NBER WORKING PAPER SERIES

DYNAMIC PREFERENCE “REVERSALS” AND TIME INCONSISTENCY

Philipp Strack
Dmitry Taubinsky

Working Paper 28961
http://www.nber.org/papers/w28961

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2021

We thank Ned Augenblick, Yoram Halevy, Paul Heidhues, Eric Koepcke, Peter Maxted, Alex Rees-Jones, Charlie Sprenger, Tomasz Strzalecki, and seminar participants at the University of Zurich, Toulouse, Toronto, and Cornell for helpful comments. Afras Sial provided impeccable research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Dynamic Preference “Reversals” and Time Inconsistency
Philipp Strack and Dmitry Taubinsky
NBER Working Paper No. 28961
June 2021
JEL No. C9,D0,D9

ABSTRACT

We study identification of time inconsistency when an agent at time 0 makes an advance commitment, and later at time 1 can revise their choice after possibly receiving additional information. Roughly speaking, we prove that the only data that reject time-consistent expected utility maximization is a time-0 choice that is always strictly dominated at time 1. This holds for rich choice sets; if the complete ranking of alternatives is observed in every period and state; when it is natural to assume additional properties like concavity; and with supplementary cardinal information. However, time inconsistency can be point identified from willingness to pay for different alternatives in both periods, if utility from money is plausibly additively-separable and independent of time-1 information.

Philipp Strack
Yale University
philipp.strack@gmail.com

Dmitry Taubinsky
University of California, Berkeley
Department of Economics
530 Evans Hall #3880
Berkeley, CA 94720-3880
and NBER
dmitry.taubinsky@berkeley.edu

In his seminal work on myopia and dynamic inconsistency, Strotz (1955) posed the following question about an individual choosing "a plan of consumption for a future period of time": "If he is free to reconsider his plan at late dates, will he abide by it or disobey it?" A fundamental intuition arising from his work is that when individuals do not discount future consumption at a constant rate, they will have time-inconsistent preferences and often choose to revise their consumption plans. For example, individuals might exhibit *present focus*, as in the quasi-hyperbolic discounting model (e.g., Laibson, 1997, O'Donoghue and Rabin, 1999), and thus revise their plans toward more immediately gratifying alternatives over time.

This intuition raises a natural question that we take up in this paper: If an analyst observes individuals who tend to revise their consumption plans in a certain *systematic* direction, while being subject to random taste shocks, when can the analyst infer that the individuals are time-inconsistent, and when can the analyst quantify the degree of time inconsistency?

Although there is a surprising dearth of theoretical work on this basic question, it is central for interpreting a number of influential empirical studies that utilize what we call *revision designs*. A classic example is the study by Read and van Leeuwen (1998), which is often cited as "a canonical example of a preference reversal."[1] Read and van Leeuwen find that when making a plan a week in advance, approximately 50 percent of individuals choose a healthy over an unhealthy snack, but this fraction declines to approximately 20 percent when individuals are given a surprise opportunity to revise their plans a week later. The implicit assertion in the conclusions drawn by Read and van Leeuwen (1998) is that *if* individuals systematically revise their plans toward some types of alternatives over others, *then* they must have time-inconsistent preferences. Continuing with this assertion, economists and psychologists have conducted numerous revision design studies, often with richer choice sets, in domains such as intertemporal allocation of work, entertainment choice, financial plan-making, opioid use, and nutrition.[2,3]

This paper makes both theoretical and methodological contributions. The theoretical contribution is a general characterization of data sets that are consistent with time-consistent expected utility maximization (TC-EU) in different types of revision design settings. We introduce three different consistency requirements (no *simple/cyclic/stochastic dominance violations*) that vary in strength, and which are necessary and sufficient for the data to be rationalized by TC-EU in environments featuring (i) concave utility functions, (ii) single-peaked preferences, and (iii) arbitrary prior information of the analyst.

Because the consistency requirements are rarely violated in practice, the implication for applied work utilizing revision designs is that in the absence of difficult-to-verify assumptions, the degree of

[1] See, e.g., Ericson and Laibson (2019).

[2] Work allocation: Augenblick et al. (2015), Abebe et al. (2021), Andreoni et al. (2020), Barton (2015), Corbett (2016), Imas et al. (forthcoming), Kölle and Wenner (2019), Augenblick and Rabin (2019), Fedyk (2018) (see Imai et al., 2021 for a review); Financial plan-making: Kuchler and Pagel (2021); Entertainment: Read et al. (1999), Milkman et al. (2009), Bartos et al. (forthcoming); Opioid use: Badger et al. (2007); Nutrition choice: Sadoff et al. (2019).

[3] Typical designs elicit preferences from an identical choice set at two different points in time in an incentive-compatible way by informing individuals that their initial preferences and revised preferences both have a positive probability of determining their outcomes. Which preference is implemented is determined after both are elicited.

time inconsistency cannot generally be point identified; we show that seemingly minor variations in the assumptions made by the analyst can drive conclusions about time inconsistency. However, the second part of the paper addresses the non-identification results by proposing some special economic environments where the assumptions required for point identification are plausible. Collectively, our formal results provide guidance for theoretically-robust measurement.

Our model considers a data set of an agent's revealed ordinal preferences over a finite set of alternatives at two different points in time: an advance commitment stage (time 0) and a revision stage (time 1). The preferences can be completely or incompletely observed. At time 1, the agent's preferences are state-dependent; e.g., the agent's rankings over different food items might depend on their level of hunger. To consider a best-case scenario for identification, we suppose that the data set includes the agent's time-1 preferences in every (decision-relevant) state of the world.[4] We say that an agent's choices can be rationalized by time-consistent expected utility maximization (TC-EU) if there exists a utility function that rationalizes the agent's preferences in every realized state of the world at time 1, if the agent correctly perceives the probabilities of each state of the world, and if the agent's time-0 ranking of alternatives is consistent with the (objective) expectation of time-1 utilities.[5]

In the first part of the paper, we show that contrary to assertions like those of Read and van Leeuwen (1998), TC-EU can rationalize a variety of different data that feature systematic choice reversals. For example, consider a choice set consisting of a healthy snack and an unhealthy but more filling and tasty snack. Suppose that 80 percent of the time the agent prefers the less healthy snack by 1 util at time 1, but that 20 percent of the time the agent feels gorged and disgusted with themselves at time 1 and thus prefers the healthy snack by 5 utils. The state is not known at time 0, and the expected utility of choosing the healthy snack is higher than the expected utility of choosing the unhealthy snack: $0.2 \times 5 - 0.8 \times 1 > 0$. This generates a data set consistent with TC-EU preferences, despite the agent revising their time-0 choice of the healthy snack toward the more immediately gratifying option 80 percent of the time.

Data sets with richer choice sets—such as allocating consumption or effort between multiple periods—similarly cannot generally identify time inconsistency. We provide an example about an agent's consumption allocation decisions where if the analyst assumes that the agent learns about their time t taste shocks at time t, then the analyst must infer that the agent is present-focused, while if the analyst instead assumes that the agent learns about their time t taste shocks at time $t-1$, then the analyst must instead infer that the agent is future-focused. The analyst can also

[4] In practice, data sets do not have more than several observations of an individual's propensity to revise their choices. In practice, a typical assumption that facilitates identification is that individuals who make the same choice in time 0 are homogeneous in their preferences and economic environments, and any differences in time-1 choices are due to independent realizations of time-1 taste shocks.

[5] To be clear, belief-based biases that generate behavior resembling time inconsistency, such as the *planning fallacy* (Kahneman and Tversky, 1982, Buehler et al., 2010, Brunnermeier et al., 2008), overoptimism (Browning and Tobacman, 2015, Breig et al., 2021) or other misperceptions of the time-1 decision environment (see, e.g., Sadoff et al. 2019 for other examples), are violations of TC-EU in our framework as the time-0 preference is not derived by taking the (correct) expectation over time-1 utilities. Our results thus imply that such biases can also not be identified from revision designs.

rationalize time consistency with an information structure in between these two extremes.

Formally, our Theorem 1 shows that when the choice set can be equated to a subset of the real line, when time-1 preferences are single-peaked, and when at least the time-0 preference is completely observed, the data can rationalized by TC-EU as long as there is no *simple dominance violation*—i.e., alternatives x_1 and x_2 such that x_1 is preferred to x_2 in time 0, but where x_2 is preferred to x_1 *with probability 1* at time 1. In the general case where preferences are incompletely observed, Theorem 1 states that the data can be rationalized by TC-EU under the more general requirement that there is no *cyclic dominance violation*—i.e., a set of alternatives $x_1, \ldots.., x_k$ such that x_{j+1} is preferred to x_j either in time 0 or with probability 1 at time 1, and such that also x_1 is preferred to x_k either in time 0 or with probability 1 at time 1.

Theorem 2 shows that under the additional assumption where the analyst knows the utility functions to be strictly concave in both time 0 and time 1, the data are consistent with TC-EU if and only if there is no simple dominance violation and the time-0 ranking of alternatives is single-peaked. Interestingly, the condition that the time-0 preference is single-peaked implies that it is not necessary to check for cyclic dominance violations more generally.[6]

Outside of the economic environments considered in our first two theorems, we show by example that TC-EU can be rejected even when there are no cyclic dominance violations. The reason is that the lack of cyclic dominance violations does not necessarily exclude the possibility of a *stochastic dominance violation*—which we define to be the case where lottery L is revealed to stochastically dominate lottery L' according to time-1 preferences, but where L' stochastically dominates L according to time-0 preferences. (However, existence of cyclic dominance violations implies existence of stochastic dominance violations.) Theorem 3 shows that in the general case, a data set is consistent with TC-EU if and only if there are no stochastic dominance violations. Theorem 4 generalizes this result to the case where the analyst obtains (or assumes) additional cardinal information by directly observing (or assuming) an agent's preferences over a set of lotteries.

Additionally, we provide sufficient conditions for when data sets do not exhibit stochastic dominance violations. One such condition is that every element of the choice set is the most preferred element at time 1 with positive probability. This *positivity* condition applies to commonly-used stochastic discrete choice models, such as the Luce model.

While the first part of the paper focuses on characterizing data sets consistent with TC-EU, the second part of the paper presents a series of results about rejecting and identifying time inconsistency. To highlight our view that an inability to reject TC-EU is not actually compelling evidence for accepting TC-EU as the "right" model, we begin the second part of the paper with a mathematically trivial result that shows that it is even more difficult to reject time inconsistency than it is to reject time consistency.

We then propose a special class of economic environments in which point identification of

[6] Theorem 1 and 2 turns out to be mathematically equivalent to the statement that under certain conditions arising from the economic assumptions, non-intersecting convex cones can be separated by a hyperplane with just two non-zero entries (or only $0, +1, -1$ entries), which might be of independent interest and have applications beyond the specific economic context studied here. We discuss possible other applications in Section 5.4.

the degree of time inconsistency—e.g., point identification of the present focus parameter in the quasi-hyperbolic discounting model—is feasible. These environments feature preferences that are additively separable over a set Y of alternatives such as food options or effort allocations and a set Z of a "numeraire" commodity such as money. The analyst must have enough range in Z to be able to measure the agent's relative willingness to pay (WTP) for one alternative of Y over another, both at time 0 and in every state at time 1. Point identification is obtained if the agent's utility function over Z is known (e.g., is quasi-linear in money) and does not vary with taste shocks to alternatives in Y. All features of this economic environment are plausibly present in the recent work of Augenblick and Rabin (2019), Augenblick (2018), and Fedyk (2018), which illustrates the feasibility of point identification.

The state-independence of utility over Z is crucial. For example, we show that if utility over both Y *and* Z is subject to multiplicative taste shocks, but the functional form of utility over Z is otherwise known, then the present focus parameter β in the quasi-hyperbolic discounting model can only be set identified. Concretely, consider an analyst who treats all differences in time-1 choices as due to differences in realized β (for example, due to population heterogeneity or stochasticity in the present focus parameter), thus estimating a separate $\hat{\beta}_i$ for each observed preference profile. We show that the true value of β, accounting for taste shocks, can be anything in the interval $[\min \hat{\beta}_i, \max \hat{\beta}_i]$. This characterization is directly applicable to studies that report distributions of such estimates $\hat{\beta}$, and as we show often leads to a wide range of estimates, including $\beta = 1$.

Our results about the difficulty of identification arise due to random taste shocks and other decision-relevant information being realized between time 0 and time 1. Toward the end of the paper, we discuss the types of empirical tests that can be used in experiments to gauge the extent to which time-1 revisions are due to random taste shocks rather than individual differences in time preferences. One test is to directly show, as in Read and van Leeuwen (1998), that changes in some state variable such as hunger lead to different choices in time 1. Another test is to note that if differences in time-1 revisions are due largely to stable individual differences in time preferences, then data sets in which choice revisions are observed multiple times for each individual should display near-perfect stability in individuals' choice *revisions*. However, we show in Section 5.1 that data from recent work suggests that only a small fraction of the variance in time-1 choice revisions can be explained by stable individual differences in time preferences, which suggests a quantitatively large role for random taste shocks or the revelation of other decision-relevant information.

The rest of this paper proceeds as follows. Section 1 presents the formal model. Section 2 presents several numerical examples illustrating why it is generally difficult to point identify the degree of time inconsistency. Section 3 presents our main theorems about the types of data sets that are consistent with TC-EU—this section is arguably the central mathematical contribution of this paper. Section 4 presents results about economic environments where it is possible to point or set identify the degree of time inconsistency. Section 5 connects our results to practical empirical questions like the degree to which random taste shocks affect behavior, discusses some extensions, and summarizes potential applications of our mathematical results to other literatures in economics.

Section 6 concludes, clarifying that while we hope to budge the frontier on theoretically-robust measurement of time inconsistency, we do not question its existence. All proofs are relegated to the appendix.

1 Model

Preference data set There is an agent who has a preference over a finite set of alternatives X at time 0 and time 1.[7] Their preference at time 0 is deterministic and denoted by $\preceq^0$. The agent's preference $\preceq^1$ at time 1 is a random draw from $(\preceq^1_1, \ldots, \preceq^1_n)$, with $n < \infty$. We denote by $(f_1, \ldots, f_n)$ the strictly positive probabilities (or frequencies) associated with each realization.

We say that a preference is complete if the ranking of any two alternatives is observed. We allow for time-0 and time-1 preferences to be incomplete to capture the practically-relevant case where the analyst observes only limited information about the agent's preferences. An example of such limited information would be if the analyst only observes the alternative the agent likes best, or the agent's top two alternatives, but has no information about the ranking of other alternatives. We do not require $\preceq^0$ or any of the $\preceq^1_j$ to relate the same pairs of alternatives, or even the same number of alternatives.

Throughout, we assume that if two alternatives $x, y \in X$ are related by the (potentially incomplete) preference $\preceq$, then the analyst observes whether the agent is indifferent or prefers one of the alternatives strictly; i.e., either $x \sim y$, or $x \prec y$, or $y \prec x$.[8] Note that if strict preferences cannot be observed, then any set of decisions is trivially consistent with TC-EU, where the utility function assigns the same value to each alternative.

A *data set*

$$(\preceq^0, \preceq^1_1, \ldots, \preceq^1_n, f_1, \ldots, f_n)$$

consists of a time-0 preference $\preceq^0$ and a probability distribution over the possible time-1 preferences, which we compactly write as $(\preceq^1, f)$. We refer to $\preceq^1$ as a random time-1 preference.

Time-Consistent Expected Utility Preferences A time-consistent expected-utility (TC-EU) agent evaluates alternatives according to the utility function

$$u : X \times \Omega \to \mathbb{R}$$

that depends on the chosen alternative $x \in X$ and state $\omega \in \Omega$. The states capture taste shocks or information that arrives between time 0 and time 1, and the agent does not know the state at time 0, but observes it before time 1. Our convention is to denote the state by a subscript and the alternative as an argument, so that $u_\omega(x)$ denotes the utility of alternative x in state ω.

[7] We make the assumption that the choice set is finite to avoid technicalities and streamline the presentation. We put no bound on the size of the choice set and in all empirical applications the choice set is necessarily finite.

[8] Formally, $\preceq$ is a preorder where we interpret $x \preceq y$ and $y \preceq x$ as indifference $x \sim y$, and $x \preceq y$ but not $y \preceq x$ as a strict preference for u over x. The preference between x,y is unobserved if neither $x \preceq y$ nor $y \preceq x$. If the ranking of any two alternatives is observed then $\preceq$ is a complete preorder.

Without loss, we assume that there as many states as realizations of the time-1 preference, with each realization of the time-1 preference corresponding to a state, so that[9]

$$\Omega = \{1, \ldots, n\}.$$

The TC-EU agent prefers alternative x over y in state ω at time 1 if and only if x has a higher associated utility; i.e., for all $x, y \in X$, $\omega \in \Omega$,

$$y \preceq^1_\omega x \quad \Leftrightarrow \quad u_\omega(y) \leq u_\omega(x)\,. \tag{1}$$

At time 0 the TC-EU agent prefers x over y if and only if the expected utility of x exceeds the expected utility of y; i.e., for all $x, y \in X$,

$$y \preceq^0 x \quad \Leftrightarrow \quad \sum_{\omega \in \Omega} f_\omega u_\omega(y) \leq \sum_{\omega \in \Omega} f_\omega u_\omega(x)\,. \tag{2}$$

Definition 1 (Consistency with TC-EU)**.** A data set $(\preceq^0, \preceq^1, f)$ is consistent with TC-EU if there exists a utility function $u : X \times \Omega \to \mathbb{R}$ that satisfies (1) and (2).

Consistency with TC-EU means that there exists a state-dependent utility function that represents $\preceq^1_\omega$ in each state ω, and such that the expectation of this utility function represents $\preceq^0$.

Remarks on the Model

Our definition of TC-EU requires the agent to correctly understand the distribution of states. Thus, time-inconsistent-like behavior generated by belief-based biases such as the *planning fallacy* (Kahneman and Tversky, 1982, Buehler et al., 2010, Brunnermeier et al., 2008) or other forms of overoptimism (Browning and Tobacman, 2015, Breig et al., 2021) is not compatible with our definition of TC-EU. Thus, our results about an inability to reject TC-EU also imply that one cannot identify belief biases from revision designs.

Relatedly, we assume the domain of TC-EU preferences to be the set of alternatives X, which rules out models of *costly self-control* (e.g., Gul and Pesendorfer, 2001, Fudenberg and Levine, 2006). Thus, data sets that we show to reject TC-EU could be consistent either with time-inconsistent preferences or with costly self-control preferences.

We assume that the agent's preference at time 0 is deterministic. However, the agent might also experience taste shocks at time 0. Assuming deterministic time-0 preferences is without loss if the analyst observes the time-1 choices conditional on the agent's time-0 choice: in this case,

[9] Assuming that each state corresponds to an observed preference profile is without loss. To see this, note that if we have a TC-EU representation (u, Ω, F), consisting of a utility u, a state space Ω, and a prior F that is consistent with the ordinal preferences $(\preceq^0, \preceq^1)$, then without loss of generality we can associate each set Ω_k of states that leads to a preference profile $\preceq^1_k$ with the newly defined state k. We define a utility function on this new state space as the conditional expectation $\tilde{u}_k(x) = \frac{\int_{\Omega_k} u_\omega(x) dF}{\int_{\Omega_k} dF}$ and obtain a new EU representation with the desired state space $\{1, \ldots, n\}$.

our results apply to the conditional choice distribution.[10] We do not formally introduce random time-0 choice in the body of the paper as it adds no insight and complicates notation, although we recognize its empirical relevance and address it in Supplementary Appendix B.

Our assumptions on the data set facilitate identification of time-inconsistent behavior in two ways. First, we assume that the analyst observes the exact distribution of time-1 preferences, which corresponds to the case where the analyst observes the agent's behavior in exactly the same situation infinitely often. In practice, the analyst can gather only finitely many observations on each individual by observing their choices over time (and assuming that the individual faces the same economic environment each time). Alternatively, the analyst could assume that all individuals making the same time-0 choice have the same random time-1 preference (with independent realizations across the agents).

Second, we allow the analyst to (potentially) not only observe the agent's most preferred alternative, but the complete ranking over alternatives, which further facilitates identification. While some empirical studies collect richer data than each individual's most preferred alternatives, not many existing data sets have complete information on agents' preferences.

2 Motivating Examples

We start with several examples, based on existing studies, to illustrate the model and the intuition behind some of our main results on the difficulty of identifying time inconsistency.

Example 2.1 (Food Choice)**.** Consider an agent who chooses between a healthy snack $x = H$ and an unhealthy snack $x = N$. In the introduction, we gave an example where at time 1, the agent feels gorged with probability 20 percent, in which case they crave cleansing healthy food, so that the utility difference between the unhealthy and healthy snack equals -5. With probability 80 percent the utility difference between the unhealthy and healthy snack equals 1. At time 0 it is optimal for the agent to choose the healthy snack if they do not know the time-1 state. In the notation of the model, the data set is $N \prec^0 H, H \prec^1_1 N, N \prec^1_2 H, f = (0.8, 0.2)$.[11]

This example suggests an alternative interpretation of the results reported by Read and van Leeuwen (1998). In their experiment, participants choose between a healthy and an unhealthy snack at time 0, to be delivered at time 1 (after seven days). Then, at time 1, participants are given a surprise opportunity to revise their time-0 choice. On average (collapsing across individuals and conditions), subjects at time 0 choose the healthy snacks 49 percent of the time, but at time 1 choose healthy snacks 17 percent of the time.[12] Moreover, Read and van Leeuwen report that switching to the unhealthy snack was far more common than switching to the healthy snack.

[10] If the analyst observes only the marginal distributions of preferences, she has strictly less information and identification of time-inconsistency becomes harder, which means that all our non-identification results apply.

[11] Here $\omega = 1$ denotes the state where the agent prefers N at time 1 and $\omega = 2$ denotes the state where they prefer H at time 1. In this example the agent is never indifferent.

[12] These summary statistics are reported in Cohen et al. (2020).

To obtain the results of Read and van Leeuwen with TC-EU preferences, consider a population where 35 percent of the agents have preferences and an information structure as summarized in our first paragraph of this example. Suppose that the remaining 65 percent have the same preferences, but already know the state of the world at time 0. Thus, an additional $0.65 \cdot 0.2 = 13$ percent choose the healthy snack at time 0, knowing that they will prefer it for sure at time 1, while $0.65 \cdot 0.8 = 52$ percent choose the unhealthy snack at time 0. Thus, overall, 48 percent choose the healthy snack at time 0, but only 20 percent choose it at time 1. Moreover, the direction of revisions is asymmetric: many switch from choosing the healthy snack at time 0 to choosing the unhealthy snack at time 1, but no one switches from choosing the unhealthy snack to choosing the healthy snack.

Example 2.2 (Consumption, Savings, Labor, and Leisure)**.** Suppose that at time 0, an agent decides what share $x \in \{0, 1/42, 2/42, \ldots, 41/42, 1\} \subset [0, 1]$ of total resources to consume at time 1, and what share $1 - x$ to leave for time 2. For example, x could correspond to the share of income spent at time 1, or the amount of work put off until time 2.

The analyst observes the most preferred alternative. At time 0, the agent chooses an equal split between time 0 and time 1, $x^0 = 1/2$. At time 1, there are two equally likely states, with choices $x^1 = 5/7$ and $x^1 = 1/3$, leading to time-1 vs time-2 consumption ratios $\frac{x^1}{1-x^1} = 5/2$ and $\frac{x^1}{1-x^1} = 1/2$, respectively. Thus, while at time 0 the agent prefers an equal allocation of consumption between time 1 and time 2, at time 1 the agent on average chooses more consumption for time 1.

Suppose it is known that the agent evaluates consumption paths at time 0 according to

$$\mathbb{E}_0 \left[\beta\, \theta^1 \log(x) + \beta\, \theta^2 \log(1 - x)\right]$$

and at time 1 according to

$$\mathbb{E}_1 \left[\theta^1 \log(x) + \beta\, \theta^2 \log(1 - x)\right],$$

where θ^1 and θ^2 are random taste shocks and $\mathbb{E}_0$ and $\mathbb{E}_1$ refer to the expectations with respect to time-0 and time-1 information, respectively. Suppose that the means of both the time-1 and time-2 taste shock equal 1, which is known (or assumed) by the analyst.

It turns out that despite significant information about the agent's utility function, the value of β cannot be identified. The data set can be rationalized with the following values of β and assumptions about the revelation of information:

Case 1: **Present focus.** The agent learns θ^t at the beginning of time t, without learning anything about θ^{t+1}. This implies $\beta = 2/3$.

Case 2: **Future focus.** The agent learns θ^t at the beginning of time $t - 1$. This implies $\beta = 1.2$.

To see the logic behind these inferences, note that with log utility, the choice of x depends only on the ratio of expected multipliers, and thus satisfies $\frac{x^1}{1-x^1} = \frac{\mathbb{E}_1\theta^1}{\beta\mathbb{E}_1\theta^2}$ at time 1 and $\frac{x^0}{1-x^0} = \frac{\mathbb{E}_0\theta^1}{\mathbb{E}_0\theta^2}$ at time 0. In the first case, the data and the assumption that $\mathbb{E}_0\theta^2 = \mathbb{E}_1\theta^2 = 1$ imply that $(\mathbb{E}_1\theta^1, \beta\mathbb{E}_1\theta^2)$ is uniform on $\{(5/2\beta, \beta), (1/2\beta, \beta)\}$; thus, $\beta = 2/3$ since the mean of θ^1 is 1. In the second case, the data and the assumption that θ^1 is learned at time 0 imply that $(\mathbb{E}_1\theta^1, \beta\mathbb{E}_1\theta^2)$ is uniform on $\{(\theta^1, {}^2\!/\!{}_5\theta^1), (\theta^1, 2\theta^1)\}$; thus, $\beta = 1.2$ since the means of the taste shocks are 1.

Other learning processes can rationalize other assumptions about β. For instance:

Case 3: **Time consistency**: $\beta = 1$ and $(\mathbb{E}_1\theta^1, \mathbb{E}_1\theta^2)$ is uniform on $\{(2.5, 1), (1.5, 3)\}$.

The time consistency case results from negatively correlated shocks. E.g., if the agent is allocating effort over time, but must wait until time 1 to find out if another time draining activity must be completed at time 1 or at time 2. In sum, these different cases illustrate that while assumptions of when and how agents learn about their taste shocks may seem subtle, they can fully drive an analyst's inferences about the degree and nature of time inconsistency.

Further Intuition for Non-Identification in the Examples

Fundamentally, the lack of identification in both examples is a consequence of the analyst not knowing how the agent "weights" the different time-1 states when they make their decision at time 0. In Example 2.1, the analyst does not know the intensity of preference for one versus the other snack in the two different states. And as Example 2.2 illustrates, richer choice sets, together with some arguably strong parametric assumptions, do not resolve this problem.

To discuss the intuition behind Example 2.2, suppose that x corresponds to the fraction of work left for time 2. At time 1, the agent might choose to put off work relative to the time-0 plan for a number of reasons, including the following two: (i) because they experience a time-draining productivity shock at time 1 (without learning anything new about time-2 shocks relative to their time-0 prior), or (ii) because they learn that they will be relatively less busy at time 2 (without learning anything new about their availability at time 1 that they didn't already know at time 0). The first reason to put off work is captured by case 1 in Example 2.2, and leads the analyst to infer that the agent is present-focused. Intuitively, in this case the agent's marginal cost of effort is relatively high in the state where they put off work; thus, in the absence of present focus, the time-0 allocation would have to closely resemble the time-1 allocation in the state of the world where the agent puts off most of the work until time 2. The second reason is captured by case 2 in Example 2.2. Here, the agent at time 0 puts more weight on the state of the world where they prefer to do more of the work at time 1.

Since both cases are consistent with the data, β cannot be identified. There is also no obvious "intuition" that can help adjudicate between whether it is more correct to assume the first or the second case, or something in between (e.g., the time-consistency rationalization in case 3). In fact, there are other plausible assumptions on the information structure that would further widen the set of possible values of β.

Together, the two examples illustrate that many of the revision design papers that produce point estimates of the degree of time inconsistency do so through an assumption—typically embedded implicitly in the estimating equations—about difficult-to-verify properties of the distributions of taste shocks.[13]

[13] For example, Sadoff et al. (2019) assume symmetric mean-zero taste shocks to utility from food options. Although they thoughtfully examine several observable determinants of the taste shocks, the core symmetry assumption of

Additional examples. Empirical designs with structures similar to Example 2.1 include the food-delivery field experiment of Sadoff et al. (2019), Read et al.'s (1999) study of choice between high-brow and low-brow video rentals, and Milkman et al.'s (2009) quasi-experimental extension of Read et al. (1999).

In Example 2.2, the choice of x could correspond to consumption and $1 - x$ to savings, as in frequently-studied consumption-savings applications of time-inconsistent preferences. For example, Kuchler and Pagel (2021) elicit people's time-0 plans for debt paydown, corresponding to a time-0 choice of x, and study deviations from people's subsequent actions, corresponding to a time-1 choice of x. Alternatively, the choice of x could correspond to time-2 effort, as in the experiments of Augenblick et al. (2015), Andreoni et al. (2020), and many others, where subjects are faced with a fixed set of tasks that they must allocate between two different points in time. Other experiments with this structure include Abebe et al. (2021), Barton (2015), Corbett (2016), Imas et al. (forthcoming), Kölle and Wenner (2019).

3 Rejecting Time Consistency

In this section, we characterize data sets that are consistent with TC-EU. Section 3.1 considers single-peaked preferences, Section 3.2 considers the case where the analyst knows that the utility function is concave in every state, and Section 3.3 considers the general case where there is no natural ordering on alternatives that would lead to single-peaked or concave preferences. In Section 3.4 we allow the analyst to possess supplementary cardinal information, either through a-priori assumptions or by observing preferences over a set of lotteries.

3.1 Single-Peaked Preferences

Here we consider the case where the alternatives are real numbers in $X \subset \mathbb{R}$, and where the time-1 preferences are single-peaked, defined as follows:

Definition 2 (Single-Peaked Preference)**.** A potentially incomplete preference $\preceq$ is single-peaked if for any alternatives $x < y < z$ either $x \prec y$ or $z \prec y$.

their estimating equation is empirically unverifiable, which is potentially consequential since the majority of the variance in decisions is attributable to these taste shocks (see Section 5.1). Similarly, Augenblick et al.'s (2015) estimating equation (6) assumes mean-zero additively separable errors, which is equivalent to a particular assumption on the distribution of taste shocks that guarantees that they can be ignored without bias. Augenblick et al. (2015) acknowledge that this is not necessarily without loss, as "the minimizer of the expectation need not be the expectation of the minimizer" and begin to explore the implications of uncertainty in Appendix B. In Appendix B, they assume that the uncertainty is only on the *curvature* of the cost of effort functions, and that the shocks to time 1 and 2 cost of effort functions are perfectly correlated. Under these assumptions, they find that uncertainty generates moderate upward bias in their estimate of the present focus parameter β. As our results show, other forms of taste shocks—in particular those that affect the cost in both periods differentially—have much more significant implications for identification of present focus, and can generate either upward or downward bias. Relatedly, our additional analysis in Supplementary Appendix B.2.1 shows that Augenblick et al.'s supplementary data on commitment contract take-up does not substantially narrow the set of time preferences consistent with their data.

We say that the random preference $\preceq^1$ is single-peaked if each possible realization $(\preceq^1_\omega)_{\omega\in\Omega}$ is single-peaked. The single-peaked property is natural in environments where agents choose how to allocate consumption or effort over time, as in Example 2.2. If the utility from consumption is concave in each period, then the agent's utility function will be concave and thus single-peaked in the share of work done in time 1. The single-peaked property also mechanically applies to binary choice sets, as in Example 2.1. We consider the slightly stronger assumption of concavity in the next subsection.

Note that our definition of single-peaked preferences applies to incomplete preferences, where for example all that is known is that an agent's most preferred alternative is some $x^* \in X$, which implies that moving further left or right of x^* leads to less attractive alternatives, but contains no information about how the agent compares alternatives to the left of x^* against those to the right of x^*.

A natural requirement on a data set is that the agents' choices do not contradict themselves. More precisely, if at time 1 the agent *always* prefers some alternative over another then they should also prefer that alternative at time 0. The next definition formalizes this idea.

Definition 3 (Simple Dominance Violations)**.** A data set $(\preceq^0, \preceq^1, f)$ exhibits simple dominance violations if there exist $x, y \in X$ such that $x \preceq^0 y$, $y \preceq^1_\omega x$ for all $\omega \in \Omega$, and either $x \prec^0 y$ or $y \prec^1_\omega x$ for some $\omega \in \Omega$.

For example, if the analyst observes that the agent *always* prefers the unhealthy over the healthy snack at time 1, then the agent should also prefer the unhealthy snack at time 0 if they are time-consistent. It follows immediately from the definition that any data set that is consistent with TC-EU cannot exhibit simple dominance violations. Theorem 1 below presents a converse of that statement when the time-0 preference is complete and the time-1 preference is single-peaked.

To provide a general characterization of incomplete preferences, we introduce a generalization of simple dominance violations. To simplify notation we denote by $\preceq^1_*$ the preorder that is generated by agreement of the agent's preferences in all states ω: $x \preceq^1_* y \Leftrightarrow x \preceq^1_\omega y$ for all ω.

Definition 4 (Cyclic Dominance Violations)**.** A data set $(\preceq^0 \preceq^1, f)$ exhibits cyclic dominance violations if there exists a sequence of alternatives $x_1, x_2 \ldots, x_k$ that are alternatingly ranked by the orders $\preceq^0, \preceq^1_*$,

$$x_1 \preceq^0 x_2 \preceq^1_* x_3 \preceq^0 \ldots \preceq^0 x_k \preceq^1_* x_1,$$

with at least one relation strict.[14]

Clearly, a simple dominance violation is a cyclic dominance violation with a cycle of length 2. And as Lemma 3 in Appendix A.1 shows, cyclic dominance violations imply simple dominance violations when the time-0 preference is complete. In general, however, data sets can exhibit cyclic

[14]We only need to consider cycles where the order is generated between alternations between $\preceq^1_*$ and $\preceq^0$ as due to transitivity we can always remove elements that are bounded from above and below in the same order. Any such cycle has an even number of elements.

dominance violations without exhibiting simple dominance violations. To illustrate, consider the following example:

Example 3.1. There are four alternatives $X = \{1, 2, 3, 4\}$ and only one state, with single-peaked time-1 preference $\preceq^1$ in that state. The time-1 (incomplete) preference is $1 \prec^1 2$ and $4 \prec^1 3$, while the time-0 (incomplete) preference is $2 \prec^0 4$ and $3 \prec^0 1$. Since each preference relates a different pair of alternatives, there is no simple dominance violation. However, no utility function is consistent with both the time-0 and time-1 preferences, as the time-1 preference would imply that $u(1) < u(2)$, the time-0 preference would imply that $u(2) < u(4)$, the time-1 preference would also imply that $u(4) < u(3)$, which would then imply that $u(1) < u(3)$, violating the time-0 preference. The cycle here is $1 \prec^1 2 \prec^0 4 \prec^1 3 \prec^0 1$.

Theorem 1 (Consistency with TC-EU)**.** *Consider $X \subset \mathbb{R}$ and let $(\preceq^0, \preceq^1, f)$ be a data set with strict and single-peaked time-1 preferences.*

(i) The data set is consistent with TC-EU if and only if it exhibits no cyclic dominance violations.

(ii) If $\preceq^0$ is complete, then the data set is consistent with TC-EU if and only if it exhibits no simple dominance violations.

Theorem 1 shows that it could be difficult to reject TC-EU using revision designs, because doing so requires a dominance violation. For example, if at time 0 the agent chooses to allocate a fraction $x = 1/2$ of resources to time 1 (and the remainder $x = 1/2$ to time 2), then TC-EU can only be rejected if at time 1 the agent *always* revises to allocate $x > 1/2$ resources to time 1.

How demanding of a test this is depends on the randomness of taste shocks at time 1. In a deterministic environment (i.e., $|\Omega| = 1$) with a rich choice set, an agent with dynamically inconsistent time preferences will typically exhibit dominance violations. However, the more variability there is in an agent's time-1 preference, the more unlikely dominance violations become, even if the agent does have dynamically inconsistent time preferences.

Thus, Theorem 1 suggests that while revision designs can provide discerning tests of time inconsistency in deterministic environments, they are fairly uninformative about whether an agent is time-consistent or time-inconsistent when there is significant stochasticity in time-1 choices.[15] We discuss how to gauge stochasticity in Section 5.1.

3.2 Concave Utilities

An additional plausible restriction is that the agent's time-1 utility is concave in each state. For example, it is natural to assume that each period, the utility from consumption is concave, or that the cost of effort is convex. If the agent decides what share of resources x to allocate to time 1,

[15]Theorem 1 and related results about ordinal preference data sets can be generalized from TC-EU to recursive preferences that nest TC-EU, such as the Epstein and Zin (1989) preferences. To see this, note that any time-0 preference that is a monotonic (but not necessarily linear) function of time-1 utilities cannot generate dominance violations. Conversely, as TC-EU is a recursive preference, any data set consistent with TC-EU is also consistent with recursive preferences. We thank Yoram Halevy for pointing this out.

and what share $1-x$ to allocate to time 2, and if $u_\omega(x) = v^1_\omega(x) + v^2_\omega(1-x)$, with v^1 and v^2 both concave, then u_ω will be concave as well. We say that a data set is consistent with *concave TC-EU* if is consistent with TC-EU for a strictly concave utility function.

It is immediate that any concave utility function leads to single-peaked preferences. Moreover, because the expectation of a strictly concave function is itself strictly concave, it follows that the time-0 utility must be concave whenever all time-1 utilities are concave. This immediately implies that a necessary condition for a data set to be consistent with concave TC-EU is that the time-0 preference must be single-peaked. Our next result shows that this condition, together with no *simple* dominance violations, is also sufficient, and thus provides a complete characterization of all data sets that are consistent with concave TC-EU.

Theorem 2. *If $X \subset \mathbb{R}$, a data set $(\preceq^0, \preceq^1, f)$ is consistent with concave TC-EU if and only if (i) time-0 and time-1 preferences are single-peaked and (ii) the data exhibit no simple dominance violations.*

We note that the additional restriction that the time-0 preference is also single-peaked implies that lack of simple dominance violations is enough to guarantee consistency with TC-EU; it is not necessary to consider cyclic dominance violations more generally in this case. In Example 3.1, the time-0 preference is not single-peaked, because single-peakness would require that $1 \preceq^0 2$ and $2 \preceq^0 3$ if $2 \prec^0 4$, which is inconsistent with $3 \prec^0 1$.

3.2.1 A Sketch of the Proofs of our Main Results

The idea behind the proofs of Theorems 1 and 2 is as follows. First, define $U^0 \subset \mathbb{R}^{|X|}$ to be the set of utilities consistent with $\preceq^0$ and $\bar{U}^1 \subset \mathbb{R}^{|X|}$ to be the set of *expected* utilities consistent with the random preference $\preceq^1$. The sets $U^0, \bar{U}^1$ have an empty intersection if and only if the data set $(\preceq^0, \preceq^1, f)$ cannot be explained by TC-EU. The sets $U^0, \bar{U}^1$ are convex, relatively open cones that are closed under the addition of constants. Thus, they have an empty intersection if and only if there exists a hyperplane that properly separates U^0 and $\bar{U}^1$; i.e., there is a vector $p \in \mathbb{R}^{|X|}$ and a constant $c \in \mathbb{R}$ such that $p \cdot u^0 \geq c \geq p \cdot u^1$ for all $u^0 \in U^0$ and $u^1 \in \bar{U}^1$, with one of the inequalities strict. Due to the structure of $U^0, \bar{U}^1$, this hyperplane must pass through the origin (i.e., $c = 0$) and the entries of p must sum to 0. Moreover, we show that that the same inequality must remain strict for *all* $u^0 \in U^0$ and $u^1 \in \bar{U}^1$. These properties imply that we can write p as the difference of two component-wise non-negative vectors whose entries sum to 1, and thus correspond to lotteries over alternatives. These lotteries have the property that one is preferred under any utility representation of the time-0 preference $\preceq^0$, while the other is preferred under any utility representation of the time-1 preference $\preceq^1$, with one of these preferences strict (we state variants of this insight as Theorems 3 and 4). The goal, then, is to establish that the existence of these lotteries, together with the assumptions of Theorems 1 or 2, implies cyclic or simple dominance violations, respectively.

For clarity, we sketch the remainder of the argument only for the case where all preferences are complete, and where there is a unique least-preferred alternative at both time 0 and in every state at time 1. Without these restrictions, the argument is substantially more involved, and makes use of supplementary results on aggregation of incomplete preorders in Appendix A.1. In the simple case, the remainder of the proof establishes that if $\bar{U}^1$ is generated by either single-peaked or concave utilities, then the existence of two such lotteries implies that one can find two *degenerate* lotteries with the same property, which then correspond to a simple dominance violation. Mathematically, this step might be of independent interest and corresponds to establishing a condition under which if a separating hyperplane exists between two cones, there also exists a hyperplane where p has only two non-zero entries (or only entries equal to $0, +1$, or -1). We prove this result by induction over the number of alternatives. It is trivially true for two alternatives.

For any number of alternatives larger than 2 we establish that the alternative that is least preferred according to the time-0 preference is such that (i) it either dominates another alternative according to the time-1 preference or (ii) its corresponding coefficient in p must equal zero. Case (i) implies a separating hyperplane with only two non-zero entries. Case (ii) reduces the problem to one with one less alternative, which then completes the proof by the induction hypothesis.

To establish the claim in case (ii) above, we show that if a given alternative x does not dominate any other alternative according to the time-1 preferences, and preferences are either single-peaked or concave, we can construct time-1 utility representations that leave the expected utility of any other alternative unchanged, and decrease the expected utility of x by an arbitrary amount. As x is chosen to be the least preferred element according to the time-0 preference, one can also arbitrarily decrease its utility in a representation of the time-0 preferences. But because the hyperplane separates U^0 and $\bar{U}^1$, the coefficient of x in the vector p must be zero.

3.3 Other Types of Ordinal Preference Data

We have thus far characterized consistency with TC-EU for data sets where the time-1 preference is single-peaked. A natural question is whether the results generalize to non-single-peaked preferences. The example below shows that they do not.

Example 3.2. There are 6 alternatives $X = \{1, 2, 3, 4, 5, 6\}$ and two states $\Omega = \{1, 2\}$. All preferences are strict and given by

$$\begin{aligned}
&1 \prec^0 2 \prec^0 3 \prec^0 4 \prec^0 5 \prec^0 6 \\
&2 \prec^1_1 3 \prec^1_1 6 \prec^1_1 1 \prec^1_1 4 \prec^1_1 5 \\
&4 \prec^1_2 1 \prec^1_2 2 \prec^1_2 5 \prec^1_2 6 \prec^1_2 3\,.
\end{aligned}$$

It is easy to check that the preferences given in Example 3.2 exhibit no simple or cyclic dominance violations. However, the preferences are not consistent with TC-EU. In both states, the time-1 preference implies that the individual strictly prefers a uniform lottery over $\{1, 3, 5\}$ to a uniform lottery over $\{2, 4, 6\}$ in each state, since the former first-order stochastically dominates the latter. However, the time-0 preference implies that the individual strictly prefers a uniform lottery over $\{2,4,6\}$ to a uniform lottery over $\{1,3,5\}$.

Example 3.2 is minimal in the following sense: For all data sets with two states ($|\Omega| = 2$) and fewer than 6 alternatives ($|X| < 6$), consistency with TC-EU is ensured when there are no simple dominance violations (i.e., the conclusion of Theorem 1 holds for non single-peaked data with $|\Omega| = 2, |X| < 6$). Furthermore, Example 3.2 is the only data set (up to relabeling of the states) with $|\Omega| = 2, |X| = 6$ where Theorem 1 does not hold.[16] But there are many other examples of data sets not consistent with TC-EU and not exhibiting simple dominance violations when $|\Omega| > 2$ or $|X| > 6$.

To formalize more general necessary and sufficient conditions for consistency with TC-EU, recall that a complete ordinal preference $\preceq$ over X induces an incomplete preference over lotteries through first-order stochastic dominance. Formally, denote by $L(x)$ the probability assigned to x by the lottery $L \in \Delta(X)$. For $x_1 \preceq x_2 \preceq \ldots \preceq x_{|X|}$ the lottery L is dominated by L' if for all $r \in \{1, \ldots, |X|\}$,

$$\sum_{s=1}^{r} L(x_s) \geq \sum_{s=1}^{r} L'(x_s)\,.$$

Strict dominance—denoted by $\prec$—holds if dominance holds, but not equality. We can also generalize the stochastic dominance order to incomplete preferences. We say that $L \preceq L'$ if and only if first-order stochastic dominance holds for all completions of the incomplete preference.

Example 3.2 motivates a stronger necessary and sufficient condition for consistency with TC-EU. The example was about a pair of lotteries $L, L' \in \Delta(X)$ such that $L \prec^0 L'$ but $L' \prec^1_\omega L$ for all states ω. This suggests that consistency with TC-EU should involve the following notion of stochastic dominance violations:

Definition 5 (Stochastic Dominance Violations)**.** A data set $(\preceq^0, \preceq^1, f)$ exhibits a stochastic dominance violation if there exist lotteries $L, L' \in \Delta(X)$ such that $L \preceq^0 L'$, $L' \preceq^1_\omega L$ for all $\omega \in \Omega$ and either $L \prec^0 L'$ or $L' \prec^1_\omega L$ for some $\omega \in \Omega$.

It follows immediately by considering degenerate lotteries that whenever there is no stochastic dominance violation that there is also no simple dominance violation. It is also easy to see that a cyclic dominance violation implies a stochastic dominance violation. Following the notation of Definition 4, let L' be a uniform lottery over the alternatives with even indices in the cycle, and let L be a uniform lottery over the alternatives with odd indices in the cycle. Then $L \preceq^0 L'$ and $L' \preceq^1_\omega L$ for all ω, with at least one of the preferences strict. Example 3.2 illustrates that the opposite direction does not hold in general. Our next theorem shows that stochastic dominance violations characterize the data sets that are incompatible with TC-EU preferences even if one does not impose that preferences are single-peaked.

[16] We verified this using a computer program that checks for each configuration of preferences if there is a simple dominance violation and solves the linear programming problems that corresponds to checking if there is an TC-EU representation.

Theorem 3 (Consistency with TC-EU)**.** *A data set* $(\preceq^0, \preceq^1, f)$ *is consistent with TC-EU if and only if it exhibits no stochastic dominance violations.*

Theorem 3 is a natural extension of our earlier two results for single-peaked and concave preferences. Although it is possible to reject TC-EU without cyclic dominance violations for certain types of non-single-peaked preferences, the theorem shows that, loosely speaking, rejecting TC-EU is not *that* much easier for other classes of preferences. Practically, the theorem is most useful for interpreting empirical designs where the alternatives are not naturally ordered along a single dimension; e.g., designs where individuals might choose between three or more different food options. Because in such designs the choice set is unlikely to be large, checking for stochastic dominance violations is not difficult, and there will generally be few cases that exhibit stochastic dominance violations despite no cyclic dominance violations.

Designs where the elements are not naturally ordered are often analyzed with convenient parametric models of stochastic discrete choice, such as the Luce model. Below, we provide a sufficient (but not necessary) condition that may be easier to check in some applications than the necessary and sufficient condition in Theorem 3—and that speaks directly to a key property of commonly-used discrete choice models. The sufficient condition extends our simple dominance consistency condition to a consistency condition over sets.

Proposition 1. *The data set* $(\preceq^0, \preceq^1, f)$ *is consistent with TC-EU if one of the following holds:*

(i) For each x *there is a* ω *such that for all* $y \succeq^0 x$ *we do not have* $x \succeq^1_\omega y$*.*

(ii) For each x *there is a* ω *such that for all* $y \preceq^0 x$ *we do not have* $x \preceq^1_\omega y$*.*

As an illustration, suppose that each alternative is the most preferred alternative with positive probability—a *positivity* condition that holds for most standard stochastic discrete choice models. Positivity holds, for example, in models with a random effect that follows a Type-1 extreme value distribution—such as the one used by Sadoff et al. (2019) to estimate the degree of dynamic inconsistency in food choice. In this case, it is immediate that all alternatives that are lower-ranked according to the time-0 preference will also be lower-ranked according to the preferences in the state where it is the most preferred alternative. Thus, the second sufficient condition of the above proposition holds, and the data set is consistent with TC-EU.

3.4 Cardinal Preferences Data

We prove Theorem 3 by establishing a more general result that applies to cases where some cardinal information is known about time-0 or time-1 preferences. For example, some cardinal information can be acquired by observing preferences over a set of lotteries, allowing the analyst to draw some conclusions about the curvature of the utility functions. Some types of cardinal information might also be assumed. For example, the analyst might use supplementary data on risk aversion, the elasticity of labor supply, or the elasticity of intertemporal substitution to decide what is a "reasonable" degree of curvature.

Formally, we associate each cardinal preference with a vector $u_\omega \in \mathbb{R}^{|X|}$. We consider a data set (U^0, U^1) where it is known that the time-1 cardinal preferences in state ω satisfy $u_\omega(\cdot) \in U^1_\omega \subseteq \mathbb{R}^{|X|}$, and we set $U^1 \subseteq \mathbb{R}^{|X|\times|\Omega|}$ to be the product of the sets U^1_ω. We make three assumptions about U^1: (i) U^1_ω is convex; (ii) U^1_ω is open relative to its affine hull; (iii) if $u_\omega \in U^1_\omega$ then $\lambda_0 + \lambda_1 u_\omega \in U^1_\omega$ for all $\lambda_0 \in \mathbb{R}$ and $\lambda_1 \in \mathbb{R}_{++}$. Assumption (iii) corresponds to the fact that cardinal information about utility functions can only be learned up to monotonic linear transformations. Assumption (ii) is an innocuous technical condition. Assumption (i) is arguably the strongest.

All three of these assumptions are satisfied for sets U^0 and U^1 of utility functions that represent the ordinal relations $\preceq^0$ and $\preceq^1$, respectively (see Lemma 7 in the Appendix). Because of this, the result in this subsection generalizes Theorem 3. The following definition generalizes what it means for the information observed by the analyst to be consistent with TC-EU.

Definition 6 (Consistency with TC-EU)**.** A data set (U^0, U^1, f) is consistent with TC-EU if there exist utility functions $u^1 \in U^1$ and $u^0 \in U^0$ such that for all $x \in X$,

$$u^0(x) = \sum_{\omega\in\Omega} f_\omega u^1_\omega(x)\,.$$

Intuitively, U^0 captures the information that the analyst has about time-0 preferences and U^1 captures the information that the analyst has about time-1 preferences. A data set (U^0, U^1, f) is consistent with TC-EU if there is a way of picking a utility function consistent with the time-1 information such that the induced time-0 expected utility function is consistent with the information the analyst has about time-0 utility.

To characterize the preferences that are consistent with TC-EU, it will be necessary to consider lotteries over the alternatives, as we have done in the previous subsection. For any lottery $L \in \Delta(X)$ and utility function $u : X \to \mathbb{R}$, we define the associated expected utility $u(L) = \sum_{x\in X} L(x)u(x)$.

Definition 7 (Dominance with Respect to U^1)**.** We say that L' weakly dominates L if $u^1_\omega(L') \geq u^1_\omega(L)$ for all $u^1 \in U^1$ and $\omega \in \Omega$. We say that L' strictly dominates L if it weakly dominates L for each $u^1 \in U^1$, and there exists ω such that $u^1_\omega(L') > u^1_\omega(L)$.

The above definition is equivalent to first-order stochastic dominance whenever U^1_ω is the set of all utility functions consistent with a given ordinal preference $\preceq^1_\omega$. We generalize our definition of stochastic dominance violations accordingly.

Definition 8. A data set (U^0, U^1, f) exhibits stochastic dominance violations if there exist lotteries $L, L' \in \Delta(X)$ such that for every $u^0 \in U^0$ either

(i) L' weakly dominates L with respect to U^1 and $u^0(L) > u^0(L')$, or
(ii) L' strictly dominates L with respect to U^1 and $u^0(L) \geq u^0(L')$.

This definition facilitates the following generalization of Theorem 3:

Theorem 4. *A data set (U^0, U^1, f) is consistent with TC-EU if and only if it exhibits no stochastic dominance violations.*

4 Estimating Time Inconsistency

4.1 Rejecting Time Inconsistency

We begin this section by pointing out that data sets consistent with TC-EU are also likely to be consistent with time-inconsistent preferences. In this subsection, we focus on *dynamic preference reversals*, which correspond to different time-0 and time-1 rankings of alternatives in X in at least one state of the world:

Definition 9. A data set $(\preceq^0, \preceq^1, f)$ is consistent with a dynamic preference reversal if there exists a utility function $u^0 : X \times \Omega \to \mathbb{R}$ such that

1. $\sum_\omega f_\omega u^0_\omega(x) \geq \sum_\omega f_\omega u^0_\omega(y)$ whenever $x \succeq^0 y$.
2. There is ω and a pair (x, y) such that $u^0_\omega(x) \geq u^0_\omega(y)$ and $x \preceq^1_\omega y$, with either $u^0_\omega(x) > u^0_\omega(y)$ or $x \prec^1_\omega y$.

In words, a data set is consistent with dynamic preference reversals if there exists a state-dependent time-0 utility such that maximization of the expectation of that utility rationalizes the time-0 preference, and at least in some state does not rationalize the time-1 preference. Clearly, whenever a data set is not consistent with TC-EU it exhibits a dynamic preference reversal. However, even if the data set is consistent with TC-EU we still cannot rule out a dynamic preference reversal when the distribution over time-1 preferences is non-degenerate.[17]

To see why, consider a utility function $u^0 : X \times \Omega$, where $u^0_\omega(x) \leq u^0_\omega(y)$ if and only if $x \preceq^0 y$. Then, the agent with these time-0 preferences exhibits a dynamic preference reversal if there exists a state ω in which $y \prec^1_\omega x$. In words, there exists a state in which the time-1 preference is not identical to the time-0 preference $\preceq^0$. But such a state exists as long as the distribution over time-1 preferences is non-degenerate. This leads to the following trivial proposition:

Proposition 2. *Any data set $(\preceq^0, \preceq^1, f)$ where there exists a state ω such that time 0 and time 1 preferences disagree, i.e. $\preceq^0 \neq \preceq^1_\omega$, is consistent with a dynamic preference reversal.*

The key lesson of this trivial proposition is that in non-deterministic environments, it is at least as difficult to reject time inconsistency as it is to reject time consistency. The proposition thus raises the natural question of whether there exist restrictions on preferences and/or the economic environment that allow for either point identification or informative set identification of time inconsistency.

[17]We focus on reversals of rankings over objects in X because the set of utility functions that are consistent with a given ordinal preference is relatively open, and thus it follows that whenever one can rationalize a data set by (u^0, u^1) such that $u^0 = \sum_\omega f_\omega u^1_\omega$ one can also rationalize this data set by a pair of utilities (u^0, u^1) such that $u^0 \neq \sum_\omega f_\omega u^1_\omega$. This conclusion is somewhat uninteresting as it merely states that one can slightly perturb the cardinal utilities without changing the ordinal ranking over alternatives.

4.2 Data Sets with Identification of Time Inconsistency

In this subsection, we propose a set of restrictions under which identification is possible. The main idea behind this identification strategy is to assume that the choice set contains one dimension Z over which preferences are known and additively separable. If Z is sufficiently rich, then the analyst can infer the strength of preferences along the other dimensions from the agent's "willingness to pay" (WTP) from Z. As a motivating example, consider the following hypothetical variation of the experiment by Read and van Leeuwen:

Example 4.1 (Read and van Leeuwen with Money)**.** Suppose again, as in Example 2.1, that the agent chooses between a healthy snack and an unhealthy snack. As before, there are two states in time 1, where the agent either feels normal ($\omega = 1$) or gorged ($\omega = 2$), with the corresponding probabilities $f_1 = 0.8$ and $f_2 = 0.2$. But suppose now that the experimenter instead elicits the maximal amount of money (to be received later at "time 2") that a person is willing to forego to receive their preferred option, both at time 0 and at time 1.[18] Given the small amounts of money involved, the agent's preferences are (approximately) quasi-linear in the monetary amounts varied in the experiment, and the marginal utility from money does not vary with the hunger state. Under these assumptions, TC-EU implies that the WTP for the healthy snack at time 0 must equal the average WTP at time 1.

Suppose that at time 0, the agent has a WTP of \$1 for the healthy snack. But suppose that at time 1, the analyst finds that the agent prefers the unhealthy snack by \$1 80 percent of the time (i.e., when the agent is feeling normal), and prefers the healthy snack by \$5 20 percent of the time (i.e., when the agent is feeling gorged). Thus, because the agent has an average WTP for the healthy snack of $0.2 \times 5 - 0.8 \times 1 = \$0.20 < \$1$ at time 1, their behavior is inconsistent with TC-EU.

We next formalize this idea for general environments, starting with two formal definitions.

Definition 10. Preferences $(\preceq^0, \preceq^1, f)$ on $X \subseteq Y \times Z$ have an *additively separable representation* (h^0, h^1, g^0, g^1) if there exist $h^0, h^1_\omega : Y \to \mathbb{R}$ and $g^0, g^1_\omega : Z \to \mathbb{R}$ such that for all ω,

$$u^0(y,z) = h^0(y) + g^0(z)$$
$$u^1_\omega(y,z) = h^1_\omega(y) + g^1_\omega(z)$$

are consistent with $\preceq^0$ and $\preceq^1_\omega$, respectively.[19]

Definition 11. A preference $\preceq$ over X is *responsive* if there exists a reference alternative y° such that for every pair $(y, y^\circ) \in Y^2$ there exists a pair $(z, z^\circ) \in Z^2$ such that $(y,z) \sim (y^\circ, z^\circ)$.

An agent's preferences in a data set $(\preceq^0, \preceq^1, f)$ are responsive if $\preceq^0$ and $\prec^1_\omega$ are responsive for all ω, with the same reference alternative y°.

[18] For the purpose of this example, assume that both time-0 and time-1 decisions concern time-2 money, to eliminate any potential issues with money discounting. As we discuss later, the assumptions of this example are most likely to be satisfied when time 2 is reasonably far away from time 1.

[19] To be clear, we use $h^1 = (h^1_\omega)_{\omega \in \Omega}$ and $g^1 = (g^1_\omega)_{\omega \in \Omega}$ to denote vectors of possible time-1 utilities.

As Example 4.1 illustrated, the assumptions of separability and responsiveness correspond to the case where the agent's preference for receiving an alternative $y \in Y$ instead of $y^\circ \in Y$ can be "priced out" in units of Z.[20] The key additional assumption needed to use the priced-out valuations to point identify time preferences is that g is state-independent; i.e., the agent's valuations of alternatives in Z are state-independent. In the context of Example 4.1, this assumption amounts to the plausible case that the agent's marginal utility of money does not vary with their appetite. The proposition below formalizes the general case.

Proposition 3. *If preferences in the data set* $(\preceq^0, \preceq^1, f)$ *are responsive, additively separable, and* $g \equiv g^0 \equiv g_1^1 \equiv \ldots \equiv g^1_{|\Omega|}$ *is known, then the following is point-identified:*

$$\frac{h^0(y) - h^0(y^\circ)}{\sum_\omega f_\omega [h^1_\omega(y) - h^1_\omega(y^\circ)]} = \frac{g(z^\circ_{0,y}) - g(z_{0,y})}{\sum_\omega f_\omega [g(z^\circ_{\omega,y}) - g(z_{\omega,y})]}\,, \tag{3}$$

where $(y, z_{0,y}) \sim^0 (y^\circ, z^\circ_{0,y})$ *and* $(y, z_{\omega,y}) \sim^1_\omega (y^\circ, z^\circ_{\omega,y})$ *for all* $\omega \in \Omega$.

We note that the right-hand-side of (3) is observable in the data set because g is known and the alternatives $z_{0,y}, z_{0,y^\circ}, z_{\omega,y}, z_{\omega,y^\circ}$ that make the agent indifferent between y and y° are observed in a responsive data set. Intuitively, the ratio on the right-hand-side captures how the agent's valuation of alternatives in Y, measured in units of Z, changes over time (on average).

For a TC-EU agent, $h^0(y) = \sum_\omega f_\omega\, h^1_\omega(y)$ for all y, and thus the expression in (3) must equal 1. Thus, under the assumptions of Proposition 3, one can exactly identify the degree of time inconsistency (or lack thereof). An identification strategy corresponding to the logic of Proposition 3 is utilized in Augenblick and Rabin (2019), Augenblick (2018), and Fedyk (2018). We summarize the main idea in the example below:

Example 4.2 (Augenblick and Rabin 2019)**.** Augenblick and Rabin elicit willingness to work for various amounts of money. Suppose that preferences are given by

$$\begin{aligned} u^0(y,z) &= \left[\sum_\omega f_\omega \beta\, h^1_\omega(y)\right] + \beta\, g(z) \\ u^1(y,z) &= \quad h^1_\omega(y) + \beta\, g(z) \end{aligned} \tag{4}$$

where y is work at time 1 and z is compensation for this work, paid out later. The analyst has no reason to believe that the marginal utility of money is related to the marginal costs of effort in the experiment. Also, given the small stakes, the analyst has good reason to believe that utility is

[20] Practically, a responsive data set can be easily generated using standard multiple price list or Becker–DeGroot–Marschak (BDM) techniques. The analyst simply needs to elicit how much money an agent is willing to forego to obtain their preferred option, and ensure enough range in monetary amounts to elicit the agent's maximum willingness to pay.

quasi-linear in money.[21] As $h^0 \equiv \sum_\omega \beta f_\omega h^1_\omega$ we have that

$$\beta = \frac{h^0(y) - h^0(y^\circ)}{\sum_\omega f_\omega [h^1_\omega(y) - h^1_\omega(y^\circ)]}$$

and Proposition 3 implies that because g is known, β is point-identified. The intuition is that quasi-hyperbolic discounting implies that estimated willingness-to-accept for an extra unit of work must be higher by a proportion $1/\beta$, on average, in time 1 versus time 0.

As the two examples illustrate, one strategy to obtain point identification is to monetize agents' preferences over alternatives in set Y, under conditions that ensure that it is plausible to assume that (i) preferences are separable over money and Y, (ii) the marginal utility of money does not vary with shocks to utility from alternatives in Y, and (iii) the analyst can estimate a utility function over money, either by assuming quasi-linearity or by using preferences over lotteries to estimate curvature. Although this set of conditions is restrictive, it is possible to plausibly approximate these conditions in the field using variants of ideas in the Augenblick and Rabin (2019) design. This has been recently done by Chaloupka et al. (2019), Carrera et al. (2021) and Allcott et al. (forthcoming), building on the ideas in DellaVigna and Malmendier (2004) and especially Acland and Levy (2015).[22]

Having a monetary domain is neither necessary nor sufficient to achieve point identification. In general, what is important is that at time 1, the agent does not update their expectation of utility over Z. This is most likely to be satisfied when (i) time 0 and time 1 are "close together," (ii) Y corresponds to consumption events realized at time 1, and (iii) Z corresponds to time-2 consumption events that are "far away" from time 1. For example, if time 1 is 1 week away from time 0, but Z corresponds to consumption or effort that is one year away, then it is unlikely that any information could be revealed between time 0 and time 1 that would alter the agent's expectation of utility from Z. This suggests strategies for point identification that don't involve money.[23] On the other hand, the conditions of Proposition 3 are unlikely to be satisfied if Y and Z correspond to large sums of money received at time 1 and time 2, with time 1 sufficiently far away from time 0

[21] Alternatively, assume that the analyst gathers additional data on preferences over monetary lotteries to estimate the curvature of g.

[22] For example, Carrera et al. (2021) develop a field analogue of Augenblick and Rabin's strategy to estimate quasi-hyperbolic discounting parameters in the exercise domain by estimating gym members' "desired" (time 0) and "realized" (time 1) attendance at varying monetary incentives for gym attendance. Like Augenblick and Rabin, Carrera et al. assume that utility over money is approximately linear, given the relatively small stakes involved. Allcott et al. (forthcoming) adapt Carrera et al.'s strategy to estimate the quasi-hyperbolic discounting parameters of payday loan borrowers. Allcott et al. (forthcoming) estimate curvature in utility over money by eliciting borrowers' preferences over monetary lotteries. See also Allcott et al. (2021).

[23] For a concrete example, consider a modification of the Augenblick et al. (2015) design, as stylized by Example 2.2, but suppose now that time 2 corresponds to work that must be completed in one year. In this modified design, it is plausible to assume that the agent does not update their beliefs about the time-2 effort costs between time 0 and time 1, and thus that the expected cost of time-2 effort is constant across all states of the world. Thus, if the analyst is justified in assuming separable effort costs over time, so that choices from multiple budget sets identify time-2 effort costs, then Proposition 3 implies that β is identified. Thus, the convex time budget approach can be used to identify time preferences as long as the time-2 consumption/effort events are sufficiently far in time that the agent does not update their beliefs about time-2 utility between time 0 and time 1.

such that an agent subject to serially-correlated liquidity and income shocks may plausibly update their beliefs at time 1 about their time-2 marginal utility of money.

4.3 Partial Identification of Quasi-hyperbolic Discounting

We end this section by noting that it is difficult to obtain point identification if the preference along one dimension is not independent of the state. For expositional clarity, we focus on the quasi-hyperbolic discounting model with multiplicative taste shocks:

Definition 12 (Quasi-hyperbolic Discounting with Multiplicative Shocks)**.** A data set $(\preceq^0, \preceq^1, f)$ is consistent with quasi-hyperbolic discounting with multiplicative taste shocks if it is consistent with utilities of the form

$$\begin{aligned} u^0(y,z) &= \sum_{\omega\in\Omega} f_\omega \left(\theta^1_\omega h(y) + \theta^2_\omega g(z)\right) \\ u^1(y,z) &= \theta^1_\omega\, h(y) + \beta\, \theta^2_\omega\, g(z)\,. \end{aligned} \tag{5}$$

Note that if preferences can be represented as in Definition 12, then there exists an additively separable representation (h^0, h^1, g^0, g^1) of the preferences where the function g^1 does not depend on the state and equals g^0. To obtain this representation, simply set $h^0 \equiv \frac{\sum_\omega f_\omega \theta^1_\omega}{\sum_\omega f_\omega \theta^2_\omega} h$, $h^1_\omega \equiv \frac{\theta^1_\omega}{\beta\theta^2_\omega} h$, $g^0 = g$ and $g^1 = \beta g$.

To obtain intuition for the types of inferences that can be made about the parameter β given data consistent with quasi-hyperbolic discounting with multiplicative taste shocks, consider an identification strategy that (wrongly) assumes no taste shocks and assumes instead that all differences in time 1 are instead due to variation in the time preference parameter β. That is, if at time 0 the agent is indifferent between $(y, z_{0,y}) \sim^0 (y^\circ, z^\circ_{0,y})$ and at time 1 is indifferent between $(y, z_{\omega,y}) \sim^1_\omega (y^\circ, z^\circ_{\omega,y})$, then

$$\begin{aligned} h(y) + \quad g(z_{0,y}) &= h(y^\circ) + \quad g(z^\circ_{0,y}) \\ h(y) + \hat\beta_\omega g(z_{\omega,y}) &= h(y^\circ) + \hat\beta_\omega g(z^\circ_{\omega,y})\,. \end{aligned}$$

Rearranging these equations yields that

$$\hat\beta_\omega = \frac{g(z^\circ_{0,y}) - g(z_{0,y})}{g(z^\circ_{\omega,y}) - g(z_{\omega,y})}\,. \tag{6}$$

We use the "hat" notation in the definition above because $\hat\beta_\omega$ can also be thought of as a "noisy" estimate of the true present focus parameter β, which is a common statistic to report in empirical studies. An immediate corollary of Proposition 3 is that if utility over the Z dimension is state-independent, then β is point-identified and given by

$$\beta = \frac{1}{\sum_{\omega\in\Omega} f_\omega \hat\beta_\omega^{-1}}\,. \tag{7}$$

Our next result shows that without state independence of utility from Z, the range of the distribution of $\hat{\beta}_\omega$ identifies the range of possible values of β consistent with the data.

Proposition 4. *A responsive data set $(\preceq^0, \preceq^1, f)$ is consistent with quasi-hyperbolic discounting with multiplicative taste shocks if and only if*

(i) it has an additively separable representation (h^0, h^1, g^0, g^1) where for all $\omega \in \Omega$, $g^0 = g^1_\omega = g$ and $\frac{h^1_\omega(y)}{h^0(y)}$ is non-negative and constant in $y \in Y$, and

(ii) $\beta \in \left(\min_\omega \hat{\beta}_\omega, \max_\omega \hat{\beta}_\omega\right)$ for $\hat{\beta}_\omega = \frac{g(z^\circ_{0,y}) - g(z_{0,y})}{g(z^\circ_{\omega,y}) - g(z_{\omega,y})}$.

The range of $\hat{\beta}_\omega$ is estimable, and is frequently reported, which facilitates the application of Proposition 4. In practice, because many experiments feature only one revision observation per person, $\hat{\beta}_\omega$ is obtained as an individual-level estimate of present focus—we discuss this in greater detail in Section 5.3 and especially Supplementary Appendix B. To illustrate what distributions of $\hat{\beta}$ look like in practice, the left panel of Figure 1 presents estimates $\hat{\beta}$ reported in Augenblick et al. (2015), while the right panel presents estimates reported in Augenblick and Rabin (2019). Under the assumptions of Proposition 3, the population (average) present focus is obtained from the average of these point estimates (see Supplementary Appendix B.2 for more detail). As we have argued, the assumptions of Proposition 3 are plausible for Augenblick and Rabin (2019), but less so for Augenblick et al. (2015). In Augenblick et al. (2015), it is plausible that there are random taste shocks affecting utility from both Y and Z. Under the assumptions of Proposition 4, the population β in Augenblick et al. (2015) can be any value between the minimum and maximum of the support of the estimated $\hat{\beta}$. As we show in Supplementary Appendix B.2, this implies the range $[0.66, 1.12]$, even after 20 percent of the most extreme observations of the estimated distribution of $\hat{\beta}$ are excluded. This range includes values of population β that imply both future focus and significant present focus.

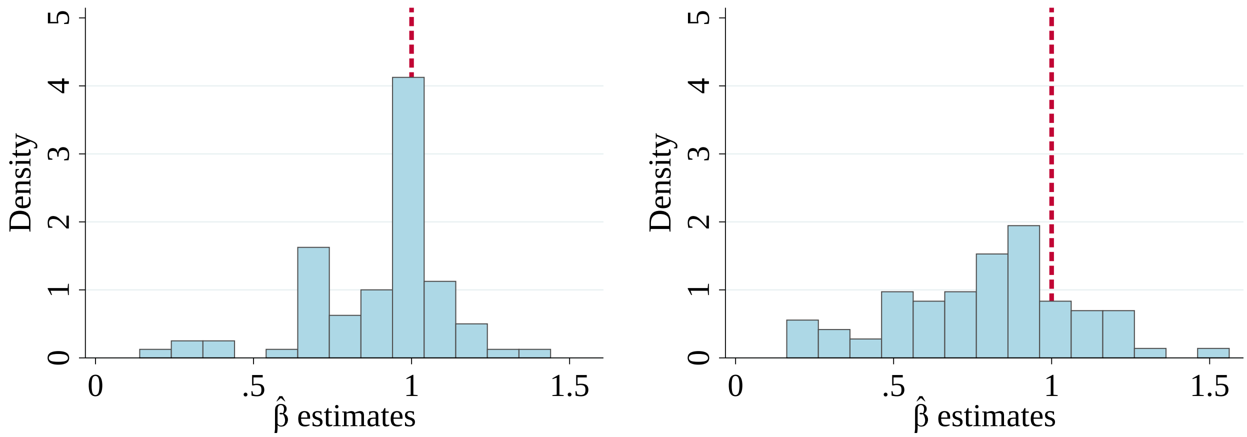
5 Discussion

5.1 Are Taste Shocks Important?

There is one special assumption that an analyst could in principle make to obtain stronger identification results that would typically lead to rejection of TC-EU: that no individuals face any uncertainty between time 0 and time 1, and thus that all differences in time-1 choices between individuals making the same time-0 choice are due solely to individual differences in time preferences. To illustrate this assumption, return to the consumption over time Example 2.2. Then, any individual i who in time 1 revises their time-1 consumption to be higher than the time-0 choice would be classified as present-focused ($\beta_i < 1$), while any individual who revises their time-1 consumption to be lower than the time-0 choice would be classified as future-focused ($\beta_i > 1$).

Both intuition and evidence suggest that this assumption is unrealistic. For example, Read and van Leeuwen (1998) experimentally show that variation in individuals' satiation has a large effect

Figure 1: Distribution of estimates $\hat{\beta}$ from two studies



Notes: The left panel presents the histogram of of individual-level estimates $\hat{\beta}$ using the first experimental block in Augenblick et al. (2015), as reported in the top panel of their Figure 6. The right panel presents the histogram of individual-level estimates $\hat{\beta}$ in Augenblick and Rabin (2019), as reported in the left panel of their Figure 6. The formatting of the histogram from Augenblick et al. (2015) is slightly altered to match the formatting choice of the histogram in Augenblick and Rabin (2019).

on their preferences for healthy versus unhealthy foods at time 1.

The strong assumption of no random taste shocks also implies that because any difference between time-0 and time-1 choice is due to (stable) time preferences, these differences should be stable within individuals over time. To test this assumption, the analyst needs to estimate individuals' time-0 and time-1 preferences on two (or more) separate occasions, $j \in \{1, 2\}$. If on both occasions the analyst obtains an identical measure of time preference for each individual (e.g., $\hat{\beta}_i^1 = \hat{\beta}_i^2$ for all agents i in the quasi-hyperbolic discounting model), then the assumption is satisfied. For example, Sadoff et al. (2019) include two sets of decisions to study the stability of their measure of time inconsistency. Sadoff et al. find correlations between 0.2 and 0.33 in agents' propensity to switch toward unhealthier foods at time 1, which suggests some stable individual differences. However, this also implies that stable individual differences explain only $0.2^2 = 0.04$ to $0.33^2 = 0.11$ of the variance in time-1 revisions of time-0 choices, leaving random taste shocks and other forms of updating to explain the remaining variation. Similarly, as we show in Supplementary Appendix B.1, stable individual differences in time preferences can explain at most 2 percent of the variation in time-1 revisions of time-0 preferences in the Augenblick et al. (2015) data.

5.2 Commitment

Our results have some potentially interesting connections to demand for *commitment devices* in the form of choice set restrictions. In particular, rejection of TC-EU in our setting implies that *sophisticated agents*—i.e., those who are aware of their time inconsistency at time 0—should demand commitment devices in the form of choice set restrictions. For example, suppose that there is a simple dominance violation, such that in time 1, the agent always chooses x over x', while in time 0 the agent chooses x' over x. Then, this agent would be better off with the choice set $\{x'\}$ than with choice set $\{x, x'\}$. If this agent accurately predicts their time-1 choices, then they should strictly prefer to face $\{x'\}$ rather than $\{x, x'\}$ at time 1. Stochastic dominance violations similarly imply a demand for restrictions of choice sets that involve lotteries.

Although demand for choice set restrictions cannot be used to point identify time preference parameters (see, e.g., Carrera et al., 2021),[24] in some cases it serves as a useful correlate of time preferences that can help refine their set identification. For example, Augenblick et al. (2015) find that their measure of demand for commitment relates negatively to their estimates of the present focus parameter β, and thus provides evidence of some stable individual differences in time preferences. In Supplementary Appendix B.2 we illustrate how this finding can be used in conjunction with Proposition 4 to refine the set of time preferences that could be consistent with individuals' full set of choices.

5.3 Heterogeneity

Our results concern a data set in which the analyst observes an agent's time-0 preference and the *full* distribution of time-1 preferences. In practice, such data sets are rarely available, which implies that our results are a "best case" for identification. A more typical data set consists of a large population of individuals who each make a single choice at time 0 and at time 1.

One approach to analyzing such data sets is to assume that all individuals who make the same time-0 choice are homogeneous both in preferences and the economic environment, and consequently treat variation in time-1 decisions as due to realization of uncertainty. Under this assumption, the set of all individuals who make the same time-0 choice constitutes the kinds of data sets that we study in this paper.

Even without this homogeneity assumption, however, most of our key results still apply. Theorems 1-4 continue to hold verbatim when it's possible to observe the joint distribution of time-0 and time-1 preferences. Because the only pattern that rejects time consistency with homogeneous preferences is a (stochastic) dominance violation, and because such a violation cannot be explained by heterogeneous preferences, a data set is consistent with TC-EU under homogeneous preferences if and only if it is consistent with TC-EU under heterogeneous preferences. To see this, fix a time-0 preference profile, allowing for individuals to be heterogeneous conditional on this time-0 preference. If the analyst observes a dominance violation (simple for Theorems 1 and 2, or stochastic for Theorems 3 and 4), then the analyst must conclude that all agents with that time-0 preference are time-inconsistent. If the analyst does not observe a dominance violation, then Theorems 1-4 imply that the data can be rationalized with a homogeneous time-consistent EU agent.

Our identification results in Propositions 3 and 4 are also straightforward to generalize to give

[24] Commitment take-up is a coarse measure that might lead to false negatives in tests of time inconsistency because uncertainty and thus demand for flexibility reduce demand for choice set restrictions (Heidhues and Kőszegi, 2009, Laibson, 2015, Carrera et al., 2021), and can also deliver false positives because of noise in take-up decisions (Carrera et al., 2021).

a measure of *average* time inconsistency when there is some unobserved heterogeneity. To give a concrete example of applying Proposition 3 to a heterogeneous population, suppose that there there is a finite number of agent types making the same time-0 choice, with agents of each type having the same preferences and receiving independent draws from the same distribution of taste shocks. Suppose the analyst observes only a single realization of the time-1 preference for each agent. Then, the logic of Example 4.1 is still identical (under the maintained assumptions of that example): In expectation, the average time-1 WTP for the healthy over the unhealthy snack must equal the average time-0 WTP for the unhealthy over healthy snack. In Supplementary Appendix B.2 we provide additional examples applying Propositions 3 and 4 to heterogeneous populations, with applications to the data sets from Augenblick et al. (2015) and Augenblick and Rabin (2019).

5.4 Other Related Literature

Mathematically, the results in Section 3 also have implications for two separate literatures: social choice and identification in random utility models.

Social Choice. The ordinal efficiency welfare theorem (McLennan, 2002, Carroll, 2010) states that for any lottery that is Pareto efficient given a vector of ordinal preferences, there exist utility functions consistent with the ordinal preferences such that this lottery maximizes the sum of utilities. This result is mathematically equivalent to the special case of Theorem 4 where the analyst only observes the most preferred time-0 alternative.[25] The sharper and more interesting characterizations that we provide for single-peaked and concave preferences in Theorems 1 and 2 do not, to our knowledge, relate to any known results in the social choice literature—although they of course have implications for that literature. For example, they imply that for complete single-peaked preferences it is not necessary to consider lotteries: an alternative is a maximand of some social welfare function as long as it is not Pareto dominated by any other alternative. Example 3.2 shows that this stronger conclusion fails for social choice problems without the single-peaked property.[26]

Random Utility Models. If we set $v = \sum_\omega f_\omega u^1_\omega$ and $\varepsilon_\omega = u^1_\omega - v$, so that $v+\varepsilon$ is a random utility representation of time-1 preferences with $\mathbb{E}[\varepsilon] = 0$, then our results characterize the set of utility functions v that are consistent with the time-1 choice data. This connects to a large literature on identification of stochastic discrete choice models, though to our knowledge the data set we consider is markedly different from those studied in the literature in two ways. First, the literature considers a random choice rule ρ that assigns a probability $\rho(x, A)$ to selecting alternative x from a subset $A \subset X$, but the joint distribution of preferred alternatives over all subsets is not known. In our case, allowing the analyst to observe the full ranking in each state is equivalent to allowing the

[25] Specifically, this is the case for the more general version stated by Carroll (2010). The original version stated by McLennan (2002) imposes a more special structure.

[26] It is perhaps also worth clarifying that to our knowledge and understanding, our results do not have a mathematical connection to the literature on aggregation of time preferences (e.g., Jackson and Yariv, 2015, Millner, 2020).

analyst to observe a joint distribution of preferred alternatives from all subsets $A \in \mathcal{P}(X)$. Second, the literature typically makes the positivity assumption that each element is the most preferred one with positive probability. This assumption is economically restrictive in many environments, as it rules out most forms of time-1 preferences that are single-peaked or concave. Furthermore, a corollary of our Proposition 1 implies that this assumption is highly consequential, as it implies that the average utility v cannot be identified without imposing additional structure on ε.[27] This generalizes the insight from Alós-Ferrer et al. (2021) who highlight a related identification issue in a setting where the analyst has less information and only observes the marginal distribution of preferences over binary choice sets. They propose to resolve it by inferring cardinal information from response times, which is similar to the additional choice dimension we propose in Section 4.3.

"Static" preference reversals and money discounting. While the focus of this paper is on dynamic preference reversals, there is also an important literature on static preference reversals. A large literature studies how people trade off between the size of a monetary reward and its delay, under the assumption that time-dated payments and time-dated utils are interchangeable (see, e.g., Cohen et al. 2020 for a review). A smaller literature conducts such studies with real consumption events, such as juice squirts (McClure et al., 2007). These studies are an important and complementary source of evidence, although some of the results concerning monetary rewards could be due to violations of stationarity rather than time consistency (Halevy, 2015), or due to uncertainty changing with delay (Halevy, 2005, 2008, Andreoni and Sprenger, 2012, Chakraborty et al., 2020).

While there is an important ongoing conversation about whether time-dated monetary rewards should be treated interchangeably with time-dated utils (see, e.g., Ericson and Laibson 2019 and Cohen et al. 2020 for reviews), our results suggest one possible additional reason why studies such as those of Augenblick et al. (2015) estimate fewer preference reversals for money versus actual consumption: utility over money may be subject to fewer random taste shocks.[28]

Other work on identification of behavioral biases. Finally, our work is thematically related to other theoretical work showing that identification of behavioral biases may be more difficult than previously intuited, such as Heidhues and Strack (forthcoming) on time preferences in optimal stopping problems and Benoit and Dubra (2011) on overconfidence. While Heidhues and Strack (forthcoming) also study time preferences, the optimal stopping environment they consider

[27]There is also a thematic, but not mathematical connection to identifying time preferences in dynamic discrete choice models. See, e.g., Magnac and Thesmar (2002), Abbring and Daljord (2020), Levy and Schiraldi (2021), Mahajan et al. (2020).

[28]There are a number of other papers (see Imai et al. 2021 for a review) that follow Andreoni and Sprenger (2012) in applying the convex time budget (CTB) approach to time-dated monetary rewards. Our paper largely focuses its discussion and examples on revision designs with real consumption events because of the ambiguities highlighted by the ongoing conversation about interpreting preferences over time-dated monetary rewards. But given an interpretation, our results apply to those designs as well. For example, if the monetary rewards are treated as real consumption events, but the subjects are not plausibly subject to liquidity shocks that change their marginal utility of money, then our results imply that monetary CTBs constitute deterministic environments that allow point identification of time preferences.

is completely different from our repeated choice environment. The mathematical techniques and the economic intuition for why identification is difficult differ substantially across these two papers.

6 Conclusion

Our general characterization of consistency with TC-EU shows that in typical revision designs it is difficult to identify the degree of time inconsistency—or to even formally reject the hypothesis of time consistency. The difficulty arises from random taste shocks or other arrival of information, which are particularly plausible in more complex and economically consequential field settings. However, we have also provided guidance on the types of economic environments where the assumptions required for point identification are plausible. Thus, while identification of time inconsistency may be more difficult than initially intuited, it is certainly theoretically and empirically feasible.

Of course, our results do not imply that nothing can be learned from data sets where we show that it is not possible to formally reject TC-EU. For example, it is unlikely to be mere coincidence or file-drawer bias that in most circulated papers, the systematic reversals tend to be toward more immediately gratifying options.[29] Just as proper Bayesian scientists reservedly update about causal relationships from all well-measured associations—even when the associations are not produced by experimental or quasi-experimental techniques—we think it is appropriate to carefully update from all revision design data. Moreover, correlational analyses that link choice revisions to supplementary proxies of time inconsistency or observable determinants of taste shocks (see, e.g., Augenblick et al., 2015, Sadoff et al., 2019) can bolster the updating. At the same time, by formally studying identification in a general theoretical framework, this paper clarifies just how strong the assumptions for (point) identification of time inconsistency have to be, and helps identify the most theoretically-robust designs. We hope that this will help further the important agenda of measuring time inconsistency.

[29]Based on their meta-analysis, Imai et al. (2021) suggest that selective reporting is modest in revision designs studying effort allocation tasks using the convex time budget approach.

A Appendix

The appendix proceeds as follows: Section A.1 derives several results on the aggregation of incomplete preorders. Section A.2 derives several results on the separation of finite-dimensional cones generated by sets of utility functions. In Section A.3 we use these results to derive the results in the body of the paper.

A.1 Results on Aggregating Incomplete Preorders

Recall that we defined a new preorder $\preceq^1_*$ that ranks one alternative y weakly higher than an alternative x if the agent ranks that alternative higher in all states ω

$$x \preceq^1 y \Leftrightarrow x \preceq^1_\omega y \text{ for all } \omega \in \Omega\,.$$

We thus denote by $\preceq^1_*$ the preorder that is generated by agreement of the preorders $\preceq^1_\omega$ in the different states ω. We define another binary relation $\trianglelefteq^*$ such that y is weakly preferred to x if it is either preferred according to the time-0 preference or according to all time-1 preferences

$$x \trianglelefteq^* y \quad \Leftrightarrow \quad x \preceq^0 y \text{ or } x \preceq^1_* y\,. \tag{8}$$

We define $\triangleleft^*$ to be the asymmetric component of $\trianglelefteq^*$. We note that $\trianglelefteq^*$ need not to be transitive and define $\trianglelefteq$ to be the smallest transitive closure of $\trianglelefteq^*$. We define $\sim_\trianglelefteq$ to be the symmetric part of $\trianglelefteq$, and $\triangleleft$ to be the asymmetric part of $\trianglelefteq$.

Lemma 1. *If the data set admits no simple dominance violations then*

$$x \triangleleft^* y \quad \Leftrightarrow \quad x \prec^0 y \text{ or } x \prec^1_* y$$

Proof. We first note that $x \triangleleft^* y$ if $x \trianglelefteq^* y$ and neither $y \preceq^1_* x$ nor $y \preceq^0 x$. We furthermore note that if there is no simple dominance violation, then $x \prec^0 y$ implies that we do not have $y \preceq^1_* x$, and $x \prec^1_* y$ implies that we do not have $y \preceq^0 x$. Hence, $x \prec^0 y$ or $x \prec^1_* y$ implies $x \triangleleft^* y$.

To see that the converse direction also holds note that $x \triangleleft^* y$ implies that neither $y \preceq^0 x$ nor $y \preceq^1_* x$, and either $x \preceq^0 y$ or $x \preceq^1_* y$, which together implies that either $x \prec^0 y$ or $x \prec^1_* y$. □

We next translate the condition of no cyclic dominance violations in the data set $(\preceq^0, \preceq^1, f)$ into a condition on the induced order $\trianglelefteq^*$.

Definition 13 (Only Weak Cycles)**.** We say that $\trianglelefteq^*$ admits *only weak cycles* if $x_1 \trianglelefteq^* \cdots \trianglelefteq^* x_n \trianglelefteq^* x_1$ implies that $x_1 \sim_{\trianglelefteq^*} \cdots \sim_{\trianglelefteq^*} x_n$.

Lemma 2. *The following are equivalent:*

(i) The data set has no cyclic dominance violation.

(ii) The data-set has no simple dominance violation and $\trianglelefteq^*$ satisfies the only weak cycles condition.

Proof. $(ii) \Rightarrow (i)$: Suppose that the elements $x_1, \ldots, x_k$ constitute a cyclic dominance violation. Without loss we can assume that the first inequality is strict (otherwise reorder the elements) and according to the time-0 order (the argument for the time-1 order is identical)

$$x_1 \prec^0 x_2 \preceq^1_* x_3 \preceq^0 \ldots \preceq^0 x_k \preceq^1_* x_1 \, .$$

If there is no simple dominance violation, then Lemma 1 implies that

$$x_1 \lhd^* x_2 \unlhd^* x_3 \unlhd^* \ldots \unlhd^* x_k \unlhd^* x_1 \, ,$$

and thus a violation of the only weak cycles condition.

$(i) \Rightarrow (ii)$: Suppose (ii) does not hold. If the data set has a simple dominance violation it also has a cyclic dominance violation, as any simple dominance violation is a cyclic dominance violation with a cycle of length 2. Thus, suppose that there is no simple dominance violation but that there is a non-weak cycle in $\unlhd^*$ involving $x_1, \ldots x_k$. Without loss, assume that $x_k \lhd^* x_1$. Then Lemma 1 implies that there exist alternatives $x_1, \ldots, x_k$ such that:

(i) For all $j \leq k-1$, either $x_j \preceq^0 x_{j+1}$ or $x_j \preceq^1_* x_{j+1}$

(ii) Either $x_k \prec^0 x_1$ or $x_k \prec^1_* x_1$

Now if $x_j \preceq^0 x_{j+1} \preceq^0 x_{j+2}$, then $x_j \preceq^0 x_{j+2}$, and thus there is a non-weak cycle over the set $\{x_1, \ldots, x_k\} \setminus \{x_{j+1}\}$. A similar statement applies to three adjacent alternatives in a cycle related by $\preceq^1_*$. Thus, any non-weak cycle can be reduced to a non-weak cycle where no three adjacent alternatives are in increasing order according to $\preceq^0$ or according to $\preceq^1_*$; this non-weak cycle amounts to a cyclic dominance violation. □

Lemma 3. *If $\preceq^0$ is complete and the data-set $(\preceq^0, \preceq^1, f)$ exhibits cyclic dominance violations then there exists a simple dominance violation.*

Proof. We prove this result by contraposition. Assume that there is no simple dominance violation. Then $\unlhd^*$ preserves the asymmetric part of $\preceq^0$ by Lemma 1. As $\preceq^0$ is by assumption complete, $\unlhd^*$ must be complete as well. As $\preceq^0$ is transitive, $\unlhd^*$ can only admit weak cycles. But then Lemma 2 implies that there is no cyclic dominance violation. □

Lemma 4. *Suppose that $\unlhd^*$ satisfies the only weak cycles condition. Then $x \lhd^* y \Rightarrow x \lhd y$.*

Proof. Suppose that $x \lhd^* y$. Then the only way for $y \unlhd x$ is if there is a cycle in $\unlhd^*$ involving x and y. Since $x \lhd^* y$, this cycle is not a weak cycle. □

Lemma 5. *Suppose that $\preceq^0$ and $\preceq^1$ are single-peaked, and that there are no simple dominance violations. Then there exists a set $X^* \subseteq \{\min X, \max X\}$ such that $x \unlhd x^*$ does not hold for any $x \in X \setminus X^*$ and $x^* \in X^*$, and such that $\min X \sim_{\unlhd} \max X$ if $X^* = \{\min X, \max X\}$.*

Proof. Consider the set of alternatives that is not strictly better than any other alternative according to the preorder $\trianglelefteq$

$$\mathcal{X} = \{x \in X\colon \text{ there does not exist } y \in X \text{ with } y \triangleleft x\}.$$

We first argue that $\mathcal{X} \cap \{\min X, \max X\} \neq \emptyset$. If not, then there would exist $y \in X$ such that either $\min X \succ^0 y$ or $\min X \succ^1_* y$. Suppose first that $\min X \succ^0 y$. Single-peakness implies that if $y \neq \max X$ then $\min X \succ^0 y \succeq^0 \max X$, so that $\min X \succ^0 \max X$. Thus, $\min X \succ^0 \max X$ if $\min X \notin \mathcal{X}$. Now if $\max X \notin \mathcal{X}$, then a similar argument shows that $\min X \prec^1_* \max X$ (since $\min X \prec^0 \max X$ is impossible by transitivity), which implies a simple dominance violation. Thus, if $\min X \succ^0 y$, then there is a simple dominance violation. A symmetric argument also shows that $\min X \succ^1_* y$ implies a simple dominance violation.

Next, pick a maximal subset of $\mathcal{X}$, X^*, such that $X^* \cap \{\min X, \max X\} \neq \emptyset$, and such that $x \sim_{\trianglelefteq} x'$ for all $x, x' \in X^*$. Without loss, assume that $\min X \in X^*$. We now argue that $X^* \subseteq \{\min X, \max X\}$. If not, then there is a $y \in X \setminus \{\min X, \max X\}$ such that $y \sim_{\trianglelefteq^*} \min X$. Then, because we assumed no simple dominance violations, either

(i) there is some $y \in X \setminus \{\min X, \max X\}$ such that either $y \sim^0 \min X$ or $y \sim^1_* \min X$, or

(ii) $\min X$ and y are part of a cycle in $\trianglelefteq^*$ that is not a weak cycle

To see why (ii) must hold if (i) does not, note that if $\min X$ were part of a *weak* cycle that relates it to y through $\sim_{\trianglelefteq^*}$, and condition (i) did not hold for *any* $y' \in X \setminus \{\min X, \max X\}$, then there would need to be simple dominance violations to generate the indifferences in the weak cycle.

We next argue that in both cases (i) and (ii) above, either $\min X \succ^0 \max X$ or $\min X \succ^1_* \max X$. We then show that neither possibility is inconsistent with there being no simple dominance violations.

Case (i) If $y \sim^0 \min X$ then single-peakness implies that $y \succ^0 \max X$ and thus $\min X \succ^0 \max X$. An identical argument shows that $y \sim^1_* \min X$ implies that $\min X \succ^1_* \max X$.

Case (ii) If $\min X$ is part of a cycle in $\trianglelefteq^*$ that is not a weak cycle, then there is some $y \in X \setminus \{\min X, \max X\}$ such that either (i) $y \sim^0 \min X$ or (ii) $y \sim^1_* \min X$ or (iii) $y \prec^0 \min X$ or (iv) $y \prec^1_* \min X$. In the first two cases, we have already shown that $\min X \succ^0 \max X$ or $\min X \succ^1_* \max X$. In the second two cases, the first paragraph of this proof shows that either $\min X \succ^0 \max X$ or $\min X \succ^1_* \max X$.

Thus, if $X^* \subseteq \{\min X, \max X\}$ does not hold, then either $\min X \succ^0 \max X$ or $\min X \succ^1_* \max X$, and thus $\max X \trianglelefteq \min X$. But since $\min X \in X^*$, the definition of X^* requires that $\min X \sim_{\trianglelefteq} \max X$, and thus that $\max X \in X^*$. Moreover, since either $\min X \succ^0 \max X$ or $\min X \succ^1_* \max X$, the only way for $\min X \sim_{\trianglelefteq} \max X$ to hold in the absence of simple dominance violations is if there is a non-weak cycle involving both $\min X$ and $\max X$. But then, reasoning identical to Case (ii) above implies that we must have $\max X \succ^0 \min X$ or $\max X \succ^1_* \min X$, which implies a simple dominance violation. $\square$

A.2 Results on the Cones Generated from Sets of Utility Functions

As there are only finitely many alternatives we will throughout identify utilities with vectors in $\mathbb{R}^{|X|}$.

For any sets A and B, we define the Minkowsky sum $A + B := \{a + b \mid a \in A, b \in B\}$. We say that a set $A \subset \mathbb{R}^n$ is *constant shift invariant* if $a \in A$ implies that $a + (\lambda, \dots, \lambda) \in A$ for all $\lambda \in \mathbb{R}$.

Lemma 6 (Cone Separation Lemma)**.** *Suppose that $A, B \subset \mathbb{R}^n$ are convex cones that are open relative to their affine hull and constant shift invariant with $A \cap B = \emptyset$. Then there exists a vector $p \in \mathbb{R}^n$ with $p \neq 0$ and $\sum_{i=1}^n p_i = 0$, such that for all $a \in A, b \in B$*

$$p \cdot a \geq 0 \geq p \cdot b$$

and one of the inequalities is strict for all $a \in A$ and $b \in B$.

Proof. As A, B are disjoint they can be properly separated by a hyperplane; i.e., there exists $p \in \mathbb{R}^n$ with $p \neq 0$ and $c \in \mathbb{R}$ such that for all $a \in A, b \in B$

$$p \cdot a \geq c \geq p \cdot b\,.$$

with at least one inequality strict. As A is constant shift invariant, $a + (\lambda, \dots, \lambda) \in A$ if $a \in A$, which implies that for all $\lambda \in \mathbb{R}$

$$p \cdot a + \lambda \sum_{i=1}^n p_i \geq c\,.$$

The above inequality can only hold for all $\lambda \in \mathbb{R}$ if $\sum_{i=1}^n p_i = 0$, which thus must hold.

Similarly, as A is a cone $a \in A$ implies that $\lambda a \in A$ for all $\lambda > 0$ and hence

$$\lambda(p \cdot a) \geq c\,.$$

Taking the limit $\lambda \to 0$ yields that $0 \geq c$. Applying the same argument using that B is a cone yields that $c \geq 0$ and hence we have that $c = 0$.

By the proper separation there exists either $a \in A$ such that $p \cdot a \neq 0$ or $b \in B$ such that $p \cdot b \neq 0$. Consider the first case and assume that $a_1 \in A$ with $p \cdot a_1 > 0$. If no $a_2 \in A$ exists with $p \cdot a_2 = 0$ we have established that $p \cdot a > 0$ for all $a \in A$ and thus completed the proof. If $a_2 \in A$ exists with $p \cdot a_2 = 0$ consider another point $a_3 = a_2 + \epsilon(a_2 - a_1)$. As A is open relative to its affine hull, $a_3 \in A$ for ϵ small enough. However, note that

$$p \cdot a_3 = (1 + \epsilon)(p \cdot a_2) - \epsilon(p \cdot a_1) = -\epsilon(p \cdot a_1) < 0\,.$$

This contradicts that $p \cdot a \geq 0$ for all $a \in A$ and thus implies that no $a_2 \in A$ with $p \cdot a_2 = 0$ can exist. Hence, we have established that $p \cdot a > 0$ for all $a \in A$.

The proof for the case where there exists a $b \in B$ with $p \cdot b < 0$ is analogous. This implies that one of the inequalities is always strict for all $a \in A$ and $b \in B$. This completes the proof. □

Lemma 7. *The set of utility functions consistent with a given (potentially incomplete) preference relation is open relative to its affine hull.*

Proof. Fix two utility functions u, v consistent with $\preceq$ and fix $\epsilon > 0$ such that

$$\epsilon < \frac{\min_{y,y' \,:\, u(y)>u(y')} u(y) - u(y')}{\max_{y,y' \,:\, v(y)>v(y')} v(y) - v(y')} .$$

We have that $u(x) > u(x')$ and $v(x) > v(x')$ implies that

$$[u(x) + \epsilon(u(x) - v(x))] - [u(x') + \epsilon(u(x') - v(x'))] \geq [u(x) - u(x')] - \epsilon[v(x) - v(x')] > 0 .$$

Similarly, $u(x) = u(x')$ and $v(x) = v(x')$ implies that

$$u(x) + \epsilon(u(x) - v(x)) = u(x') + \epsilon(u(x') - v(x')) .$$

Hence, the utility $u(x) + \epsilon(u(x) - v(x))$ is also consistent with the preference $\preceq$ for every ϵ small enough and the set of utilities consistent with $\preceq$ is open relative to its affine hull (c.f. Aliprantis and Border, 2006, page 277). $\square$

Lemma 8. *The set of strictly concave utility functions is open.*

Proof. A utility function u is strictly concave if for all $x, y, z \in X$

$$\frac{y-z}{y-z}u(x) + \frac{z-x}{y-z}u(y) < u(z) .$$

Note that if $v \in B_\epsilon(u)$ (i.e., an ϵ ball around u) we have that

$$\left[\frac{y-z}{y-z}v(x) + \frac{z-x}{y-z}v(y)\right] - v(z) \leq \left[\frac{y-z}{y-z}u(x) + \frac{z-x}{y-z}u(y)\right] - u(z) + 2\epsilon .$$

Thus we have that the utility v is strictly concave for

$$\epsilon < \frac{1}{2} \min_{x,y,z \in X} \left| \left[\frac{y-z}{y-z}u(x) + \frac{z-x}{y-z}u(y)\right] - u(z) \right| . \qquad \square$$

We also make use of the following straightforward properties of Minkowski sums of sets of utilities in our proofs. Define $\bar{U}^1 = \sum_\omega U^1_\omega$ and note that $\bar{U}^1$ is convex and open relatively to its affine hull if all $(U^1_\omega)_\omega$ are convex and open relative to their affine hull. Note also that if each of the U^1_ω are cones and constant shift invariant, then so is $\bar{U}^1$: if $\bar{u} \in \bar{U}$, then $(\lambda_0, \cdots, \lambda_0) + \lambda_1 \bar{u} \in \bar{U}$, for any $\lambda_1 > 0$ and any real λ_0. Finally, note that $v = \sum_\omega f_\omega u_\omega$ for some choices of $u_\omega \in U^1_\omega$ if and only if $v = \sum_\omega u_\omega$ for some choices of $u_\omega \in U^1_\omega$.

Lemma 9. *Let U^1 denote all utility functions consistent with the single-peaked time-1 preferences. Fix an alternative $m \in \{\min X, \max X\}$ such that for every $x \neq m$ there exists ω such that $x \preceq^1_\omega m$ does not hold. Then for every $v \in U^1$ there is $u \in U^1$ such that $\mathbb{E}\left[u_\omega(x)\right] = \mathbb{E}\left[v_\omega(x)\right] - \mathbf{1}_{x=m}\Delta$.*

Similarly, if U^1 denotes all concave utility functions consistent with the single-peaked time-1 preferences, then for every $v \in U^1$ there is $u \in U^1$ such that $\mathbb{E}[u_\omega(x)] = \mathbb{E}[v_\omega(x)] - \mathbf{1}_{x=m}\Delta$.

Proof. Without loss, assume that $m = \min X$. By assumption, there exists at least one state ω' such that $\max X \preceq^1_{\omega'} m$ does not hold. For this ω', it is then more generally true that $x \preceq^1_{\omega'} m$ does not hold for any $x \neq m$ by the definition of single-peaked preferences. To see this, note that if $x \neq \max X$ and $x \preceq^1_{\omega'} m$, then single-peakness requires that $\max X \preceq^1_{\omega'} x$, and thus that $\max X \preceq^1_{\omega'} \min X$.

Now define

$$u_\omega(x) = v_\omega(x) - \mathbf{1}_{x=m \text{ and } \omega=\omega'} \frac{\Delta}{f_{\omega'}}.$$

By construction, u_ω is identical to v_ω in all states $\omega \neq \omega'$. Moreover, because m is not weakly preferred to any other alternative $x \neq m$ in state ω', it follows that if $v_{\omega'}$ is consistent with $\preceq^1_{\omega'}$ then subtracting a positive constant from $v'_\omega(m)$ still preserves consistency with $\preceq^1_{\omega'}$ as well as single-peakness. Thus, $v \in U^1$ if $u \in U^1$, and $\mathbb{E}[u_\omega(x)] = \mathbb{E}[v_\omega(x)] - \mathbf{1}_{x=m}\Delta$ by construction.

Finally, note that if v_ω is concave for all ω, including $\omega = \omega'$, then subtracting a positive constant from $v_{\omega'}(m)$ also preserves concavity. This establishes the last part of the Lemma. □

Lemma 10. *Assume that time-1 preferences are single-peaked and have no indifferences. Let U^1 denote all single-peaked utility functions with no indifferences that are consistent with the time-1 preferences. Fix an alternative $m \notin \{\min X, \max X\}$ such that for every $x \neq m$ there exists ω such that $x \prec^1_\omega m$ does not hold. Then for every $v \in U^1$ there is $u \in U^1$ such that $\mathbb{E}[u_\omega(x)] = \mathbb{E}[v_\omega(x)] - \mathbf{1}_{x=m}\Delta$.*

Proof. By definition, there exist states ω' and ω'' such that $\min X \prec^1_{\omega'} m$ and $\max X \prec^1_{\omega''} m$ do not hold. We will show that $u : X \times \Omega \to \mathbb{R}$ defined as below belongs to U^1 if $v \in U^1$:

$$u_\omega(x) = \begin{cases} v_\omega(x) + \Delta & \text{if } \omega \notin \{\omega', \omega''\} \\ v_\omega(x) - \mathbf{1}_{x \geq m} \frac{\Delta}{f_{\omega'}} + \Delta & \text{if } \omega = \omega' \\ v_\omega(x) - \mathbf{1}_{x \leq m} \frac{\Delta}{f_{\omega''}} + \Delta & \text{if } \omega = \omega'' \end{cases}.$$

First, note that if $\min X \prec^1_{\omega'} m$ does not hold, then $y \prec^1_{\omega'} m$ cannot hold for any $y < m$. Otherwise, if $y \prec^1_{\omega'} m$, then $\min X \prec^1_{\omega'} y$ by definition of single peakness, and thus $\min X \prec^1_{\omega'} m$. Similarly, $m \prec^1_{\omega'} x$ cannot hold for any $x > m$. We now argue that $y \prec^1_{\omega'} x$ cannot hold for $y < m < x$. If it did, then $m \succ^1_{\omega'} y$ by single-peakedness. But we have already shown that this cannot hold.

Together, this implies that $y \prec^1_{\omega'} x$ cannot hold for any $x \geq m$ and $y < m$ (and indifference cannot occur by the assumption of the lemma). Thus, subtracting a constant from $v_{\omega'}$ for all alternatives $x \geq m$ leads to another utility function compatible with the preference $\preceq^1_{\omega'}$. A symmetric argument implies that subtracting a constant from $v_{\omega''}$ for all alternatives $x \leq m$ leads to another utility compatible with the preference $\preceq^1_{\omega''}$.

Thus, the utility function u defined above belongs to U^1 if $v \in U^1$ (where we also us the obvious fact that adding Δ to the utility from all elements preserves inclusion in U^1).

Observe that

$$\begin{aligned}\mathbb{E}\left[u_\omega(x)\right] &= \mathbb{E}\left[v_\omega(x)\right] - \mathbb{E}\left[\mathbf{1}_{\omega=\omega'}\right]\mathbf{1}_{x\geq m}\frac{\Delta}{f_{\omega'}} - \mathbb{E}\left[\mathbf{1}_{\omega=\omega''}\right]\mathbf{1}_{x\leq m}\frac{\Delta}{f_{\omega''}} + \Delta \\ &= \mathbb{E}\left[v_\omega(x)\right] - \mathbf{1}_{x\geq m}\Delta - \mathbf{1}_{x\leq m}\Delta + \Delta \\ &= \mathbb{E}\left[v_\omega(x)\right] - \mathbf{1}_{x=m}\Delta\,,\end{aligned}$$

which completes the proof. □

Lemma 11. *Let $L_1, L_2 \in \Delta(|X|)$ be two lotteries over outcomes.*

(i) L_2 dominates L_1 with respect to U^1 if and only if $\bar{u}(L_2) \geq \bar{u}(L_1)$ for all $\bar{u} \in \bar{U}^1$.

(ii) L_2 strictly dominates L_1 with respect to U^1 if and only if $\bar{u}(L_2) > \bar{u}(L_1)$ for all $\bar{u} \in \bar{U}^1$.

Proof. If L_2 weakly dominates L_1 then $u_\omega(L_2) \geq u_\omega(L_1)$ for all $\omega \in \Omega, u \in U^1$. Thus, $\sum_\omega u_\omega(L_2) \geq \sum_\omega u_\omega(L_1)$ for all u in U^1 and hence weak dominance implies $\bar{u}(L_2) \geq \bar{u}(L_1)$ for all $\bar{u} \in \bar{U}^1$. The argument for strict dominance is analogous.

For the opposite direction observe that $\bar{u}(L_2) \geq \bar{u}(L_1)$ for all $\bar{u} \in \bar{U}^1$ implies that for all $u \in U^1$ we have $\sum_\omega u_\omega(L_2) \geq \sum_\omega u_\omega(L_1)$. Note that if $u_\omega \in U^1_\omega$ then $\alpha_\omega u_\omega \in U^1_\omega$ for all $\alpha_\omega > 0$. Thus, for all $\alpha \in \mathbb{R}^{|\Omega|}_{++}$, we have that

$$\sum_\omega \alpha_\omega u_\omega(L_2) \geq \sum_\omega \alpha_\omega u_\omega(L_1)\,.$$

Choosing $\alpha_\omega = \mathbf{1}_{\omega=\tilde{\omega}} + \epsilon\mathbf{1}_{\omega\neq\tilde{\omega}}$ for $\epsilon > 0$ yields that

$$u_{\tilde{\omega}}(L_2) - u_{\tilde{\omega}}(L_1) \geq \epsilon\left(\sum_{\omega\neq\tilde{\omega}} u_\omega(L_1) - u_\omega(L_2)\right)\,.$$

Taking the limit $\epsilon \to 0$ yields that for each state $\tilde{\omega}$ we have that $u_{\tilde{\omega}}(L_2) \geq u_{\tilde{\omega}}(L_1)$, and thus that L_2 weakly dominates L_1. This establishes part (i) of the Lemma. Furthermore, note that if $\bar{u}(L_2) > \bar{u}(L_1)$, and if $u_\omega(L_2) \geq u_\omega(L_1)$ for all ω, then the inequality must be strict for at least one ω, which establishes part (ii) of the Lemma. □

A.3 Proofs of Our Main Results

To simplify exposition we will say that a set is *relatively open* if it is open relative to its affine hull.

Proof of Theorem 1. Let $\bar{U}^1 = \sum_\omega U^1_\omega$ be the set of all utility functions that can be rationalized as sums of single-peaked utility functions representing the time-1 preferences $\preceq^1_\omega$. Let U^0 be the set of all utility functions that are consistent with $\preceq^0$. We will show that if U^0 and $\bar{U}^1$ do not intersect and the data set exhibits no violation of the only weak cycles condition, then there is a simple dominance violation. By Lemma 2 this implies a cyclic dominance violation; hence, if U^0 and $\bar{U}^1$ do not intersect, there must be a cyclic dominance violation.

Because the set of all utility functions consistent with a given (potentially incomplete) preference relation is relatively open by Lemma 7, the set of single-peaked utility functions is relatively open, and the intersection of relatively open sets is relatively open, it follows that U^0 and each U^1_ω are relatively open. Because the Minkowski sum of relatively open sets is relatively open it follows that $\bar{U}^1$ is relatively open.

By Lemma 6 there exists a separating hyperplane $p \in \mathbb{R}^{|X|}$ with $p \neq 0$ and $\sum_{x \in X} p_x = 0$ (throughout, we denote by p_x the entry of p corresponding to alternative $x \in X$) such that for all $u^0 \in U^0, u^1 \in \bar{U}^1$

$$p \cdot u^0 \geq 0 \geq p \cdot u^1 \tag{9}$$

and at least one of the inequalities is strict for all u^0, u^1.

We next argue that this implies the existence of another hyperplane $\tilde{p} \in \mathbb{R}^{|X|}$ that satisfies the same properties and only two non-zero entries, which is equivalent to the existence of a simple dominance violation and thus the statement of the theorem. We prove this by induction, showing that if the statement holds for a set with $|X| - 1$ objects then it holds for a set with $|X|$ objects.

Induction hypothesis We first prove the result for two alternatives $|X| = 2$. In this case, by definition p is either $(+\alpha, -\alpha)$ or $(-\alpha, +\alpha)$ for some $\alpha > 0$, and thus the result holds.

Induction step We next prove the induction step. Assume that the result holds whenever the number of alternatives is not more than $|X| - 1$. We consider the preorder $\trianglelefteq$ defined above.

Consider the set of alternatives the is not strictly better than any other alternative according to the preorder $\trianglelefteq$

$$\mathcal{X} = \{x \in X \colon \text{ there does not exist } y \in X \text{ with } y \vartriangleleft x\}$$

Note that $|\mathcal{X}| > 0$ because $\trianglelefteq$ is by construction transitive. Clearly, the elements in $\mathcal{X}$ are either related by indifference $\sim_{\trianglelefteq}$ or unrelated. Pick a maximal subset of this set X^* such that $x \sim_{\trianglelefteq} x'$ for all $x, x' \in X^*$. By definition for all $x' \notin X^*$ and $x \in X^*$ we have that $x' \trianglelefteq x$ can not hold as otherwise there exists an element $y \vartriangleleft x' \trianglelefteq x$ which contradicts that $x \in X^*$ as $\trianglelefteq$ is transitive.

Fix any $x^* \in X^*$. As $x \preceq^0 x^*$ implies $x \trianglelefteq^* x^*$, and thus $x \trianglelefteq x^*$ we have that $x \preceq^0 x^*$ implies $x \in X^*$. By the same argument $x \preceq^1_* x^*$ implies $x \in X^*$.

Thus, if $x^* \in X^*$ and $x \notin X^*$, either $x^* \prec^0 x$, or x is unrelated to x^* by $\preceq^0$. Thus, if $u^0 \in U^0$ then $u^0 - \lambda \mathbf{1}_{x \in X^*} \in U^0$ for all $\lambda \geq 0$, as we can always make alternatives that are either least preferred or unranked relative to others worse without violating the ranking of the alternatives implied by $\preceq^0$. Thus, for every $u^0 \in U^0$ we have that

$$0 \leq p \cdot (u^0 - \lambda \mathbf{1}_{x \in X^*}) = p \cdot u_0 - \lambda \sum_{x \in X^*} p_x \,.$$

Taking $\lambda \to \infty$ implies that

$$\sum_{x \in X^*} p_x \leq 0. \tag{10}$$

Because time-1 preferences are strict, it follows that if there are no simple dominance violations, then for any $x \in X^*$ and any $x' \in X$, there is a state ω' such that $x \succeq^1_{\omega'} x'$ does not hold. Otherwise it would have to be that $x' \preceq^1_* x$, which implies $x' \prec^1_* x$ because we rule out indifference; thus Lemma 1 implies $x' \lhd^* x$, and Lemma 4 $x' \lhd x$, which contradicts $x \in X^*$. If there is a simple dominance violation then we have reached our desired contradiction. So assume that there is not, so that for any $x \in X^*$ and any $x' \in X$, there is a state ω' such that $x \succeq^1_{\omega'} x'$ does not hold.

Then, Lemmas 9 and 10 imply that if $x^* \in X^*$ and $u^1 \in \bar{U}^1$, then $u^1 - \lambda \mathbf{1}_{x=x^*} \in \bar{U}^1$ for all $\lambda \geq 0$. Thus, for every $u^1 \in \bar{U}^1$ we have that

$$0 \geq p \cdot (u^1 - \lambda \mathbf{1}_{x=x^*}) = p \cdot u_1 - \lambda \mathbf{1}_{x=x^*} p_{x^*} \,.$$

Taking $\lambda \to \infty$ implies that $p_{x^*} \geq 0$ for every $x^* \in X^*$. Together with (10), this implies that $p_{x^*} = 0$.

Thus, there exists a vector $p \in \mathbb{R}^{|X|-|X^*|}$ such that $\sum_{x \in X \setminus X^*} p_x = 0$ with $p \neq 0$ such that (9) is satisfied on the set $X \setminus X^*$. As the preferences are single-peaked on $X \setminus X^*$ and this set contains only $|X| - |X^*|$ alternatives, there exists a vector $p \neq 0$ with $\sum_{x \in X} p_x = 0$ and only two non-zero entries, satisfying (9) on that set of alternatives. As this vector corresponds to a simple dominance violation this completes the proof. □

Proof of Theorem 2. As in the proof of Theorem 1, we will show that if U^0 and $\bar{U}^1$ do not intersect, then there must be a simple dominance violation. Because the set of all utility functions consistent with a given (potentially incomplete) preference relation is relatively open by Lemma 7, the set of strictly concave utility functions is relatively open by Lemma 8, and the intersection of relatively open sets is relatively open, it follows that U^0 and U^1_ω are relatively open. Because the Minkowski sum of relatively open sets is relatively open it follows that $\bar{U}^1$ is relatively open.

By Lemma 6 there exists a separating hyperplane $p \in \mathbb{R}^{|X|}$ with $p \neq 0$ and $\sum_{x \in X} p_x = 0$ such that equation (9) is satisfied for all $u^0 \in U^0, u^1 \in \bar{U}^1$, with at least one of the inequalities strict. As in the proof of Theorem 1, we show that this implies the existence of another hyperplane $\tilde{p} \in \mathbb{R}^{|X|}$ that consists of only two non-zero entries, which is equivalent to the existence of a simple dominance violation and thus the statement of Theorem 2. We prove this by induction, showing that if the statement holds for a set with $|X| - 1$ objects then it holds for a set with $|X|$ objects.

As in the proof of Theorem 1, the statement holds trivially when $|X| = 2$. We next prove the induction step and assume that the result holds whenever the number of alternatives is less than $|X| - 1$. If there is a simple dominance violation then the statement of the theorem holds. If there is not, then Lemma 5 implies that there exists a set $X^* \subseteq \{\min X, \max X\}$ such that if $x \in X \setminus X^*$ and $x^* \in X^*$, then $x \trianglelefteq x^*$ cannot hold, and such that $x \sim_{\trianglelefteq} x'$ for $x, x' \in X^*$.

As in the proof of Theorem 1, if $u^0 \in U^0$ then $u^0 - \lambda \mathbf{1}_{x \in X^*} \in U^0$ for all $\lambda \geq 0$, as we can always make alternatives that are either least preferred or unranked relative to others worse without violating the ranking of the alternatives implied by $\preceq^0$.

Now for every $u^0 \in U^0$ we have that

$$0 \leq p \cdot (u^0 - \lambda \mathbf{1}_{x \in X^*}) = p \cdot u_0 - \lambda \sum_{x \in X^*} p_x \, .$$

Taking $\lambda \to \infty$ implies that

$$\sum_{x \in X^*} p_x \leq 0. \tag{11}$$

We divide the remainder of the proof into five cases. We rely on the induction step in the last three cases, which imply a vector $p \in \mathbb{R}^{|X|-|X^*|}$ such that $\sum_{x \in X \setminus X^*} p_x = 0$ with $p \neq 0$ such that (9) is satisfied on the set $X \setminus X^*$.

Case (i): Suppose that $X^* = \{\min X, \max X\}$ and either $\min X \prec_*^1 \max X$ or $\max X \prec_*^1 \min X$. In the first case, this implies that $\max X \preceq^0 \min X$; otherwise, we would have $\min X \trianglelefteq \max X$, which violates the assumption that $X^* = \{\min X, \max X\}$. However, if $\max X \preceq^0 \min X$ and $\min X \prec_*^1 \max X$, then there is a simple dominance violation, which establishes the claim. The second case follows analogously.

Case (ii): Suppose that $X^* = \{\min X, \max X\}$, $\min X \sim_*^1 \max X$ and the time-0 preference relates $\min X$ and $\max X$. Then as the time zero preference relates $\min X$ and $\max X$ it must also be that $\min X \sim^0 \max X$ if there is not a simple dominance violation.

Thus, we can just identify the alternatives $\min X$ and $\max X$ with each other to arrive at a problem with $|X|-1$ alternatives. Formally, for any p satisfying (9), note that p' given by $p'_{\min X} = 0$ and $p'_{\max X} = 2p_{\max X}$ also satisfies (9) but belongs to $\mathbb{R}^{|X|-1}$. By the induction hypothesis, there must thus exist a simple dominance violation on $X \setminus \{\min X\}$.

Case (iii): Suppose that $X^* = \{\min X, \max X\}$, $\min X \sim_*^1 \max X$ and the time-0 preference does not relate $\min X$ and $\max X$.Note that by construction, if $x \in X \setminus X^*$ and $x^* \in X^*$, then either $x^* \prec^0 x$ or x^* is unrelated to x. Thus, if the time-zero preference does not relate $\min X$ and $\max X$, then for any $x^* \in X^*$ and *any* $x \neq x^*$, $x \preceq^0 x^*$ cannot hold. Because of this, $u^0 \in U^0$ implies that $u^0 - \lambda \mathbf{1}_{x=x^*} \in U^0$ for all $\lambda \geq 0$ and each $x^* \in X^*$, as we can always make alternatives that are either least preferred or unranked relative to others worse without violating the ranking of the alternatives implied by $\preceq^0$. Thus, for every $u^0 \in U^0$ and $x^* \in X^*$ we have that

$$0 \leq p \cdot (u^0 - \lambda \mathbf{1}_{x=x^*}) = p \cdot u^0 - \lambda \mathbf{1}_{x=x^*} p_{x^*}$$

and taking $\lambda \to \infty$ implies that $p_{x^*} \leq 0$ for each $x^* \in X^*$.

Now because $\min X \sim_*^1 \max X$ and they do not dominate any other alternative in X, we can identify them with each other in $\bar{U}^1$, and thus Lemma 9 implies that if $u^1 \in \bar{U}^1$ then

$$u^1 - \lambda \mathbf{1}_{x \in X^*} \in \bar{U}^1$$

Thus, for every $u^1 \in \bar{U}^1$ we have that

$$0 \geq p \cdot (u^1 - \lambda \mathbf{1}_{x \in X^*}) = p \cdot u_1 - \lambda \sum_{x \in X^*} p_x \, .$$

Taking $\lambda \to \infty$ implies that $\sum_{x \in X^*} p_x \geq 0$, which implies that $p_x = 0$ for each $x \in X^*$.

Case (iv) Suppose that X^* has only one element x^*. Since $x \preceq^1_* x^*$ cannot hold for any $x \neq x^*$, Lemma 9 implies that if $u^1 \in \bar{U}^1$ then $u^1 - \lambda \mathbf{1}_{x=x^*} \in \bar{U}^1$ for all $\lambda \geq 0$. Thus, for every $u^1 \in \bar{U}^1$ we have that

$$0 \geq p \cdot (u^1 - \lambda \mathbf{1}_{x=x^*}) = p \cdot u_1 - \lambda \mathbf{1}_{x=x^*} p_{x^*} \, .$$

Taking $\lambda \to \infty$ implies that $p_{x^*} \geq 0$ for every $x^* \in X^*$. Together with (11), this implies that $p_{x^*} = 0$.

Case (v): Suppose that $X^* = \{\min X, \max X\}$, and that neither $\min X \preceq^1_* \max X$ nor $\max X \preceq^1_* \min X$. Then application of Lemma 9 as in Case (iv) implies that $p_{x^*} \geq 0$ for each $x^* \in X^*$. Together with (11), this again implies that $p_{x^*} = 0$ for all $x^* \in X^*$.

Completing the proof in Cases (iii), (iv), and (v) In all these cases there exists a vector $p \in \mathbb{R}^{|X|-|X^*|}$ such that $\sum_{x \in X \setminus X^*} p_x = 0$ with $p \neq 0$ such that (9) is satisfied on the set $X \setminus X^*$. As the preferences are single-peaked on $X \setminus X^*$ and this set contains only $|X| - |X^*|$ alternatives, there exists a vector p with $|p_x| = 1$ or $p_x = 0$ $\forall x \in X$, satisfying (9) on that set of alternatives. As this vector corresponds to a simple dominance violation this completes the proof. $\square$

Proof of Theorem 4. Assume that $U^0 \cap \bar{U}^1 = \emptyset$ and the data set can thus not be rationalized by TC-EU. By Lemma 6, there exists $p \in \mathbb{R}^{|X|}$ with $p \neq 0$ and $\sum_{x \in X} p_x = 0$ such that $p \cdot u^0 \geq 0 \geq p \cdot u$ for $u^0 \in U^0$ and all $u \in \bar{U}^1$, such that at least one inequality is strict for all $u^0 \in U^0$ and $u \in \bar{U}^1$. Define $L_1, L_2 \in \mathbb{R}^{|X|}$

$$L_1(x) = \frac{\max\{p_x, 0\}}{\sum_{\tilde{x} \in X} \max\{p_{\tilde{x}}, 0\}}$$
$$L_2(x) = \frac{\max\{-p_x, 0\}}{\sum_{\tilde{x} \in X} \max\{-p_{\tilde{x}}, 0\}} \, .$$

Note that by definition, the entries of L_1, L_2 are non-negative and sum up to one, which implies that L_1, L_2 are well-defined lotteries over the alternatives X. Furthermore, we have that

$$0 = \sum_{\tilde{x} \in X} p_{\tilde{x}} = \sum_{\tilde{x} \in X} \max\{p_{\tilde{x}}, 0\} - \sum_{\tilde{x} \in X} \max\{-p_{\tilde{x}}, 0\} \, .$$

This implies that

$$L_1(x) - L_2(x) = \frac{\max\{p_x, 0\}}{\sum_{\tilde{x} \in X} \max\{p_{\tilde{x}}, 0\}} - \frac{\max\{-p_x, 0\}}{\sum_{\tilde{x} \in X} \max\{-p_{\tilde{x}}, 0\}} = p_x \frac{1}{\sum_{\tilde{x} \in X} \max\{p_{\tilde{x}}, 0\}} \, .$$

Thus, $p \cdot u^0 \geq 0$ implies that $u^0(L_1) \geq u^0(L_2)$, and $p \cdot u^0 > 0$ implies that $u^0(L_1) > u^0(L_2)$.

Similarly, $p \cdot u \leq 0$ for all $u \in \bar{U}$ implies that $\bar{u}(L_2) \geq \bar{u}(L_1)$ for all $u \in \bar{U}$ and $p \cdot u < 0$ for all $u \in \bar{U}$ implies that $\bar{u}(L_2) > \bar{u}(L_1)$ for all $u \in \bar{U}$. Thus, by Lemma 11, for all $u^0 \in U^0$, either (i) $u^0(L_1) \geq u^0(L_2)$ and L_2 strictly dominates L_1 with respect to U^1 or (ii) $u^0(L_1) > u^0(L_2)$ and L_2 weakly dominates L_1 with respect to U^1. Hence, according to Definition 8, the data set exhibits a stochastic dominance violation if $U^0 \cap \bar{U}^1 = \emptyset$; i.e., if the data set (U^0, U^1, f) is inconsistent with TC-EU then it exhibits a stochastic dominance violation.

The opposite direction is immediate: Suppose that the data set is consistent with TC-EU and for $u^0 \in U^0, u^1 \in U^1$

$$u^0(x) = \sum_{\omega \in \Omega} f_\omega u^1_\omega(x)\,.$$

Then, $u^1_\omega(L_2) \geq u^1_\omega(L_1)$ for all ω implies that $u^0(L_2) \geq u^0(L_1)$, and thus the data set cannot exhibit dominance violations. □

Proof of Proposition 1. First note that we can restrict to the case where preferences are complete. This is because if the preferences are incomplete and satisfy one of the two conditions, then there exist completions of the preferences that satisfy one of the conditions. And the incomplete preferences can of course be rationalized by TC-EU if their completions can be.

We shall prove the proposition under the assumption that condition 1 holds. The proof for condition 2 uses identical arguments that start with the most preferred time-0 alternatives rather than the least preferred time-0 alternatives.

As in the proof of Theorem 1, Lemma 7 implies that U^0, U^1_ω and $\bar{U}^1$ can be shown to be relatively open. We shall show that under the conditions of the Proposition, if $p \in \mathbb{R}^{|X|}$ satisfies (i) $\sum_{x \in X} p_x = 0$ and (ii) $p \cdot u^0 \geq 0 \geq p \cdot u^1$ for all $u^0 \in U^0, u^1 \in \bar{U}^1$, then $p \equiv 0$.This then implies that $U^0 \cap \bar{U}^1 \neq \emptyset$, which is the statement of the Proposition.

We shall prove this by induction, showing that if the statement holds for a set with no more than $|X| - 1$ objects, then it holds for a set with $|X|$ objects.

Let X^* denote the set of least preferred elements of the time-0 preference. Observe that if $u^0 \in U^0$ then $u^0 - \lambda \mathbf{1}_{x \in X^*} \in U^0$ for all $\lambda \geq 0$: we can always make least preferred alternatives worse without changing the ranking of the alternatives. Thus, for every $u^0 \in U^0$ we have that

$$0 \leq p \cdot (u^0 - \lambda \mathbf{1}_{x \in X^*}) = p \cdot u_0 - \lambda \sum_{x \in X^*} p_x\,. \tag{12}$$

Taking $\lambda \to \infty$ implies that $\sum_{x \in X^*} p_x \leq 0$.

Now by the assumption of the proposition, for each $x^* \in X^*$ there exists a state ω^* in which x^* is the least preferred element. We have that if $u^1 \in U^1$ then $u^1 - \frac{\lambda}{f_{\omega^*}} \mathbf{1}_{x = x^* \text{ and } \omega = \omega^*} \in U^1$ for all $\lambda > 0$. This implies that if $u \in \bar{U}^1$ then $u - \lambda \mathbf{1}_{x = x^*} \in \bar{U}^1$ for all $\lambda \geq 0$. Thus, for every $u \in \bar{U}^1$, we have that

$$0 \geq p \cdot (u - \lambda \mathbf{1}_{x = x^*}) = p \cdot u - \lambda p_{x^*}\,.$$

Taking $\lambda \to \infty$ implies that $p_{x^*} \geq 0$. Together with the condition that $\sum_{x \in X^*} p_x \leq 0$, this implies that $p_{x^*} = 0$ for all $x^* \in X^*$.

Now when $|X| = 2$, the above implies that at least one of the elements of p must equal 0, which implies that all elements of p must equal zero by the condition that $\sum_{x \in X} p_x = 0$.

When $|X| > 2$, the above implies that at least one of the elements of p must equal 0. Say that this element corresponds to an element x^*, and note that the condition of the proposition still applies to preferences on the set $X \setminus \{x^*\}$. Thus, if the result holds for sets of size $|X| - 1 \geq 2$, it must hold for sets of size $|X|$. $\square$

Proof of Proposition 3. As preferences are responsive, there exists $z_{\omega,y}$ and $z^\circ_{\omega,y}$ known to the analyst such that

$$(y, z_{\omega,y}) \sim^1_\omega (y^\circ, z^\circ_{\omega,y}) .$$

As preferences are additive, this means that

$$h^1_\omega(y) + g(z_{\omega,y}) = h^1_\omega(y^\circ) + g(z^\circ_{\omega,y}) .$$

This implies that $h^1_\omega(y) - h^1_\omega(y^\circ) = g(z^\circ_{\omega,y}) - g(z_{\omega,y})$ for each ω, and thus that

$$\sum_\omega \left(h^1_\omega(y) - h^1_\omega(y^\circ)\right) = \sum_\omega \left(g(z^\circ_{\omega,y}) - g(z_{\omega,y})\right)$$

By the same argument, there exist $z_{0,y}$ and $z^\circ_{0,y}$ such that

$$h^0(y) - h^0(y^\circ) = g(z^\circ_{0,y}) - g(z_{0,y}) .$$

Dividing the terms yields that

$$\frac{h^0(y) - h^0(y^\circ)}{\sum_\omega f_\omega (h^1_\omega(y) - h^1_\omega(y^\circ))} = \frac{g(z^\circ_{0,y}) - g(z_{0,y})}{\sum_\omega f_\omega (z^\circ_{\omega,y} - z_{\omega,y})} .$$

As all terms on the right-hand-side are observable to the analyst it follows that the left-hand-side is point-identified. $\square$

Proof of Proposition 4. We first argue necessity of the conditions. Revealed additive separability is implied by additive separability, by simply setting $h^0 \equiv \frac{\theta^1_\omega}{\theta^2_\omega} h$ and $h^1_\omega \equiv \frac{\theta^1_\omega}{\beta \theta^2_\omega} h$. To see that the condition on β is necessary, observe that for any representation (h, g) of preferences consistent with (5)

$$\begin{aligned}
\beta &= \frac{(\sum_\omega f_\omega \theta^1_\omega)(h(y) - h(y^\circ))}{\sum_\omega f_\omega \theta^1_\omega \frac{1}{\beta}(h(y) - h(y^\circ))} = \frac{(\sum_\omega f_\omega \theta^1_\omega)/(\sum_\omega f_\omega \theta^2_\omega)(h(y) - h(y^\circ))}{(\sum_\omega f_\omega \theta^2_\omega)^{-1}(\sum_\omega f_\omega \theta^1_\omega \frac{1}{\beta}(h(y) - h(y^\circ)))} \\
&= \frac{g(z^\circ) - g(z_{0,y})}{(\sum_\omega f_\omega \theta^2_\omega)^{-1} \sum_\omega f_\omega \theta^2_\omega (g(z_{\omega,y}) - g(z^\circ_{\omega,y}))} = \frac{g(z^\circ) - g(z_{0,y})}{\sum_\omega \alpha_\omega (g(z_{\omega,y}) - g(z^\circ_{\omega,y}))} \\
&\in \left(\min_\omega \frac{g(z^\circ) - g(z_{0,y})}{g(z_{\omega,y}) - g(z^\circ_{\omega,y})}, \max_\omega \frac{g(z^\circ) - g(z_{0,y})}{g(z_{\omega,y}) - g(z^\circ_{\omega,y})} \right) = (\min \hat{\beta}_\omega, \max \hat{\beta}_\omega) ,
\end{aligned}$$

where we define $\alpha_\omega = f_\omega \theta^2_\omega / (\sum_{\omega'} f_{\omega'} \theta^2_{\omega'})$.

We next prove sufficiency. By the assumptions of the proposition there exist weights $\alpha \in \Delta^{|\Omega|}$ such that

$$\beta = \frac{1}{\sum_{\omega \in \Omega} \alpha_\omega \hat{\beta}_\omega^{-1}} \,.$$

We define $\theta^1_\omega = \frac{\alpha_\omega}{f_\omega} \hat{\beta}_\omega^{-1}$, $\theta^2_\omega = \frac{\alpha_\omega}{f_\omega}$. We note that as $\frac{h^1_\omega(y)}{h^0(y)} = \frac{h^1_\omega(y^\circ)}{h^0(y^\circ)}$ by the assumptions of the proposition,

$$\hat{\beta}_\omega = \frac{g(z^\circ_{0,y}) - g(z_{0,y})}{g(z^\circ_{\omega,y}) - g(z_{\omega,y})} = \frac{h^0(y) - h^0(y^\circ)}{h^1_\omega(y) - h^1_\omega(y^\circ)} = \frac{h^0(y) - h^0(y^\circ)}{h^0(y)\frac{h^1_\omega(y)}{h^0(y)} - h^0(y^\circ)\frac{h^1_\omega(y^\circ)}{h^0(y^\circ)}} = \frac{h^0(y^\circ)}{h^1_\omega(y^\circ)} \,.$$

We thus have that

$$h^1_\omega(y) + g(z) = h^0(y)\frac{h^1_\omega(y^\circ)}{h^0(y^\circ)} + g(z) = h^0(y)\hat{\beta}_\omega^{-1} + g(z) = h^0(y)\frac{\theta^1_\omega}{\theta^2_\omega} + g(z) \,.$$

Consequently, the utility

$$\theta^1_\omega h^0(y) + \theta^2_\omega g(z)$$

represents the preference $\preceq^1_\omega$. Finally, we have that

$$\sum_\omega f_\omega \theta^1_\omega h^0(y) = h^0(y) \sum_\omega \alpha_\omega \hat{\beta}_\omega^{-1} = \frac{1}{\beta} h^0(y)$$
$$\sum_\omega f_\omega \theta^2_\omega g(z) = g(z) \sum_\omega \alpha_\omega = g(z) \,,$$

which completes the proof. □

B Online Supplementary Appendix

In this appendix, we illustrate the additional tests that researchers can conduct to gauge the importance of random taste shocks in their data. We also illustrate how researchers can apply Propositions 3 and 4 to data sets featuring these taste shocks to point or set identify the quasi-hyperbolic discounting model. We provide these illustrations by reanalyzing the data from the important work of Augenblick et al. (2015) and Augenblick and Rabin (2019), which generate ideal data sets for these types of analyses.

B.1 Gauging Uncertainty

In the Augenblick et al. (2015) experiment, participants are asked to allocate effort—in the form of units of unpleasant tasks such as transcriptions—between weeks 2 and 3 of the experiment. Participants do this in week 1 (time 0) and week 2 (time 1). The week 1 preference is implemented with probability 10 percent, and the week 2 preference is implemented with probability 90 percent.

Individuals also complete this task a second time, in weeks 4-6. Consistent with Augenblick et al., we call weeks 1-3 block 1, denoted $j = 1$, and we call weeks 4-6 block 2, denoted $j = 2$.

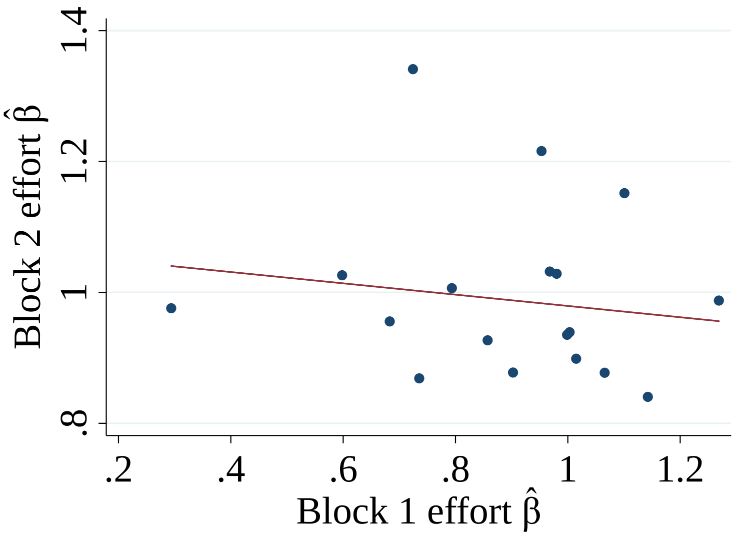
Throughout this appendix, we follow Augenblick et al. in limiting analysis to the "80 subjects who completed all aspects of the experiment with positive variation in their responses in each week." We use $\hat{\beta}_i^1$ to denote an individual-level estimate of person i's present focus from block 1, and we use $\hat{\beta}_i^2$ to denote an individual-level estimate of person i's present focus from block 2.

We find a statistically insignificant and directionally negative correlation between $\hat{\beta}_i^1$ and $\hat{\beta}_i^2$ of -0.08, with a 95% confidence interval of $[-0.29, 0.15]$. This rules out correlations larger than 0.15 with high confidence, and suggests that taste shocks and other forms of updating explain most of the variation in individuals' time-1 revisions of their time-0 choices. The upper bound of 0.15 on the correlation suggests that stable time preferences cannot explain more than about $0.15^2 \approx 2$ percent of the variance in time-1 revisions of time-0 preferences. Figure B1 presents the relationship between the two estimates of present focus in the real-effort task.

Another variable analyzed by Augenblick et al. is an indicator for present focus: $PB_i^j = \hat{\beta}_i^j < 1$. We find that the correlation between PB_i^1 and PB_i^2 is 0.04, with a 95% confidence interval of $[-0.18, 0.26]$.

We also note that this low correlation is unlikely to be due to mere "noise" in the data. As Figure B2 shows, subjects' revised (time 1) allocations of effort are very similar to subjects' initial (time 0) allocations of effort. The correlations between revised and initial effort share allocations are 0.76 and 0.83 in blocks 1 and 2, respectively. A regression of revised on initial allocations produces coefficients of 0.83 (cluster-robust SE 0.05) and 0.82 (cluster-robust SE 0.04) in blocks 1 and 2, respectively, and these decrease only modestly to 0.57 (cluster-robust SE 0.08) and 0.68 (cluster-robust SE 0.06) when including dummies for the five different task rates and the type of task (Tetris versus transcription). This persistence in allocation preferences across time is consistent with large individual differences in preferences over effort allocation. Sources of these differences

Figure B1: Within-person stability of the individual-level estimates of present focus over effort



Notes: This figure presents a binned scatter plot of the relationship between $\hat{\beta}_i^2$, the block 2 individual-level estimates of present focus over effort (y-axis) and $\hat{\beta}_i^1$, the block 1 individual-level estimates of present focus over effort (x-axis).

could include individuals already knowing at time 0 whether they will be relatively more busy in week 2 or in week 3, or stable individual differences in the curvature of the effort cost function. In sum, these results suggest that stable individual differences can be well-measured in the Augenblick et al. (2015) design, but that individual differences in the present focus parameter β are not the primary explanation for the variation in the differences between time 1 and time 0 preferences.

B.2 Identification

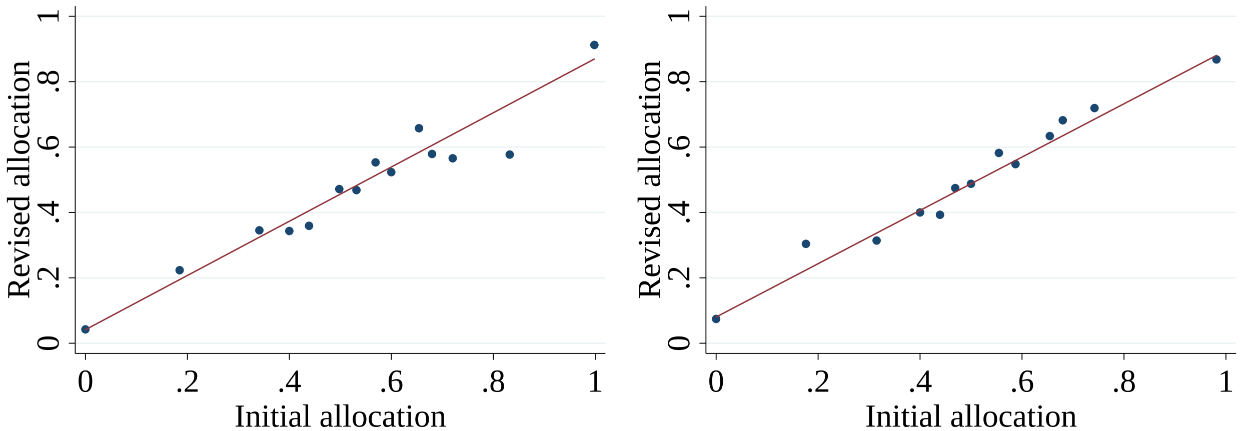
We next show how Propositions 3 and 4 can be used to draw inferences about present focus using data in Augenblick et al. (2015) and Augenblick and Rabin (2019).

B.2.1 Set Identification in Augenblick et al. (2015)

We begin by discussing Augenblick et al. (2015). As different decisions are taken at time 0, we define a state (α, ω) to capture both the variation in time 0 and time 1 choices. We denote by $f_{\omega|\alpha}$ the probability of the time-1 state ω conditional on the time-0 state α. Suppose that the cost of effort is additively separable and isoelastic, as assumed in Augenblick et al. (2015). Suppose, moreover, that taste shocks are multiplicative, as in the assumptions of Proposition 4; i.e.,

$$u^0(y,z) = \delta\bar{\theta}^1_\alpha y^\gamma + \quad \delta^2\bar{\theta}^2_\alpha z^\gamma$$
$$u^1_\omega(y,z) = \quad \theta^1_{\omega,\alpha} y^\gamma + \beta\delta\theta^2_{\omega,\alpha} z^\gamma$$

Figure B2: Revised vs. Initial Effort Allocations in Block 1 (left) and Block 2 (right)



Notes: This figure presents binned scatter plots of the revised versus the initial allocations of effort, pooled over the five task rates and the two types of tasks. The first panel presents results for block 1 (weeks 1-3), while the second panel presents results for block 2 (weeks 4-6). The x-axes correspond to the average effort share designated for weeks 2 and 5 by subjects in weeks 1 and 4, respectively. The y-axes correspond to the average effort share designated for weeks 2 and 5 by subjects in weeks 2 and 4, respectively.

where $\bar{\theta}^1_\alpha := \sum_\omega f_{\omega|\alpha}\theta^1_\omega$, $\bar{\theta}^2_\alpha := \sum_\omega f_{\omega|\alpha}\theta^2_\omega$. We assume that γ is known; Augenblick et al. identify this parameter by observing how preferred allocations vary with the returns to time-2 effort.

The analysis in Supplementary Appendix B.1 above has shown that empirically, much of the variation in individual-level point estimate $\hat{\beta}_i$ is *not* due to variation in time preferences, but rather due to taste shocks. We thus begin by assuming that β is homogeneous in the population. Note that for each value of α that fixes a time-0 choice, $\hat{\beta}_{\omega,\alpha}$—as defined in equation (6) in the body of the paper—is the key object in Proposition 4. In terms of the assumed model primitives, $\hat{\beta}_{\omega,\alpha}$ is given by

$$\hat{\beta}_{\omega,\alpha} = \frac{g(z^\circ_{0,\alpha,y}) - g(z_{0,y})}{g(z^\circ_{\omega,\alpha,y}) - g(z_{\omega,\alpha,y})} = \frac{\frac{1}{\sum_\omega f_{\omega|\alpha}\theta^2_{\omega,\alpha}}\left(\sum_{\omega\in\Omega} f_{\omega|\alpha}\theta^1_{\omega,\alpha}h(y) - \sum_{\omega\in\Omega} f_{\omega|\alpha}\theta^1_{\omega,\alpha}h(y^\circ)\right)}{\frac{1}{\beta\theta^2_{\omega,\alpha}}\left(\theta^1_{\omega,\alpha}h(y) - \theta^1_{\omega,\alpha}h(y^\circ)\right)}$$

$$= \beta\frac{\bar{\theta}^1_\alpha/\bar{\theta}^2_\alpha}{\theta^1_{\omega,\alpha}/\theta^2_{\omega,\alpha}}\,.$$

Similarly, taking first-order conditions for each realization of (α,ω) implies an equation analogous to equation (6) of Augenblick et al., except with β replaced by

$$\hat{\beta}_{\omega,\alpha} = \beta\frac{\bar{\theta}^1_\alpha/\bar{\theta}^2_\alpha}{\theta^1_{\omega,\alpha}/\theta^2_{\omega,\alpha}}\,.$$

Thus, the individual-level estimates $\hat{\beta}$ produced by Augenblick et al. give the $\hat{\beta}_{\omega,\alpha}$ that Proposition 4 applies to. Intuitively, this is because under the assumption of isoelastic cost functions with equal curvature, Augenblick et al.'s budget set variation allows them to estimate g up to multiplicative taste shocks. Thus, their data set is informationally equivalent to the data set assumed in Proposition 4.

To facilitate identification of β assume that the distribution of $\frac{\bar{\theta}^1_\alpha/\bar{\theta}^2_\alpha}{\theta^1_{\omega,\alpha}/\theta^2_{\omega,\alpha}}$ conditional on α is independent of α. This assumption is satisfied, for example, if taste shocks are multiplicatively separable over time: $\theta^t_{\omega,\alpha} = \bar{\theta}^t_\alpha \times \epsilon^t_\omega$ for $t \in \{1, 2\}$. If one does not make such an assumption, more values of β are consistent with the data. This independence assumption implies that the distribution of $\hat{\beta}_{\omega,\alpha}$ does not depend on α. Proposition 4 then implies that $\beta \in [\min \hat{\beta}_{\omega,\alpha}, \max \hat{\beta}_{\omega,\alpha}]$.

Because Augenblick et al. produce an estimate of $\hat{\beta}_{\omega,\alpha}$ for each individual when they estimate equation (6) individual by individual (under the homogeneity assumption), the distribution of $\hat{\beta}_{\omega,\alpha}$ is estimated by the empirical distribution of Augenblick et al.'s individual-level estimates $\hat{\beta}_i$, where i indexes the subjects.[30]

Using their estimates we obtain that the 90-percent intervals that exclude the 5 percent highest and 5 percent lowest estimates of $\hat{\beta}_i$ are [0.43,1.17] in block 1, and [0.57, 1.25] in block 2. The 80-percent intervals that exclude the 10 percent highest and 10 percent lowest estimates of $\hat{\beta}_i$ are [0.66,1.12] in block 1 and [0.69,1.14] in block 2.[31] Thus, the data are consistent both with future focus, as well as with present focus if one allows for taste shocks.

Of course, the complete homogeneity assumption may be too strong, even if the analysis surrounding Figure B1 above suggests that it might be a good approximation. Clearly, relaxing this assumption can only make identification harder. Augenblick et al. (2015) report that individuals who desire a choice set restriction in week 4 have a lower estimate of $\hat{\beta}_i$. The average value of $\hat{\beta}_i$ is 0.86 and 0.95 in blocks 1 and 2, respectively, for individuals who want the choice set restriction; it is 0.96 and 1.05 in blocks 1 and 2, respectively, for individuals who do not want the choice set restriction. Thus, there are at least some individual differences in present focus. Our results are easily compatible with the incorporation of observed heterogeneity, as set identification is possible when splitting individuals by their demand for a commitment device in week 4 of the experiment.

Concretely, assume instead that present focus is homogeneous conditional on the decision to take up or not the commitment device. We can then apply Proposition 4 to each of the subgroups. We obtain the following identified sets if we base them on 80-percent intervals that exclude the 10 percent highest and 10 percent lowest estimates. For those who take up commitment, we infer that their β could belong to the set $[0.44, 1.11]$ based on block 1 data, and to the set $[0.66, 1.13]$ based on block 2 data. For those who do not take up commitment, we obtain $[0.73, 1.13]$ and $[0.84, 1.16]$ for blocks 1 and 2, respectively. In particular, note that it is possible that present focus is large for both those who do and don't take up commitment. It is also possible that present focus is small for both groups, and in fact that those who do not take up commitment contracts are future-focused. Both possibilities are realistic. As Carrera et al. (2021) show, uncertainty about the future erodes demand for commitment even among present-focused individuals, so that large

[30] Formally, the empirical distribution converges to the distribution of $\hat{\beta}_{\omega,\alpha}$ as the number of subjects goes to infinity.

[31] We report our results for the two blocks separately to facilitate immediate comparison to the results in Augenblick et al. (2015), who focus on block 1 in the body of the paper and relegate separate analysis of block 2 to the appendix.

present focus is possible even among those who do not take up commitment. On the other hand, Carrera et al. (2021) also show that there may be noise and confusion in commitment take-up, so that even individuals who are time-consistent might erroneously take up commitment contracts. Finally, note that individuals who are future-focused are still time-inconsistent, and thus might take up commitment contracts as well. Thus, theoretically, any value of β can be consistent with commitment take-up, and our identified set is consistent with this theoretical ambiguity.

B.2.2 Point Identification in Augenblick and Rabin (2019)

Assume that there is no variation in θ^2 and that preferences over the second dimension are known. As the second consumption dimension in Augenblick and Rabin (2019) corresponds to small monetary amounts, this assumption is equivalent to assuming that taste shocks do not affect preferences over money and the agent is (approximately) risk neutral over small monetary amounts. We relax the assumption that β is homogeneous and allow for β to be a random variable (capturing variation of β in the population). We first observe that for a fixed value of β and a fixed time-0 choice captured by α, we have that

$$\frac{1}{\beta} = \sum_{\omega \in \Omega} f_{\omega|\alpha} \frac{1}{\hat{\beta}_{\omega,\alpha}} = \mathbb{E}\left[\frac{1}{\hat{\beta}_{\omega,\alpha}} \middle| \alpha, \beta \right].$$

The first equality we derived in equation (7) in the body of the paper and the second equality follows by definition because $f_{\omega|\alpha}$ is the probability of the time-1 state realization ω conditional on the time-0 state realization α. Then taking iterated expectations, we obtain the following immediate corollary of equation (7):

$$\mathbb{E}\left[\frac{1}{\beta}\right] = \mathbb{E}\left[\mathbb{E}\left[\frac{1}{\hat{\beta}_{\omega,\alpha}} \middle| \alpha, \beta \right] \right] = \mathbb{E}\left[\frac{1}{\hat{\beta}_{\omega,\alpha}} \right]. \tag{13}$$

Thus, the average over the observed estimates $\frac{1}{\hat{\beta}_i}$—i.e., present focus estimates produced from each individual's data—constitute an unbiased estimator of $\frac{1}{\beta}$. The logic above does not rely on multiplicative taste shocks. Nearly-identical reasoning can be used to derive (13) under the more general assumptions of Proposition 3.

References

Abbring, Jaap H. and Øystein Daljord, "Identifying the Discount Factor in Dynamic Discrete Choice Models," *Quantitative Economics*, 2020, *11* (2), 471–501.

Abebe, Girum, Stefano Caria, and Esteban Ortiz-Ospina, "The Selection of Talent: Experimental and Structural Evidence from Ethiopia," *American Economic Review*, 2021, *111* (6), 1757–1806.

Acland, Dan and Matthew R. Levy, "Naiveté, Projection Bias, and Habit Formation in Gym Attendance," *Management Science*, 2015, *61* (1), 146–160.

Aliprantis, Charalambos D. and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, 3 ed., Springer, 2006.

Allcott, Hunt, Joshua Kim, Dmitry Taubinsky, and Jonathan Zinman, "Are High-Interest Loans Predatory? Theory and Evidence from Payday Lending," *Review of Economic Studies*, forthcoming.

_ , Matthew Gentzkow, and Lena Song, "Digital Addiction," *Working Paper*, 2021.

Alós-Ferrer, Carlos, Ernst Fehr, and Nick Netzer, "Time will tell: Recovering preferences when choices are noisy," *Journal of Political Economy*, 2021, *129* (6), 1828–1877.

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

_ , Michael Callen, Karrar Hussain, Muhammad Yasir Khan, and Charles Sprenger, "Using Preference Estimates to Customize Incentives: An Application to Polio Vaccination Drives in Pakistan," 2020. NBER Working Paper No. 22019.

Augenblick, Ned, "Short-Term Discounting of Unpleasant Tasks," *Working Paper*, 2018.

_ and Matthew Rabin, "An Experiment on Time Preference and Misprediction in Unpleasant Tasks," *The Review of Economic Studies*, 2019, *86* (3), 941–975.

_ , Muriel Niederle, and Charles Sprenger, "Working Over Time: Dynamic Inconsistency in Real Effort Tasks," *The Quarterly Journal of Economics*, 2015, *130* (3), 1067–1115.

Badger, Gary J., Warren K. Bickel, Louis A. Giordano, Eric A. Jacobs, George Loewenstein, and Lisa Marsch, "Altered States: The Impact of Immediate Craving on the Valuation of Current and Future Opioids," *Journal of Health Economics*, 2007, *26* (5), 865–876.

Barton, Blake, "Interpersonal Time Inconsistency and Commitment," *Working Paper*, 2015.

Bartos, Vojtech, Michal Bauer, Julie Chytilová, and Ian Levely, "Psychological Effects of Poverty on Time Preferences," *The Economic Journal*, forthcoming.

Benoit, Jean-Pierre and Juan Dubra, "Apparent Overconfidence," *Econometrica*, 2011, *79* (5), 1591–1625.

Breig, Zachary, Matthew Gibson, and Jeffrey G. Shrader, "Why Do We Procrastinate? Present Bias and Optimism," *Working Paper*, 2021.

Browning, Martin and Jeremy Tobacman, "Discounting and Optimism Equivalances," *Working Paper*, 2015.

Brunnermeier, Markus K., Filippos Papakonstantinou, and Jonathan A. Parker, "An Economic Model of the Planning Fallacy," *Working Paper*, 2008.

Buehler, Roger, Dale Griffin, and Johanna Peetz, "Chapter One - The Planning Fallacy: Cognitive, Motivational, and Social Origins," in Mark P. Zanna and James M. Olson, eds., *Advances in Experimental Social Psychology*, Vol. 43, Academic Press, 2010, pp. 1–62.

Carrera, Mariana, Heather Royer, Mark Stehr, Justin Sydnor, and Dmitry Taubinsky, "Who Chooses Commitment? Evidence and Welfare Implications," 2021. NBER Working Paper No. 26161.

Carroll, Gabriel, "An Efficiency Theorem for Incompletely Known Preferences," *Journal of Economic Theory*, 2010, *145* (6), 2463–2470.

Chakraborty, Anujit, Yoram Halevy, and Kota Saito, "The Relation between Behavior under Risk and over Time," *American Economic Review: Insights*, 2020, *2* (1), 1–16.

Chaloupka, Frank J., Matthew R. Levy, and Justin S. White, "Estimating Biases in Smoking Cessation: Evidence from a Field Experiment," *Working Paper*, 2019.

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

Corbett, Colin, "Preferences for Effort and Their Applications," 2016. Unpublished.

DellaVigna, Stefano and Ulrike Malmendier, "Contract Design and Self-Control: Theory and Evidence," *The Quarterly Journal of Economics*, 2004, *119* (2), 353–402.

Epstein, Larry G. and Stanley E. Zin, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 1989, *57* (4), 937–969.

Ericson, Keith M. and David Laibson, "Intertemporal Choice," in B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, eds., *Handbook of Behavioral Economics*, Vol. 2, Elsevier, 2019, chapter 1, pp. 1–67.

Fedyk, Anastassia, "Asymmetric Naivete: Beliefs about Self-Control," *Working Paper*, 2018.

Fudenberg, Drew and David K. Levine, "A Dual-Self Model of Impulse Control," *American Economic Review*, 2006, *96* (5), 1449–1476.

Gul, Faruk and Wolfgang Pesendorfer, "Temptation and Self-Control," *Econometrica*, 2001, *69*, 1403–1435.

Halevy, Yoram, "Diminishing Impatience: Disentangling Time Preference from Uncertain Lifetime," *Working Paper*, 2005.

—, "Strotz Meets Allais: Diminishing Impatience and the Certainty Effect," *American Economic Review*, 2008, *98* (3), 1145–162.

— , "Time Consistency: Stationarity and Time Invariance," *Econometrica*, 2015, *83* (1), 335–352.

Heidhues, Paul and Botond Kőszegi, "Futile Attempts at Self-Control," *Journal of the European Economic Association*, 2009, *7* (2), 423–434.

— **and Philipp Strack**, "Identifying Present Bias from the Timing of Choices," *American Economic Review*, forthcoming.

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

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

Jackson, Matthew O. and Leeat Yariv, "Collective Dynamic Choice: The Necessity of Time Inconsistency," *American Economic Journal: Microeconomics*, 2015, *7* (4), 150–78.

Kahneman, Daniel and Amos Tversky, "Intuitive Prediction: Biases and Corrective Procedures," in Daniel Kahneman, Paul Slovic, and Amos Tversky, eds., *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge University Press, 1982, pp. 414–421.

Kölle, Felix and Lukas Wenner, "Time-Inconsistent Generosity: Present Bias Across Individual and Social Contexts," *Working Paper*, 2019.

Kuchler, Theresa and Michaela Pagel, "Sticking to Your Plan: The Role of Present Bias for Credit Card Paydown," *Journal of Financial Economics*, 2021, *139* (2), 359–388.

Laibson, David, "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 1997, *112* (2), 443–478.

— , "Why Don't Present-Biased Agents Make Commitments?," *American Economic Review*, 2015, *105* (5), 267–272.

Levy, Matthew and Pasquale Schiraldi, "Identification of Intertemporal Preferences from Choice Set Variation," *Working Paper*, 2021.

Magnac, Therry and David Thesmar, "Identifying Dynamic Discrete Decision Processes," *Econometrica*, 2002, *70* (2), 801–816.

Mahajan, Aprajit, Christian Michel, and Alessandro Tarozzi, "Identification of Time-Inconsistent Models: The Case of Insecticide Treated Nets," *Working Paper*, 2020.

McClure, Samuel M., Keith M. Ericson, David I. Laibson, George Loewenstein, and Johanthan D. Cohen, "Time Discounting for Primary Rewards," *Journal of Neuroscience*, 2007, *27*, 5796–5804.

McLennan, Andrew, "Ordinal Efficiency and the Polyhedral Separating Hyperplane Theorem," *Journal of Economic Theory*, 2002, *105* (2), 435–449.

Milkman, Katherine L., Todd Rogers, and Max H. Bazerman, "Highbrow Films Gather Dust: Time-Inconsistent Preferences and Online DVD Rentals," *Management Science*, 2009, *55* (6), 1047–1059.

Millner, Antony, “Nondogmatic Social Discounting,” *American Economic Review*, 2020, *110* (3), 760–75.

O’Donoghue, Ted and Matthew Rabin, “Doing It Now or Later,” *American Economic Review*, 1999, *89* (1), 103–124.

Read, Daniel and Barbara van Leeuwen, “Predicting Hunger: The Effects of Appetite and Delay on Choice,” *Organizational Behavior and Human Decision Processes*, 1998, *76* (2), 189–205.

_, George Loewenstein, and Shobana Kalyanaraman, “Mixing Virtue and Vice: Combining the Immediacy Effect and the Diversification Heuristic,” *Journal of Behavioral Decision Making*, 1999, *12* (4), 257–273.

Sadoff, Sally, Anya Samek, and Charles Sprenger, “Dynamic Inconsistency in Food Choice: Experimental Evidence from a Food Desert,” *Review of Economic Studies*, 2019, *87* (4), 1–35.

Strotz, Robert H., “Myopia and Inconsistency in Dynamic Utility Maximization,” *The Review of Economic Studies*, 1955, pp. 165–180.