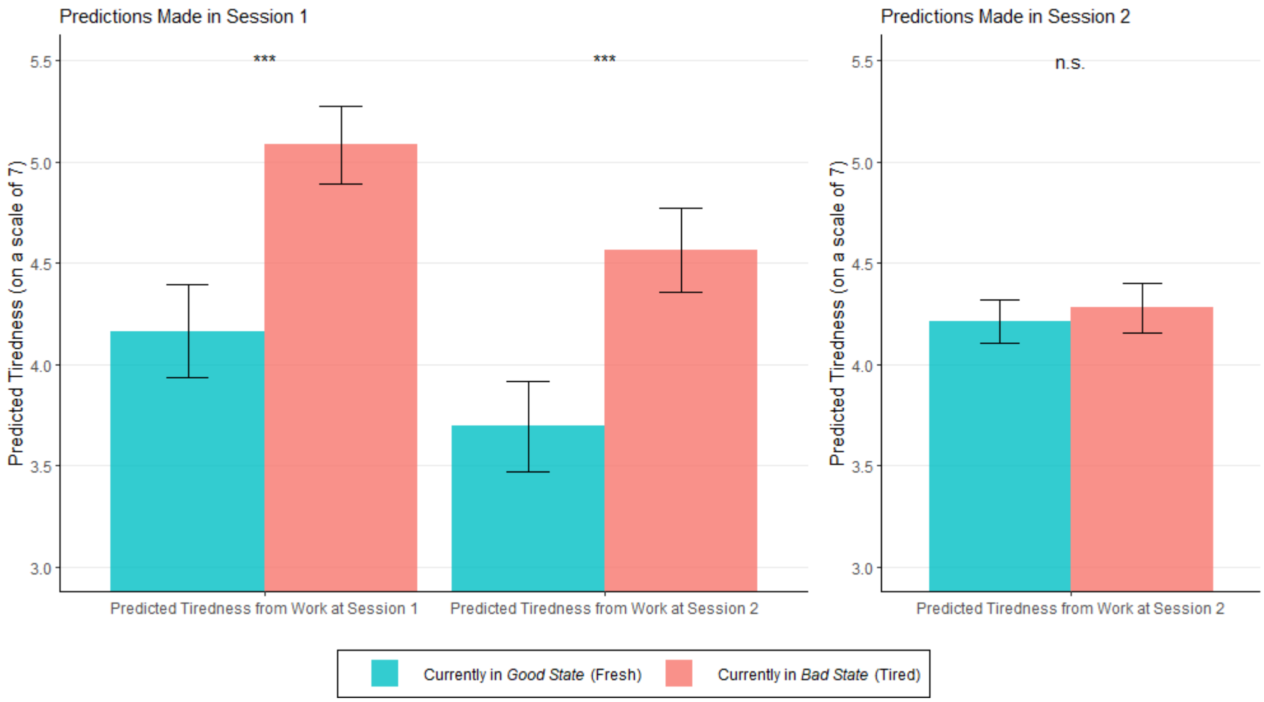
Figure E.2: Comparing Benchmark Perception with Predictions Made in Session 1




*Notes*: The left figure shows participants' predicted tiredness from additional work at end of Session 1 and Session 2 elicited in different treatments. The right figure shows the actually perceived tiredness elicited after the 2-minute rest and right before the additional task at the end of Session 2 as a benchmark. Error bars represent standard errors.

## Appendix E.4 Predictions Corrected by Experiences

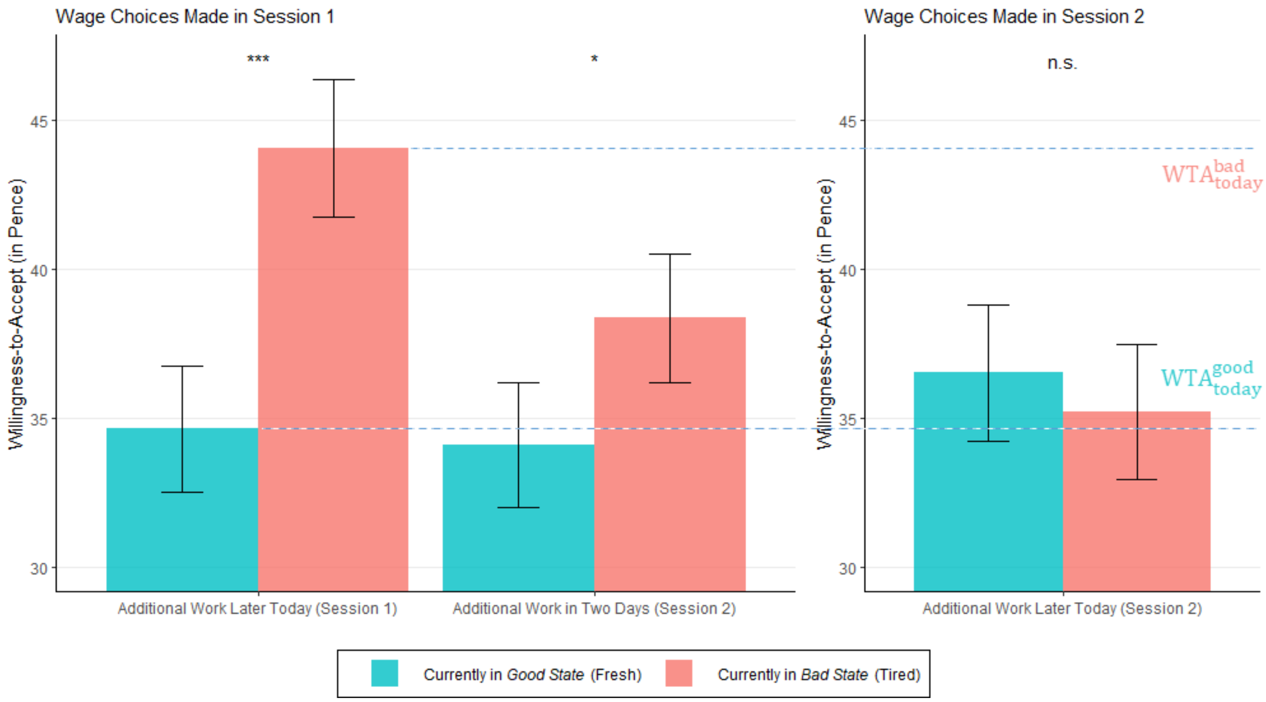
Figure E.3: Predictions Made in Session 1 (w/o Experience) and 2 (w/ Experience)




*Notes*: The left figure shows participants' predicted tiredness from additional work at end of Session 1 and Session 2 reported in Session 1 in different treatments. The right figure shows participants' predicted tiredness from additional work at the end of Session 2 reported in Session 2 in different treatments. Error bars represent standard errors.

## Appendix E.5 Wage Choices and Learning

Figure E.4: Wage Choices Made in Session 1 (w/o Experience) and 2 (w/ Experience)




*Notes*: The left figure shows participants' WTAs for additional work at end of Session 1 and Session 2 reported in Session 1 in different treatments. The right figure shows participants' WTAs for additional work at the end of Session 2 reported in Session 2 in different treatments. Error bars represent standard errors.

## Appendix E.6 The Null Effect of Task Performance

Table E.2: Wage Choices Predicted by Task Completion Time

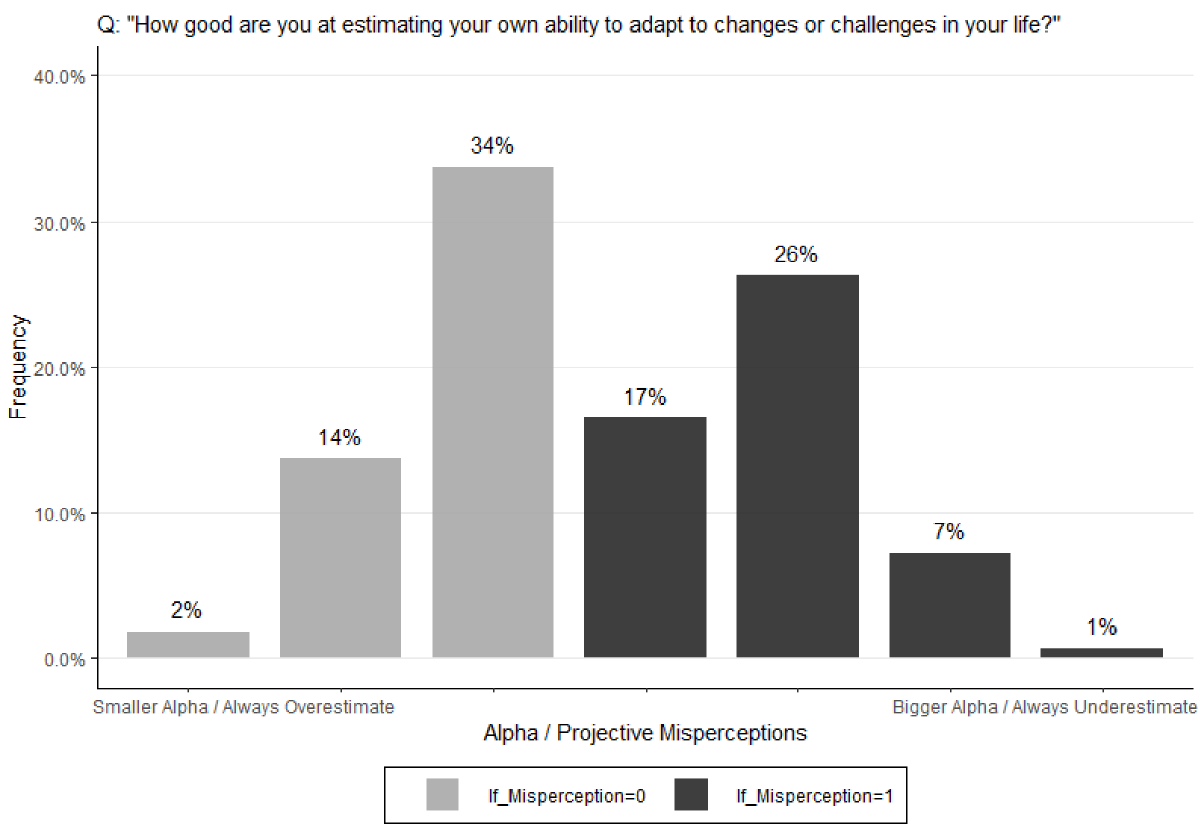
| Dep Var | (1)<br>WTA 1 | (2)<br>Prediction 1 | (3)<br>WTA 2 | (4)<br>Prediction 2 | (5)<br>WTA 3 | (6)<br>Prediction 3 |
|---|---|---|---|---|---|---|
| Completion Time | 0.00155 | 0.000336 | 0.000431 | 0.000436 | 0.00661 | 0.000629 |
| | (0.00763) | (0.000361) | (0.00715) | (0.000371) | (0.00757) | (0.000388) |
| Bad State | 9.559*** | 0.948*** | 4.312 | 0.905*** | -1.023 | 0.0961 |
| | (3.206) | (0.154) | (3.092) | (0.158) | (3.232) | (0.165) |
| Constant | 34.13*** | 4.056*** | 33.94*** | 3.554*** | 34.90*** | 4.055*** |
| | (3.344) | (0.158) | (3.177) | (0.162) | (2.927) | (0.150) |
| Observations | 411 | 430 | 409 | 430 | 389 | 408 |
| R-squared | 0.022 | 0.082 | 0.005 | 0.071 | 0.002 | 0.007 |

*Notes*: Robust standard errors in the parentheses. *** p<0.01, ** p<0.05, * p<0.1

## Appendix E.7 Survey Measure of Projective Misperceptions

The theoretical construct of misperception presumes that decision makers have little self-awareness. In contrast, participants in the experiment at least need to have some awareness of their general tendency for the measure of misperceptions to be meaningful, since the measure itself is based on self-reported answer to a single survey question. While a better understanding of misperceptions with sophistication is worth further research, the other way to reconcile this conflict is to appeal to self-knowledge as one sort of meta-awareness. While participants may not necessarily be aware of their misperceptions, their self-knowledge about the general tendency allows the survey measure to pick up meaningful information beyond noises.[61] In line with this interpretation, I find a positive correlation ($p < 0.01$) between the measure of misperceptions and the measure of sophistication about self-control (Ameriks et al., 2007; John, 2020) as both measures are based on self-reports, as shown in Table E.5. Following Figure E.5 shows the empirical distribution of projective misperceptions.

Figure E.5: Distribution of Projective Misperceptions




*Notes*: The figure shows the empirical distribution of projective misperceptions measured using the survey question.

[61] See Falk et al. (2021) for a theoretical framework on how self-knowledge influences survey response behavior.

## Appendix E.8 Placebo Tests of Other Control Variables

Table E.3: Wage Choices: Placebo Tests

| | Dep Var: WTA, 0 to 100 Pence | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Placebo Variables | Confidence | Regret | Empathy | Risktaking | Memory |
| Bad State | 7.825*** | 7.755** | -2.109 | 5.921* | 3.690 |
| | (2.958) | (3.208) | (4.812) | (3.030) | (2.719) |
| In Two Days | -3.240** | -3.179** | -3.223** | -3.188** | -3.141** |
| | (1.559) | (1.561) | (1.563) | (1.563) | (1.554) |
| Bad State × In Two Days | 0.120 | 0.0349 | 0.0882 | 0.0367 | -0.0374 |
| | (2.498) | (2.504) | (2.510) | (2.507) | (2.503) |
| Experience | -3.399** | -3.445** | -3.516** | -3.539** | -3.527** |
| | (1.385) | (1.395) | (1.383) | (1.383) | (1.389) |
| If_Confident | 3.707 | | | | |
| | (3.440) | | | | |
| Bad State × If_Confident | -7.006 | | | | |
| | (4.391) | | | | |
| If_Regret | | 0.333 | | | |
| | | (3.477) | | | |
| Bad State × If_Regret | | -7.040 | | | |
| | | (4.383) | | | |
| If_Empathetic | | | -4.369 | | |
| | | | (3.893) | | |
| Bad State × If_Empathetic | | | 7.871 | | |
| | | | (5.439) | | |
| If_Risktaking | | | | 4.000 | |
| | | | | (3.458) | |
| Bad State × If_Risktaking | | | | -3.289 | |
| | | | | (4.374) | |
| If_Recalled | | | | | -2.870 |
| | | | | | (3.355) |
| Bad State × If_Recalled | | | | | 0.768 |
| | | | | | (4.276) |
| Constant | 35.21*** | 37.08*** | 40.72*** | 35.19*** | 38.97*** |
| | (2.487) | (2.565) | (3.430) | (2.459) | (2.344) |
| Observations | 1,209 | 1,209 | 1,209 | 1,209 | 1,209 |
| R-squared | 0.010 | 0.013 | 0.010 | 0.009 | 0.009 |

*Notes*: Standard errors clustered at the individual level. *** p<0.01, ** p<0.05, * p<0.1

## Appendix E.9 Correlations among Survey Measures

Table E.4: Correlation among All Measures: Binary

| | If_Misperception | If_Patient | BetaDelta_Lower | Beta_Lower | If_Sophisticated |
|---|---|---|---|---|---|
| If_Misperception (1 or 0) | 1.000 | | | | |
| If_Patient (1 or 0) | -0.014 (0.771) | 1.000 | | | |
| BetaDelta_Lower (1 or 0) | -0.007 (0.892) | -0.001 (0.978) | 1.000 | | |
| Beta_Lower (1 or 0) | 0.028 (0.569) | 0.037 (0.439) | -0.408*** (0.000) | 1.000 | |
| If_Sophisticated (1 or 0) | 0.062 (0.201) | 0.011 (0.826) | -0.013 (0.792) | -0.016 (0.734) | 1.000 |

*Notes*: This table reports correlations between the survey measure of misperceptions and multiple measures of time preferences elicited in the post-experiment survey. All measures are in binary form and constructed based on a median split. $p$-values in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table E.5: Correlation among All Measures: Continuous

| | Misperception | Patience | BetaDelta | Beta | Sophistication |
|---|---|---|---|---|---|
| Misperception (from -3 to 3) | 1.000 | | | | |
| Patience (from -3 to 3) | 0.018 (0.713) | 1.000 | | | |
| BetaDelta (above 0) | -0.112** (0.020) | -0.070 (0.149) | 1.000 | | |
| Beta (above 0) | -0.028 (0.560) | -0.003 (0.948) | 0.207*** (0.000) | 1.000 | |
| Sophistication (from -10 to 10) | 0.131*** (0.007) | 0.002 (0.965) | -0.057 (0.241) | 0.021 (0.657) | 1.000 |

*Notes*: This table reports correlations between the survey measure of misperceptions and multiple measures of time preferences elicited in the post-experiment survey. All measures are in continuous form. $p$-values in parentheses. *** p<0.01, ** p<0.05, * p<0.1

In both Table E.4 and Table E.5, *BetaDelta* and *Beta* are significantly correlated ($p < 0.01$) since they are elicited using similar questions but with different time frames. The opposite sign, however, could be due to either the process of constructing the binary form or the inherent noisy nature of the measure.

In Table E.5, *Misperception* is significantly negatively correlated with *BetaDelta* ($p < 0.05$). This means that those who are more susceptible to projective misperceptions appear to discount the future more based on the corresponding measures, in line with the identification issue as characterized in Proposition 3.

## Appendix E.10 Predicting Exercise Using Continuous Measures

Table E.6: Predicting Self-reported Daily Behavior Using Continuous Measures

| | (1) Covid | (2) Procrastination | (3) Saving |
|---|---|---|---|
| Misperception | 0.105*** | 0.0881** | 0.0257 |
| | (0.0391) | (0.0342) | (0.0418) |
| Patience | 0.0217 | -0.0588 | 0.0961** |
| | (0.0464) | (0.0366) | (0.0421) |
| BetaDelta | 0.00944** | -0.0127*** | 0.00729* |
| | (0.00410) | (0.00268) | (0.00383) |
| Beta | -0.00254 | 0.0482*** | 0.00476 |
| | (0.0262) | (0.0131) | (0.0128) |
| Sophistication | 0.0190 | 0.0640*** | -0.0310 |
| | (0.0263) | (0.0213) | (0.0245) |
| Constant | 1.022*** | 0.392*** | 2.332*** |
| | (0.0829) | (0.0665) | (0.0732) |
| Observations | 430 | 430 | 430 |
| R-squared | 0.021 | 0.068 | 0.018 |

*Notes*: This table reports results from OLS regressions of three self-reported variables on measures of preference misperceptions and time preferences in continuous form. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

In above regressions using measures in continuous form, if anything, the coefficient of *Misperception* in predicting procrastination becomes stronger in isolation (from $p < 0.05$ to $p < 0.01$), and the coefficient of *BetaDelta* in predicting health investment becomes weaker in isolation (from $p < 0.05$ to $p < 0.1$). These changes in coefficients are consistent with the finding that *BetaDelta* is significantly negatively correlated with *Misperception* ($p < 0.05$).

The positive coefficient of *Beta* when predicting procrastination is inconsistent with the view that people procrastinate because they are present-biased. On interpretation is that *Beta* may have picked up very limited "present bias" presumably as the hypothetical measures are quite noisy. For instance, some participants trading off 100 Pounds over a time span of 30 days provide answers way beyond a reasonable range (e.g. above 2).