Temptation, Welfare and Revealed Preference*

Jawwad Noor

October 25, 2004

Abstract

Due to temptation, an agent’s choices may not respect his normative preference (his view of what is best for his welfare). Gul and Pesendorfer [7, 8] are the first to provide behavioral foundations for a theory of temptation. Adopting a preference over choice problems as the primitive, they hypothesize that temptation creates a preference for commitment, which in turn reveals the agent’s normative preference. For instance, an addict may enter rehabilitation, thereby also revealing his normative preference for abstinence. However, this paper notes that temptation may in fact create the absence of a preference for commitment: the addict may be tempted to prolong his addiction, and thus, may postpone rehabilitation indefinitely.

An alternative approach to providing foundations for a theory of temptation is introduced. Motivated by the experimental evidence on preference reversals, it is hypothesized that delayed temptations are easier to resist than immediate temptations. Normative preference is derived via choices between sufficiently delayed alternatives. With a choice correspondence as the primitive, agents who are ‘tempted not to commit’ are modeled. The foundations of the model are used to provide evidence supporting such temptation. The model is contrasted with one that adopts preferences over choice problems as the primitive. It is argued that such a primitive is not empirically meaningful since it requires us to observe behavior in the absence of temptation.

Keywords: Self-Control, Commitment, Preference Reversals, Revealed Preference.

JEL classification number: D11, D60.

*Department of Economics, University of Rochester, Rochester, NY 14627. Email: jwdb@troi.cc.rochester.edu. An earlier version of this paper was circulated under the title ‘Temptation and Revealed Preference’. I am greatly indebted to Larry Epstein for his many suggestions and his guidance throughout the project. I have also benefitted from comments from Paulo Barelli, Barton Lipman, Josef Perktold, Alvin Roth, Alvaro Sandroni, William Thomson, Gábor Virág, seminar participants at the University of Rochester, and participants at the Canadian Economic Theory Conference 2004 and the Risk, Uncertainty and Decisions Conference 2004. Any errors are my responsibility.
1 Introduction

An addict may desire drugs despite viewing abstinence as the right choice, and a consumer may want to make an impulsive purchase while simultaneously telling himself that he really should not. Such examples suggest that it is not one, but two preference orderings that underlie choice behavior: one that reflects urges (call it the temptation preference) and another that reflects the agent’s view of what is in his best interest, that is, his view of what he should do (call it his normative preference). Choice behavior is the outcome of an aggregation of temptation preference and normative preference. An agent is said to experience temptation when his temptation preference conflicts with his normative preference. He is said to have self-control problems when his choices do not necessarily respect his normative preference.

The idea that choices may not respect normative preference is in contrast with the traditional view that identifies normative preference with choice. Indeed, the choices of an agent with self-control problems may not serve as a guide for welfare policy – we cannot conclude just by observing an addict’s choice to take drugs that making drugs freely available to him would make him better off. The aim of welfare policy is to maximize an agent’s normative preference, that is, to make him better off according to his own view of what is best for his welfare. Therefore, one would like to know if normative preference is observable.

This paper is concerned with the behavioral foundations of theories of choice under temptation, and in particular, the behavioral foundations of normative preference. What behavior reveals that an agent struggles with two preference orderings? How can normative preference be elicited from choices? How can we identify what he finds tempting? Pinning down these concepts in terms of observable behavior makes it possible to supply empirical evidence in support of hypotheses about agents with self-control problems. We consider such an application also. More specifically, this paper achieves the following:

- It evaluates the approach to providing foundations that exists in the literature, and argues that the behavioral characterizations yielded by the approach are not necessarily based on observable behavior.

- It introduces a different approach to providing foundations which, in turn, yields behavioral characterizations in terms of observable behavior. The relationship with the existing approach is studied.

- While the literature has typically assumed that agents are tempted only by immediate consumption, this paper considers a different assumption: agents may be tempted also by future consumption, that is, they may be tempted by opportunities to indulge temptations tomorrow. Empirical evidence is provided and behavioral and welfare implications are discussed.
1.1 The Commitment Approach

Gul and Pesendorfer (henceforth GP) are the first to provide foundations for a model of temptation and self-control [7, 8]. The primitive of their model is a preference $\succ$ over $Z$, the space of choice problems or menus; a menu $x \in Z$ is a nonempty compact subset of some consumption set $C$. A two period time-line is implicit in their model:

$$
\begin{array}{c}
\bullet \\
 t=0 \quad x \succ y \\
\bullet \\
 t=1 \quad x \in x
\end{array}
$$

The preference $\succ$ dictates the agent’s choice of menu in period 0. After choosing a menu $x$, he subsequently makes a choice $c$ from $x$ in period 1. Period 1 choice is subject to temptation. GP hypothesize that the agent anticipates this in period 0, and that $\succ$ contains information about both the agent’s normative and temptation preferences over $C$.

Specifically, GP hypothesize that temptation creates a preference for commitment in period 0. To illustrate this ‘commitment approach’, consider an addict who has a normative preference for abstinence $n$, but cannot resist the temptation by drugs $d$. He is faced with a period 0 choice of whether or not to enter rehabilitation. If he enters rehabilitation, he faces the menu $\{n\}$ in period 1; otherwise he faces $\{n, d\}$. He understands that he will be tempted to choose $d$ if he faces $\{n, d\}$. Hence, in order to avoid temptation, he chooses to enter rehabilitation:

$$\{n\} \succ \{n, d\}.$$ 

That is, period 1 temptation leads the addict to choose commitment in period 0. This choice in turn reveals to an observer the addict’s normative preference for $n$, and his temptation preference for $d$.

GP employ these ideas to behaviorally characterize agents with self-control problems, and also to characterize concepts such as normative preference and temptation: roughly, an agent has a normative preference for $n$ over $d$ if and only if he has a period 0 preference for committing to $n$ rather than committing to $d$, and he is tempted by $d$ if and only if he has a period 0 preference for menus that exclude $d$. However, to better understand these characterizations, we must study the assumptions underlying them.

The assumption underlying the commitment approach is that the agent experiences no temptation in period 0. In GP’s words, “period 0 is ‘special’ in the sense that it is a period prior to the experience of temptation” [8, p 129]. To see this, suppose that period 0 is not

---

1What we refer to as normative preference, GP refer to as commitment preference; they interpret commitment preference as representing the agent’s view of his long-run self-interest [8, p 120].

2Also see Strotz [22], Laibson [15] and O’Donoghue and Rabin [19].

3That is, if he enters rehabilitation, he commits himself to not taking drugs tomorrow, whereas if he declines rehabilitation, drugs are an option tomorrow. Disulfiram treatment for alcoholics is an example of a treatment procedure that removes the option $d$ from the menu since, after taking disulfiram, ingesting even small quantities of alcohol leads to a severe reaction.
special. Imagine that although the addict understands the value of rehabilitation, a part of him does not wish to enter rehabilitation just yet – he is tempted to choose \( \{n, d\} \) over \( \{n\} \). This could be because tomorrow’s \( d \) tempts the addict today, so that he is tempted in period 0 by the opportunity to indulge temptations tomorrow. Indeed, if this temptation is strong enough, the addict declines rehabilitation, and thus exhibits the preference

\[
\{n, d\} \succ {n}.
\]

This demonstrates that when period 0 is not special, temptation may not create a preference for commitment.

Thus, a special period 0 is a necessary ingredient in the commitment approach. It follows that the primitive \( \succ \) of the approach is not just any ranking of menus, but rather an agent’s ranking of menus in the absence of temptation. This immediately raises some questions: Is \( \succ \) observable? Can period 0 be identified? Can we deduce how an agent would rank menus if he did not experience temptation? Is it even possible to distinguish between a ranking that is in the absence of temptation, and one that is not?

To the extent that the preference \( \succ \) is not observable, adopting it as the primitive of a model of temptation is problematic. The purpose of behavioral characterizations is to allow an observer to identify agents who experience temptation, and to elicit the normative and temptation preferences underlying their choice. But, in order to serve such a purpose, it is essential that the characterizations be in terms of observable behavior. They must be empirically meaningful.

### 1.2 The Preference Reversal Approach

This paper notes that there is another source of data that an observer can use to elicit an agent’s normative and temptation preference. We describe this next. For any alternative \( \mu \), let \( \mu^{+t} \) represent another alternative that provides \( \mu \) in \( t \) periods, and let the ranking of such alternatives be captured by a preference ordering \( \preceq \).

Research in psychology has documented a behavior called preference reversals (see Ainslie [2] for a survey of the evidence). In a typical experiment, subjects exhibit the following kinds of choices:

\[
20^{+0} > 30^{+1}, \quad (20^{+0})^{+1} < (30^{+1})^{+1},
\]

where a unit of time is a month. That is, they prefer receiving an immediate $20 to receiving $30 in one month, but reverse their preferences when both these alternatives are pushed into the future by a month. Psychologists since Ainslie [1], Rachlin [20] and Rachlin and Green [21] have interpreted preference reversals in terms of temptation by immediate gratification: A preference for $30 in two months over $20 in one month reveals that the
subjects find the smaller earlier $20 reward inferior. However, they switch preferences in favor of this same inferior reward when it is available immediately.

Our main observation is that, although the temptation by the earlier reward could not be resisted when subjects chose between $20 now and $30 in a month, resisting it became possible when both rewards were (sufficiently) delayed. That is, preference reversals lend themselves to the observation that delayed temptations are easier to resist than immediate temptations.

This suggests a way of deriving an agent’s normative preference from his choices. Let $C$ be a consumption set and let $\succ^*$ denote the agent’s normative preference over $C$. From his ranking $\succeq_t$ of delayed consumption alternatives, derive a set of preference relations $\{\succ_t\}_{t=0}^\infty$ over $C$ as follows: for each $c, \hat{c} \in C$ and $t \geq 0$,

$$c \succ_t \hat{c} \iff c + t \succ \hat{c} + t$$

Thus, for each $t$, $\succ_t$ ranks alternatives in $C$ when these alternatives are delayed by $t$ periods. By the above observation, as $t$ grows, the influence of temptation on the agent’s ranking $\succ_t$ of alternatives diminishes. That is, as $t$ grows, the temptation component underlying $\succ_t$ becomes less significant, and so, each $\succ_t$ provides an increasingly better approximation of the agent’s underlying normative preference $\succ^*$. For this reason, it is intuitive that normative preference be identified with the limit of the sequence $\{\succ_t\}_{t=0}^\infty$:

$$\succ^* \equiv \lim_{t \to \infty} \succ_t,$$

given a suitably defined space of preferences over $C$. This serves as a behavioral characterization of normative preference, and a starting point for providing foundations for a model of temptation.

To summarize, the hypothesis that ‘delayed temptations are easier to resist than immediate temptations’ is an alternative to GP’s hypothesis that ‘temptation creates a preference for commitment’. It constitutes an alternative starting point for behaviorally characterizing agents with self-control problems, and eliciting their normative and temptation preferences. A noteworthy feature of our approach is that its primitive is empirically meaningful – the primitive is a preference ordering $\succeq$ that describes choices that may well be subject to temptation. This is in contrast to the commitment approach that takes choices in the absence of temptation as its primitive.

1.3 An Application

The literature on temptation has focused almost exclusively on agents who are tempted only by immediate consumption (call them Current Temptation or CT agents). But consider the possibility that agents may be tempted also by future consumption, that is, they may be tempted today by the opportunity to consume a tempting alternative tomorrow (refer
to such agents as Future Temptation or FT agents). A behavioral implication of such
temptation was already presented in Section 1.1: when tomorrow’s drugs tempt an addict
today, he may be tempted to prolong his addiction by retaining the possibility of consuming
drugs later. That is, temptation by future consumption can induce agents with self-control
problems to avoid taking advantage of commitment opportunities.

Such behavior is not exhibited by CT agents – they do not care for temptations that
lie in the future, and therefore have no reason not to commit. Consequently, as we observe
next, the welfare implications of a model that allows for temptation by future consumption
are much different from those of a model that rules it out.

• Having restricted attention to CT agents, the literature has typically taken the
view that agents who do not seek commitment are in fact agents who do not have self-
control problems. Conclusions arising from this view include the claim that there is no
room for any welfare improvements for agents who do not seek commitment (see GP [9]),
and also the claim that the temptation literature has doubtful significance since a demand
for commitment appears to be empirically insignificant (see Gale [3] and Kocherlakota [13]).
These conclusions are misplaced if one considers the fact that for FT agents, self-control
problems may be responsible for an absent demand for commitment.

• Since CT agents always take advantage of commitment opportunities, a model
that excludes temptation by future temptation has a very simple and effective welfare
policy prescription: introduce commitment mechanisms into the market. In this way, agents
with self-control problems are able to improve their welfare, and those without self-control
problems are unaffected. However, such a policy prescription may not be effective if one
allows for temptation by future consumption. For instance, FT addicts may not take
advantage of commitment mechanisms despite desiring an end to their addiction. Narcotics
control may be a strictly better policy to improve their welfare.

These observations establish that temptation by future consumption may be worth
studying. But a prior question is whether there exists any evidence of such temptation.
This question is addressed in this paper. The ideas in Section 1.2 are used to characterize
CT and FT agents in terms of observable behavior. By contrasting the axioms that char-
acterize the two agents, we are able to identify the behaviors that distinguish them. This,
in turn, tells us what kind of evidence constitutes support for temptation by future con-
sumption. The result of our analysis is that supporting evidence does indeed exist (Section
6 presents the evidence). We find that the peculiarities of temptation by future consump-
tion include its implications for preference reversals and the demand for mechanisms that
ensure commitment in the future. The evidence supporting such behavior comes from the
preference reversals literature and from experiments on saving behavior.

4Think of addicts who do not seek rehabilitation, or people who do not take advantage of 401(k)s and
IRAs in order to save for their retirement.
An auxiliary aim of this application is to explore what relationship exists, if any, between the commitment approach and the approach outlined earlier. Roughly, this is done in the following way. The two models we axiomatize are dynamic GP-style models where menus are suitably defined dynamic objects (infinite horizon choice problems). Instead of adopting a GP-style preference \( \succ \) over menus as the primitive, we adopt as the primitive a choice correspondence \( C \) that describes choices from menus. These choices are possibly subject to temptation, and thus, arguably, they are observable in principle. We find restrictions on the choice correspondence \( C \) that are necessary and sufficient for it to be ‘generated’ or ‘rationalized’ by a GP-style preference \( \succ \), in a sense made precise in Section 2. This \( \succ \) is interpreted as describing how the agent would rank menus in a hypothetical period 0 where no temptation is experienced. At this point we can ask, for instance, how the normative preference derived from \( \succ \) by the commitment approach is related to the normative preference derived from \( C \) by our approach.

The remainder of the paper is organized as follows. Section 2 provides formal details of our model of temptation by future consumption, and Section 3 presents axioms and the representation result. Section 4 outlines the proof of the representation theorem and describes the relationship between GP’s and our definition of normative preference in the context of the model. A model that excludes temptation by future consumption is axiomatized in Section 5. Evidence supporting temptation by future consumption is discussed in Section 6. We conclude in Section 7 with some discussion of our definition of normative preference. All proofs are collected in appendices.

2 A Model of Future Temptation

In this section we first describe a GP-style model of temptation by future consumption, that is, a model constructed using the commitment approach. We then demonstrate the ‘problem’ of a special period 0 in the context of the model, and conclude by describing a way to avoid the problem.

For any compact metric space \( X \), \( \Delta(X) \) denotes the set of all probability measures on the Borel \( \sigma \)-algebra of \( X \), endowed with the weak convergence topology; \( \Delta(X) \) is compact and metrizable [4]. Let \( \mathcal{K}(X) \) denote the set of all nonempty compact subsets of \( X \). When endowed with the Hausdorff topology, \( \mathcal{K}(X) \) is a compact metric space [6, p. 222].

The set of consumption items is given by \( C \), a compact metric space. The set of menus is \( Z \). Each menu \( z \in Z \) is a compact set of lotteries, where each lottery is a measure over current consumption and a continuation menu – \( Z \) is homeomorphic to \( \mathcal{K}(\Delta(C \times Z)) \). It is also a compact metric space. See [8] for the formal definition of \( Z \). We often let \( \Delta \) denote \( \Delta(C \times Z) \).
Adopt a binary relation $\succsim$ over $\mathbb{Z}$ as the primitive, and consider the following time-line.

![Time-line diagram](image)

The preference $\succsim$ dictates the choice of menu in period 0. The chosen menu $x$ is faced in period 1, and a choice from $x$ is made. If the choice from $x$ is $(c, z)$, the agent receives immediate consumption $c$, and a continuation menu $z$.\(^5\) The continuation menu $z$ is faced in period 2 and a choice is made from it. The process is repeated ad infinitum.

**Future Temptation Preferences**

In [18], we axiomatize *Future Temptation (FT) preferences.* Say that $\succsim$ is an FT preference if it has a representation $W : \mathbb{Z} \rightarrow \mathbb{R}$ of the following form: there exist $\delta$ and $\gamma$, $0 < \gamma < \delta < 1$, continuous functions $u, v : C \rightarrow \mathbb{R}$, and continuous linear functions $U, V : \Delta(C \times Z) \rightarrow \mathbb{R}$ and $\overline{V} : Z \rightarrow \mathbb{R}$ such that for all $x \in \mathbb{Z}$,

\[
W(x) = \max_{\mu \in x} \left\{ U(\mu) + \left( V(\mu) - \max_{\eta \in x} V(\eta) \right) \right\},
\]

where

\[
U(\mu) = \int_{C \times Z} (u(c) + \delta W(y)) \, d\mu(c, y),
\]

\[
V(\mu) = \int_{C \times Z} (v(c) + \gamma \overline{V}(y)) \, d\mu(c, y),
\]

\[
\overline{V}(x) = \max_{\eta \in x} V(\eta).
\]

To understand the representation, first focus on the functional form of $W(\cdot)$. Note that $W(\{\mu\}) = U(\mu)$ for any singleton menu $\{\mu\}$. Thus, $U(\cdot)$ captures the agent’s utility under commitment. Anticipating results in Section 4, we interpret $U(\cdot)$ as a representation of the agent’s normative preference, and henceforth refer to it as normative utility.\(^6\) Interpreting $V(\cdot)$ as temptation utility, the term $|V(\mu) - \max_{\eta \in z} V(\eta)|$ can be understood as the cost of self-control, that is, the cost incurred when the most tempting item in $z$ is not chosen. Hence, (1) states that the utility $W(x)$ of a menu $x$ is the maximum value of normative utility net of self-control cost.

According to the functional form of $U(\cdot)$, normative utility depends on utility from current consumption, and the utility $W(\cdot)$ from a continuation menu discounted by $\delta$. Temptation utility $V(\cdot)$ depends on current consumption and the temptation value $\overline{V}(\cdot)$ of a continuation menu discounted by $\gamma$. As the form of $\overline{V}(\cdot)$ shows, a continuation menu is

\(^5\)If the chosen item is a nondegenerate lottery $\mu$, then the uncertainty plays out before the next period, yielding some $(c, z)$. This leaves the agent with immediate consumption $c$ and the menu $z$ to face in period 2.

\(^6\)GP call it ‘commitment utility’.  

8
as tempting as the most tempting alternative contained in it. Thus, in this model, future consumption tempts the agent via menus. The restriction $\gamma < \delta$ embodies the property that it is easier to resist a temptation when it is pushed into the future. The functional form with $\gamma = 0$ is a representation of GP’s Dynamic Self-Control (DSC) Preferences \cite{8}, which exhibit temptation by immediate consumption only. DSC preferences are not a special case of FT preferences since the latter require $\gamma > 0$.

Recall the time-line. Though (1) only represents the agent’s choice of menu in some period 0, it is suggestive of how he determines his choice from a menu in any subsequent period $t > 0$: he maximizes normative utility net of self-control costs. Since the term ‘$\max_{\eta \in x} V(\eta)$’ in (1) is a constant when $x$ is given, when maximizing over $\mu$ in $x$ the agent essentially maximizes

$$U(\cdot) + V(\cdot).$$

That is, when choosing from a menu $x$, the agent tries to find a compromise between respecting his normative preference and submitting to his temptation preference.

Thus, while the agent’s period 0 preference $\succsim$ is represented by the function $W(\cdot)$, his choices in each subsequent period $t > 0$ are captured by the choice correspondence $C(\cdot, \succsim)$ over $Z$ defined by

$$C(x, \succsim) = \arg \max_{\mu \in x} \{U(\mu) + V(\mu)\}, \quad (2)$$

for all $x \in Z$.\textsuperscript{7} That is, for the choices in the time-line to be produced by an FT agent, there must exist an FT preference $\succsim$ such that:

\begin{align*}
&\bullet \quad \bullet \quad \bullet \\
&x \succsim y \quad (c, x) \in C(x, \succsim) \quad (c', x') \in C(x, \succsim) \\
&t=0 \quad t=1 \quad t=2
\end{align*}

The distinction between the period 0 choice of a menu and period $t > 0$ choice from a menu is important in what follows.

A Question of Foundations

We demonstrate in the context of the FT model that period 0 is special in that it is a period where no temptation is experienced. Recall that the period 0 ranking $\succsim$ of menus is represented by $W(\cdot)$.\textsuperscript{3}

A ranking of menus can also be obtained in any period $t > 0$ by observing the agent’s choice from menus of the following type:

$$\{(c, x), (c, y)\}.$$
Since no choice of current consumption is involved, the choice from this menu is a choice between menus $x$ and $y$ in some period $t > 0$. The choice is determined by $C\{((c, x), (c, y)), \succsim\}$, that is, by solving
\[
\max_{\{(c, x), (c, y)\}} \{u(\cdot) + \delta W(\cdot) + v(\cdot) + \gamma V(\cdot)\}.
\]
Therefore, the period $t > 0$ ranking of menus is represented by
\[
W(\cdot) + \frac{\gamma}{\delta} V(\cdot). \tag{4}
\]

Compare (3) and (4) and conclude that, in general, the ranking of menus in period 0 is different from that in any period $t > 0$. That is, period 0 is special.\(^8\) Since $V(\cdot)$ captures temptation utility from menus, it is evident that the agent’s ranking of menus in period 0 is not subject to temptation.

Thus, the characterization of the FT model offered in [18] involves restrictions on the ranking of menus in the absence of temptation. In order to verify that an agent has FT preferences, one needs to obtain this ranking. But how can this be done? If there exists some period 0, how can it be identified? How do we tell whether an agent’s choices are in the absence of temptation or not? More to the point, can we expect a period 0 to exist? An agent who is tempted by menus will, in general, experience such temptation in all periods. In such a scenario, we would need to deduce how he would behave in the absence of temptation. How this can be done is not obvious, and thus it is not clear that the FT model’s primitive preference $\succsim$ is observable.\(^9\)

**Alternative Foundations**

In order to avoid this problem associated with a special period 0, we drop the period 0 preference $\succsim$ as the primitive of the model. We consider an alternative characterization of FT preferences that is in terms of restrictions on period $t > 0$ choices from menus, instead of period 0 choices between menus. These period $t > 0$ choices are subject to temptation, and thus describe choices that are observable in principle. Let the choices in each period $t > 0$ be summarized by a (time-invariant) choice correspondence $C(\cdot)$ over $Z$. For any choice correspondence $C(\cdot)$ and any FT preference $\succsim$, say that $\succsim$ generates $C(\cdot)$ if,
\[
C(\cdot) = C(\cdot, \succsim),
\]
where $C(\cdot, \succsim)$ is defined by (2). Our problem is to find restrictions on $C(\cdot)$ that imply the existence of an FT preference $\succsim$ that generates $C(\cdot)$. That is, under what conditions on $C(\cdot)$ can we say that, in a hypothetical period 0 where no temptation is experienced, the agent’s ranking of menus $\succsim$ is an FT preference?

\(^8\)Though $V(\cdot)$ has a specific functional form in this example, the argument reveals that a special period 0 arises generically in extensions of DSC preferences that permit temptation by menus.

\(^9\)GP note that DSC preferences do not suffer from this problem. When only immediate consumption is tempting ($\gamma = 0$), problems (3) and (4) are identical, and hence, the ranking of menus in any period $t > 0$ reveals the period 0 ranking of menus to an observer.
These restrictions on $C(\cdot)$ constitute revealed preference foundations for the FT model in the following sense. In standard revealed preference theory, we start with some class of preferences defined over a set of alternatives, say $\Delta$, and for each preference $\lesssim$ in this class we define a choice correspondence $C^*(\cdot, \lesssim)$ over $K(\Delta)$ by

$$C^*(x, \lesssim) = \{\mu \in x : \mu \lesssim \eta \text{ for all } \eta \in x\},$$

for each $x \in K(\Delta)$. Then, we characterize the choice correspondences that can be ‘rationalized’ or ‘generated’ by a preference in the class, that is, we find conditions that must be satisfied by a correspondence $C(\cdot)$ in order for there to exist a preference $\lesssim$ such that

$$C(\cdot) = C^*(\cdot, \lesssim).$$

In an analogous fashion, we are starting with a class of FT preferences, defining an appropriate choice correspondence $C(\cdot, \succsim)$, and seeking to characterize the choice correspondences that may be ‘generated’ by FT preferences. The exercise is nonstandard only in that an FT preference is defined over the space of menus $K(\Delta)$ rather than over the space of alternatives $\Delta$.

It is worth mentioning that the primitive of the model is a choice correspondence $C(\cdot)$ that describes choices that are subject to temptation, and yet we want to derive from these choices a preference ordering $\succsim$ that represents how the agent would rank menus in the absence of temptation. As we will see, such a derivation is achieved by exploiting the ideas contained in Section 1.2.

### 3 Axioms and Representation Result

Generic elements of $Z$ are $x, y, z$ whereas generic elements of $\Delta$ are $\mu, \eta, \nu$. For $\alpha \in [0, 1]$, $\alpha \mu + (1 - \alpha) \eta \in \Delta$ is the measure that assigns $\alpha \mu(A) + (1 - \alpha) \eta(A)$ to each $A$ in the Borel $\sigma$-algebra of $C \times Z$. Similarly, $\alpha x + (1 - \alpha) y \equiv \{\alpha \mu + (1 - \alpha) \eta : \mu \in x, \eta \in y\} \in Z$ is a mixture of the choice problems $x$ and $y$. Denote these mixtures more simply by $\mu^{\alpha \nu}$ and $x^{\alpha y}$ respectively.

The primitive of the model is a closed-valued choice correspondence $C : Z \rightharpoonup \Delta(C \times Z)$ where, for all $x \in Z$, $C(x) \neq \emptyset$ and $C(x) \subset x$. This is a time-invariant choice correspondence that captures the choices an agent would make out of menus at any time $t = 1, 2,...$(see the time-line in Section 2). Thus, our model is dynamic. We introduce some notation to aid exposition:

- Fix $\bar{c} \in C$ throughout. For any $x$, define $x^{t+1} \equiv (\bar{c}, x)$ and inductively for $t > 1$, $x^{t+1} = (\bar{c}, x^{(t-1)})$. Then $x^{t+1} \in \Delta$ is the alternative that yields menu $x$ after $t > 0$ periods, and $\bar{c}$ in all periods between time 0 and $t$. We write $\{\mu\}^{++t}$ as $\mu^{++t}$ and identify $\mu^{++0}$ with $\mu$. The reader should keep in mind that $x^{++t}$ is not a menu, but a degenerate lottery. That is, it is not an element of $Z$ but rather an element of $\Delta$. 

11
• The option that gives \( x \cup y \) (resp. \( x \alpha y \)), after \( t \) periods is denoted \( (x \cup y)^+t \) (resp. \( (x \alpha y)^+t \)).

• Let \( \gtrless \) denote the revealed preference relation (defined on \( \Delta \)) that is generated by choices from binary menus, that is,

\[
\mu \gtrless \eta \iff \mu \in C(\{ \mu, \eta \}) .
\]  

(5)

The indifference relation \( \approx \) and the strict preference relation \( > \) are derived from \( \gtrless \) in the usual way.

Consider the following axioms on \( C(\cdot) \). The quantifiers ‘for all \( \mu, \eta \in \Delta, x, y \in Z, c, c', c'' \in C, \) and \( \alpha \in [0, 1] \)’ should be understood.

**Axiom 1 (WARP)**

If \( \mu, \eta \in x \cap y, \mu \in C(x) \) and \( \eta \in C(y) \), then \( \mu \in C(y) \).

This is the familiar Weak Axiom of Revealed Preference. It is a minimal consistency requirement on choices. Though WARP is a standard axiom in standard choice theory, we must inquire whether it is appropriate for a theory of choice under temptation. Consider the following example. Assuming a static set-up, let \( s \) represent a small stakes gamble, \( l \) a large stakes gamble, and \( n \) the option of not gambling. The agent normatively prefers \( n \) to \( s \) and \( s \) to \( l \), but finds \( l \) more tempting than \( s \), and \( s \) more tempting than \( n \). For simplicity, assume that \( C(\cdot) \) is single-valued. The following choices violate WARP:

\[
\{n\} = C(\{n, s\}) \text{ and } \{s\} = C(\{n, s, l\}) .
\]

Yet, the choices are not unreasonable. Since \( s \) is not so tempting, he is able to apply self-control and choose \( n \) out of \( \{n, s\} \). But when faced with \( \{n, s, l\} \), he is greatly tempted by \( l \). In order to compromise between his craving for a gamble and his normative preference for \( n \), he settles for \( s \).

This example is one where self-control is menu-dependent – the presence of \( l \) affects the agent’s ability to resist \( s \). Although WARP excludes such behavior, we note that it is still consistent with other aspects of decision-making under temptation and thus it may be an acceptable axiom. In the above example, WARP would not be violated if the choice from \( \{n, s, l\} \) was different from \( s \), that is, if either \( \{l\} = C(\{n, s, l\}) \) or \( \{n\} = C(\{n, s, l\}) \).

These capture the following stories: In the first case, \( l \) is so tempting that it cannot be resisted, and in the second case, \( l \) is not that much more tempting than \( s \), and thus can be resisted along with \( s \).

**Axiom 2 (Continuity)**

\( C(\cdot) \) is upper hemicontinuous.

Upper hemicontinuity of \( C(\cdot) \) is implied by choices being determined by the maximization of a continuous preference.\(^{10}\)

We impose upper hemicontinuity as an axiom, with the intention of establishing that choices are determined in such a way.

\(^{10}\)Formally, upper hemicontinuity implies that if \( \{x_n\} \) is a sequence of menus converging to \( x \), and \( \mu_n \in C(\{x_n\}) \) for each \( n \), then the sequence \( \{\mu_n\} \) has a limit point in \( C(\{x\}) \).
Axiom 3 (Independence) \[ \mu > \eta \implies \mu \alpha \nu > \eta \alpha \nu. \]

This is the familiar Independence axiom.

Axiom 4 (Separability) For all \( t \geq 0 \),

\[
\left( \frac{1}{2}(c, x) + \frac{1}{2}(c', x') \right)^+ \approx \left( \frac{1}{2}(c, x') + \frac{1}{2}(c', x) \right)^+.
\]

Separability states that when comparing two lotteries (delayed by \( t \geq 0 \) periods), the agent only cares about the marginal distributions on \( C \) and \( Z \) induced by the lotteries. That is, only marginals matter, and correlations between consumption and continuation menus do not affect the agent’s choices.

Axiom 5 (Indifference to Timing) For all \( t > 0 \),

\[ x^{+t} \alpha y^{+t} \approx (x \alpha y)^{+t}. \]

Under both rewards \( x^{+t} \alpha y^{+t} \) and \( (x \alpha y)^{+t} \), the agent faces \( x \) after \( t \) periods with probability \( \alpha \) and \( y \) after \( t \) periods with probability \( (1 - \alpha) \). However, under \( x^{+t} \alpha y^{+t} \), the uncertainty will be resolved today, whereas under \( (x \alpha y)^{+t} \), the uncertainty will be resolved after \( t \) periods. That is, the two rewards differ only in the timing of resolution of uncertainty. Indifference between the rewards corresponds to indifference to the timing of resolution of uncertainty.

Axiom 6 (Set-Betweenness) For all \( t > 0 \),

\[ x^{+t} \gtrless y^{+t} \implies x^{+t} \gtrless (x \cup y)^{+t} \gtrless y^{+t}. \]

Moreover, there exists \( x, y \) and \( t > 0 \) such that \( x^{+t} > (x \cup y)^{+t} > y^{+t} \).

Set-Betweenness expresses the idea that the agent anticipates experiencing temptation when choosing out of some menus in the future, and may exert self-control when making such choices. To illustrate, fix some delay \( t > 0 \), and consider \( \mu, \eta \in \Delta \) such that \( \{\mu\}^{+t} > \{\mu, \eta\}^{+t} \). The preference for commitment to \( \mu \) reveals that \( \eta \) is tempting. If we also observe the ranking \( \{\mu, \eta\}^{+t} \approx \{\eta\}^{+t} \), then it implies that the agent would choose the same item whether faced with \( \{\mu, \eta\} \) or \( \{\eta\} \). That is, choice from \( \{\mu, \eta\} \) is \( \eta \) and so, the agent succumbs to temptation. On the other hand, the ranking \( \{\mu, \eta\}^{+t} > \{\eta\}^{+t} \) would suggest that \( \mu \) is chosen from \( \{\mu, \eta\} \) and so, the agent resists temptation, that is, he exerts self-control.

GP would interpret the ranking \( \{\mu\}^{+t} \approx \{\mu, \eta\}^{+t} \gtrapprox \{\eta\}^{+t} \) as saying that no temptation is experienced in the menu \( \{\mu, \eta\} \). However, this ranking permits another interpretation as well: when future consumption is tempting, it is also consistent with an overwhelming temptation by \( \mu \). The story is that the reward \( \mu \) is so tempting that he prefers \( \{\mu, \eta\}^{+t} \).
over \( \{\eta\}^{+t} \), that is, he submits to the temptation of the menu \( \{\mu, \eta\} \) that contains the tempting reward \( \mu \). The indifference between \( \{\mu\}^{+t} \) and \( \{\mu, \eta\}^{+t} \) is another expression of the overwhelming temptation by \( \mu \) – he foresees that he will choose \( \mu \) in either menu, which is why he is indifferent between them.

The second part of Set-Betweenness is a nondegeneracy condition. It states that there is a menu \( x \cup y \) in which the agent anticipates experiencing temptation and exerting self-control.

**Axiom 7 (Sophistication)** For any \( t > 0 \) and \( \mu, \eta \) such that \( \{\mu\}^{+t} \not\approx \{\eta\}^{+t} \),
\[
\mu \gtrless \eta \iff \{\mu\}^{+t} \gtrless \{\mu, \eta\}^{+t} > \{\eta\}^{+t} \quad \text{or} \quad \{\eta\}^{+t} > \{\mu, \eta\}^{+t} \approx \{\mu\}^{+t}.
\]

As the name suggests, this axiom connects the agent’s expectation of his future choices with his actual choices.\(^{11}\) Each of the two right-hand-side pairs of rankings reveal that the agent expects himself to choose \( \mu \) if, after \( t \) periods, he faces the menu \( \{\mu, \eta\} \). Since \( C(\cdot) \) is time-invariant, the actual choice from \( \{\mu, \eta\} \) after \( t \) periods is given by \( C(\cdot) \). Thus, \( \mu \gtrless \eta \) says that after \( t \) periods, the agent’s actual choice from \( \{\mu, \eta\} \) would be \( \mu \). That is, the axiom says that the agent’s anticipated choice and actual choice coincide.

**Axiom 8 (Menus Can Tempt)** For all \( t > 0 \),
\[
x^{+t} > (x \cup y)^{+t} \iff \{(c, x)\}^{+t} > \{(c, x), (c, y)\}^{+t}.
\]

The axiom embodies the idea that the presence of tempting items in a menu makes the menu tempting. The preference for commitment exhibited by the left-hand-side ranking implies that the menu \( y \) contains tempting alternatives. The right-hand-side ranking states that \( y \) is a tempting menu. Thus, the axiom states that \( y \) is tempting if and only if \( y \) contains tempting alternatives.

**Axiom 9 (Reversal)** If \( \mu \gtrless \eta \) and \( \mu^{+T} \gtrless \eta^{+T} \) with at least one ranking strict, then \( \mu^{+t} > \eta^{+t} \) for all \( t > T \). Moreover, there exists some \( \mu, \eta \) and \( T \) such that \( \mu \gtrless \eta \) and \( \mu^{+T} > \eta^{+T} \).

This axiom imposes the structure of preference reversals on \( C(\cdot) \). That is, if pushing a pair of rewards into the future changes its ranking, then the reversed ranking is maintained for all subsequent delays in the rewards. Following the evidence on preference reversals, the axiom allows no more than one reversal for any pair of rewards. Post-preference reversal indifference is ruled out. The second part of the axiom requires that the agent exhibit a preference reversal for at least one pair of rewards.

---

\(^{11}\)It should be noted that this axiom is dynamic in that it relates choice across different times. All the other axioms are static as they deal with choice at a given time. Also, it should be clear that the axiom does not contradict preference reversals (Axiom 9 below). Preference reversals are a restriction on static choice.
As a simple consequence of Reversal we obtain a function \( \tau : \Delta \times \Delta \to \mathbb{N} \cup \{0\} \) such that for any pair of rewards \( \mu, \eta \), \( \tau(\mu, \eta) \) is the number of periods that \( \mu \) and \( \eta \) need to be delayed before a preference reversal is observed; if no reversal is observed, then \( \tau(\mu, \eta) = 0 \).\(^\dagger\) For instance, if \( \mu > \eta \), \( \mu^{t+1} \gtrless \eta^{t+1} \) and \( \mu^{t} < \eta^{t} \) for all \( t > 1 \), then \( \tau(\mu, \eta) = 2 \). For a precise definition of \( \tau(\cdot, \cdot) \), see Lemma 1 in Appendix B.

For any \( \mu, \eta \), the ranking \( \mu > \eta \) reflects either the underlying temptation ranking, or the normative ranking. If the preference for \( \mu \) is the result of an overwhelming temptation, then \( \mu > \eta \) reflects the temptation ranking. If no temptation is experienced, or if temptation does not overwhelm the agent, then \( \mu > \eta \) reflects the normative ranking. The next axiom requires us to be able to distinguish the two possibilities.

The distinction is achieved by recognizing the information contained in the function \( \tau(\cdot, \cdot) \). For any \( \mu, \eta \), say that the ranking \( \mu > \eta \) is normative if \( \tau(\mu, \eta) = 0 \) and \( \tau \) is continuous at \( (\mu, \eta) \).

To understand this definition, recall that preference reversals occur when the choice between two rewards is overwhelmed by temptation, and when pushing the rewards into the future makes it possible to resist temptation. If the choice between \( \mu \) and \( \eta \) is not overwhelmed by temptation, then pushing \( \mu \) and \( \eta \) into the future should not affect their ranking. That is, if the choice between \( \mu \) and \( \eta \) respects normative preference, we should observe no preference reversal, \( \tau(\mu, \eta) = 0 \).

However, the condition \( \tau(\mu, \eta) = 0 \) is not enough to conclude that the ranking \( \mu > \eta \) reflects normative preference: if there is normative indifference and \( \mu \) is more tempting than \( \eta \), then we can expect to observe \( \mu > \eta \) and \( \tau(\mu, \eta) = 0 \). Thus, in order to conclude that the ranking \( \mu > \eta \) reflects normative preference, we have to rule out the possibility of normative indifference and temptation non-indifference between \( \mu \) and \( \eta \).

When this possibility arises, one can expect to find in any neighborhood of \( (\mu, \eta) \) a pair of rewards \( (\mu^*, \eta^*) \) with the two properties. First, normative preference and temptation preference are both strict and in opposite directions; this is possible since normative indifference is a knife-edge case. Second, temptation is overwhelming relative to normative preference; this is possible since \( (\mu^*, \eta^*) \) being close to \( (\mu, \eta) \) implies that normative preference is close to indifference, whereas temptation preference is not. These conditions imply \( \tau(\mu^*, \eta^*) > 0 \) since overwhelming temptation can always be made resistible by delaying the rewards, thereby causing a reversal. Thus, when there is normative indifference and temptation non-indifference, in every neighborhood of \( (\mu, \eta) \) we can expect to find a pair of rewards for which a preference reversal is observed. This is ruled out by the continuity of \( \tau \) at \( (\mu, \eta) \).

Now we can state the final axiom.

\(^\dagger\)The set of nonnegative integers \( \mathbb{N} \cup \{0\} \) is endowed with the discrete topology.
**Axiom 10 (Commitment is Normative)** The ranking $x^t > (x \cup y)^t$ is normative. If $x^t > (x \cup y)^t > y^t$, then the ranking $(x \cup y)^t > y^t$ is also normative.

The ranking $x^t > (x \cup y)^t$ reveals not only that $x \cup y$ is a menu that contains temptations, but also that $x \cup y$ is not an overwhelmingly tempting menu; if it were, we would have observed $(x \cup y)^t \succeq x^t$. Hence a $t$ period delay makes the temptation by $x \cup y$ resistible. Consequently, choice between $x^t$ and $(x \cup y)^t$ is not overwhelmed by temptation, and thus, their ranking is normative. Observe also that the ranking $x^t > (x \cup y)^t$ reveals that $y$ contains temptations, and that a $t$ period delay in $y$ makes the temptation by $y$ resistible. It follows that the choice between $(x \cup y)^t$ and $y^t$ is not overwhelmed by temptation, and thus, their ranking is normative.

The main idea behind the axiom is that the choice to commit reflects normative preference and not temptation preference. But it is conceivable that commitment could be a property of temptation preference as well. One may be tempted to restrict one’s menu in order to reduce guilt.\(^{13}\) For instance, a pre-paid vacation package may be more tempting than a regular vacation package because the former makes it possible to indulge without experiencing guilt. Commitment is Normative rules out such considerations.

Finally, say that a binary relation $\succsim$ defined over $Z$ is an *FT preference* if, firstly, it has an FT representation (Section 2) and, secondly, it is *nondegenerate* in the sense that there exists $x, y \in Z$ such that

$$x \succ x \cup y \succ y.$$  

Recall from Section 2 that an FT preference $\succsim$ generates $C(\cdot)$ if,

$$C(\cdot) = C(\cdot, \succsim),$$

where $C(\cdot, \succsim)$ is defined by (2).

**Theorem 1** A choice correspondence $C(\cdot)$ satisfies Axioms 1-10 if and only if there exists an FT preference $\succsim$ that generates it. Furthermore, each such $C(\cdot)$ is generated by a unique FT preference $\succsim$.

Theorem 1 states that an agent whose choice correspondence satisfies Axioms 1-10 can be viewed as an FT agent, and conversely, the choices of an FT agent satisfy Axioms 1-10. In particular, if a choice correspondence $C(\cdot)$ satisfies Axioms 1-10, then there exist

---

\(^{13}\)The FT model may be extended to allow for guilt. Just as self-control cost is captured by $|V(\mu) - \max_{\eta \in Z} V(\eta)|$, guilt cost can be captured by $|U(\mu) - \max_{\eta \in Z} U(\eta)|$. That is, departing from normative preference leads to guilt, just as departing from temptation preference leads to a self-control cost. Thus, to incorporate guilt, replace $V(\cdot)$ in (1) by $\hat{V}(z) = \max_{\mu \in Z} \{V(\mu) + (U(\mu) - \max_{\eta \in Z} U(\eta))\}$, so that menus that contain more tempting items and less normatively preferable items are more tempting.

\(^{14}\)In terms of the representation, this is captured by the condition that $U(\cdot)$ is not an affine transformation of $V(\cdot)$, where $U(\cdot)$ and $V(\cdot)$ are as in the representation of FT preferences.
functions $U(\cdot)$ and $V(\cdot)$ as in the FT representation, such that for any menu $x$, the agent’s choice $C(x)$ from $x$ maximizes $U(\cdot) + V(\cdot)$.$^{15}$ The next section outlines the proof of the Theorem and shows that $U(\cdot)$ represents the agent’s normative preference.

Since choice maximizes $U(\cdot) + V(\cdot)$, and since $U(\cdot)$ represents normative preference, we interpret $V(\cdot)$ as a representation of the agent’s temptation preference. That is, it is the component of choice not ‘explained’ by normative preference. Thus, the functions $U(\cdot)$ and $V(\cdot)$ constitute a decomposition of the agent’s choices into its normative and temptation component. The Theorem assures us that this decomposition is unique, since two distinct FT agents cannot share the same choice correspondence.

We have provided foundations for the FT model that are an alternative to those in [18]. In order to check whether an agent is an FT agent, we do not need any data on how he would behave in the absence of temptation. All we need is to check that his actual choices, summarized by a choice correspondence, satisfy Axioms 1-10. In this sense, the foundations provided here are empirically more meaningful.

4 FT Preference and Normative Preference

Turn to the construction of the FT preference $\succsim$ in Theorem 1. The construction is based on the behavioral definition of normative preference presented in the Introduction. In this section we also show that this definition allows us to interpret $U(\cdot)$ as normative utility.

The FT preference $\succsim$ is derived from $C(\cdot)$ in 3 steps.

**Step 1:** Derive a set of preference relations $\{\succsim_t\}_{t=0}^{\infty}$ over $\Delta$, where for each $t \geq 0$ and $\mu, \eta \in \Delta$,

$$\mu \succsim_t \eta \iff \mu^{+t} \in C(\{\mu^{+t}, \eta^{+t}\}).$$

For any $\mu, \eta \in \Delta$, the preference $\succsim_t$ ranks $\mu$ and $\eta$ when both rewards are to be received $t$ periods later. Thus, $\{\succsim_t\}$ captures how the agent’s current period preference over $\Delta$ changes as $\Delta$ is pushed into the future, so to speak.

**Step 2:** Derive the normative preference $\succsim^*$ over $\Delta$.

By the Reversal axiom, $C(\cdot)$ exhibits preference reversals. As argued in the Introduction, preference reversals manifest the ability to resist delayed temptations. Thus, the ‘limit’ of the sequence $\{\succsim_t\}$ as $t$ goes to infinity describes how an agent ranks elements of $\Delta$ in the absence of temptation. Since choice respects normative preference when temptation is absent, the agent’s normative preference $\succsim^*$ over $\Delta$ may be identified with this limit.

To be formal about the definition of $\succsim^*$, say that a binary relation $B$ on $\Delta$ is nonempty if $\mu B \eta$ for some $\mu, \eta \in \Delta$. Adapting Hildenbrand [11], identify any nonempty continuous
binary relation on \( \Delta \) with its graph, a nonempty compact subset of \( \Delta \times \Delta \). Thus, the space of nonempty continuous preferences on \( \Delta \) can be identified with \( \mathcal{P} = \mathcal{K}(\Delta \times \Delta) \), the space of nonempty compact subsets of \( \Delta \times \Delta \) endowed with the Hausdorff metric topology. See Appendix A for details about the topology. Think of \( \{ \tilde{\succ}_t \} \) as a sequence in \( \mathcal{P} \).

**Definition 1**  

*The normative preference \( \succ^* \) over \( \Delta \) is the limit of the sequence \( \{ \tilde{\succ}_t \} \).*

The existence of normative preference obtains under WARP, Continuity and Reversal alone (see Section 7). Reversal requires that for every \( \mu, \eta \in \Delta \), the post-preference reversal preferences agree on the ranking of \( \mu \) and \( \eta \). This implies that for every \( \mu, \eta \in \Delta \), there exists \( T \) such that all preferences \( \tilde{\succ}_t, t \geq T \), agree on the ranking of \( \mu \) and \( \eta \). In a sense, the difference between \( \tilde{\succ}_t \) and \( \tilde{\succ}_{t+1} \) decreases as \( t \) grows, since \( \tilde{\succ}_t \) and \( \tilde{\succ}_{t+1} \) agree on the ranking of more and more pairs of rewards. This provides some intuition for why normative preference exists.

**Step 3:**  

*Derive \( \succ \) over \( Z \).*

The preference \( \succ \) in Theorem 1 is meant to capture the agent’s ranking of menus in the absence of temptation. Since choice in the absence of temptation respects normative preference, the desired preference \( \succ \) is in fact the agent’s normative preference over menus. That is, \( \succ \) is simply the ranking over \( Z \) induced by \( \succ^* : \)

\[ x \succ y \iff x^{+1} \succ^* y^{+1}, \]

for all \( x, y \in Z \). Thus, we say \( x \succ y \) if and only if normative preference ranks the \( x^{+1} \) higher than \( y^{+1} \), where \( x^{+1} \) and \( y^{+1} \) are degenerate lotteries in \( \Delta \) that give some common immediate consumption \( c \), and the respective menus \( x \) and \( y \) in the next period.

Theorem 1 establishes that under Axioms 1-10 imposed on \( \mathcal{C}(\cdot) \), the preference \( \succ \) is a well-defined FT preference. The latter is verified by checking that \( \succ \) satisfies the axioms in [18]. Commitment is Normative and Sophistication play key roles in establishing that \( \succ \) generates \( \mathcal{C}(\cdot) \). To see this, recall from the discussion following Set Betweenness that a ranking of menus embodies the choices an agent anticipates making from menus. That is, a ranking of menus generates an ‘anticipated choice correspondence’. Like \( \succ^* \), each \( \tilde{\succ}_t \) induces a ranking over \( Z \), and thus generates a choice correspondence.\(^{16}\) Sophistication ensures that each \( \tilde{\succ}_t \) generates the same choice correspondence, namely, \( \mathcal{C}(\cdot) \). Commitment is Normative ensures that \( \succ \) generates the same choice correspondence as each \( \tilde{\succ}_t \). The assertion follows.

This completes the derivation of \( \succ \). Now turn to the justification of our interpretation of the function \( U \) as normative utility.

\(^{16}\)For instance, suppose \( \{ \mu, \eta \}^{+1} \succ_t \{ \eta \}^{+1} \). By definition of \( \tilde{\succ}_t \), this is equivalent to \( \{ \mu, \eta \}^{+(t+1)} \succ \{ \eta \}^{+(t+1)} \). This ranking contains the following information: if the agent faces the menu \( \{ \mu, \eta \} \) after \( t + 1 \) periods, he anticipates choosing \( \eta \) from it.
For any FT preference $\succeq$, a normative preference $\succeq^*$ is said to be elicited from $C(\cdot, \succeq)$ if it is derived as follows: First, a choice correspondence $C(\cdot, \succeq)$ is derived from $\succeq$ by defining it as in (2). Then, a set of preference relations $\{\succeq^t\}_{t=0}^{\infty}$ defined over $\Delta$ is obtained from $C(\cdot, \succeq)$ as in Step 1 above, and finally, $\succeq^*$ is defined as in Step 2.

Call $(U, V)$ a representation of the FT preference $\succeq$ if $U$ and $V$ are as in the FT functional form. Theorem 2 establishes that $U$ represents the normative preference $\succeq^*$ elicited from $C(\cdot, \succeq)$.

**Theorem 2** If $(U, V)$ represents an FT preference $\succeq$, and $\succeq^*$ is the normative preference elicited from $C(\cdot, \succeq)$, then $U$ represents $\succeq^*$.

Put differently, if we start with a choice correspondence $C(\cdot)$ and derive a normative preference $\succeq^*$ from it, and if this choice correspondence is generated by an FT preference $\succeq$ with representation $(U, V)$, then $U$ represents $\succeq^*$. This justifies our referring to $U$ as normative utility. However, Theorem 2 is of interest also for other reasons. Recall from Section 2 that for any $\mu$, $U(\mu) = W(\{\mu\})$, where $W$ is defined by (1). Thus, $U$ represents the agent’s period 0 preference under commitment: for any $\mu, \eta$,

$$U(\mu) \geq U(\eta) \iff \{\mu\} \succeq \{\eta\}.$$  

GP [8] informally interpret $U$ as capturing the agent’s view of his long-run self-interest, that is, they identify normative preference with commitment preference. Theorem 2 provides formal justification for such a definition of normative, albeit in the context of the FT model.

Theorem 2 also establishes that our definition of normative preference is, in some sense, dual to GP’s. The normative preference that would be derived by applying GP’s definition in a (special) period 0 is the same as that derived by applying our definition in any period $t > 0$. The discussion in Sections 1 and 2 about the non-observability of the period 0 preference $\succeq$ implies that GP’s definition of normative preference may be based on unobservables. Theorem 2 tells us that ours is a dual definition in terms of observables.

### 5 A Model of Current Temptation

Having modeled an agent who is tempted by future consumption, we now model one who is tempted only by immediate consumption. The primitive is a choice correspondence $C(\cdot)$ as before. Ten axioms are impose on $C(\cdot)$. They coincide with the FT axioms, except that Menus Can Tempt and Reversal are replaced with the following two axioms.

As noted in Section 3, Menus Can Tempt captures temptation by future consumption. Observe that the axiom permits,

$$\{(c, x)\}^{t+t} > \{(c, x), (c, y)\}^{t+t}.$$
for some \(x, y\) and \(t\). Such a ranking expresses a demand for delayed commitment: Under \(\{(c, x), (c, y)\}^{t+1}\), the agent has the opportunity to decide after \(t\) periods whether to face \(x\) in the \((t + 1)\)th period or \(y\). Under \(\{(c, x)\}^{t+1}\), the agent is committed to facing \(x\) in the \((t + 1)\)th period. Therefore, the above ranking indicates the agent’s desire to commit to a future menu in advance. Menus Can Tempt reveals the reason for this: since \(y\) contains tempting alternatives, it is a tempting menu, and thus he would rather avoid having to choose between \(x\) and \(y\). The following axiom rules out a demand for delayed commitment.

**Axiom 8' (Menus Do Not Tempt)** For all \(c, x, y\) and \(t > 0\),

\[\{(c, x), (c, y)\}^{t+1} \succeq \{(c, x)\}^{t+1}.\]

That is, while agents who are tempted by future consumption benefit from restricting their choice of continuation menus, agents who are not tempted by future consumption have no motivation to do the same. Continuation menus do not tempt them – continuation menus contain future consumption items, and such items do not tempt the agent today. Hence, they find no benefit in restricting the choice of continuation menus available to them in the future.

**Axiom 9' (Preferences Reverse Tomorrow)** For all \(x, y\) and \(t > 0\),

\[x^{t+1} \succeq y^{t+1} \iff x^{t} \succeq y^{t}.\]

The axiom strengthens (the first part of) Reversal by imposing a stationarity property on the ranking of any pair of delayed alternatives.\(^{17}\) Thus, if \(x^{t+1} \succeq y^{t+1}\), then no subsequent delay in \(x\) and \(y\) leads to a preference reversal. Stated differently, the axioms implies \(\tau(\mu, \eta) \leq 1\) for all \(\mu, \eta\), where \(\tau\) captures the time of a reversal, as before. This says that delaying a pair of alternatives by a single period suffices to induce a reversal, and if no reversal is observed, no subsequent delay will induce one. This is to be expected when only immediate consumption is tempting. To see why, suppose \(\mu > \eta\), where \(\mu\) is overwhelmingly tempting. When both \(\mu\) and \(\eta\) are delayed by one period, they become future rewards, and since the agent is tempted only by immediate rewards, he experiences no temptation by \(\mu^{t+1}\) and so reverses his preferences: \(\mu^{t+1} < \eta^{t+1}\).

Turn to the representation. Let \(\succ_{CT}\) be a binary relation over \(Z\) that is represented by some function \(W^{CT} : Z \rightarrow R\). Say that \(\succ_{CT}\) is a Current Temptation (CT) preference if it is nondegenerate (in the sense of Section 3), and there exists \(\delta \in (0, 1), \gamma \geq 0\) and continuous functions \(u, v : C \rightarrow R\) such that for all \(x \in Z\),

\[W^{CT}(x) = \max_{\mu \in x}\{U(\mu) + V(\mu) - \max_{\eta \in x} V(\eta)\},\]

\(^{17}\)Given Preferences Reverse Tomorrow, Axioms 5-7 and 8' can be weakened by restricting their statements to hold only for \(t = 1\), and Axiom 10 can in fact be dropped altogether. However, in order to facilitate comparison with the FT model, we keep the stronger statements.
where \( U(\mu) = \int_{C \times Z} (u(c) + \delta W^{CT}(y)) \, d\mu(c, y) \)
\( V(\mu) = \int_{C \times Z} (v(c) + \gamma W^{CT}(y)) \, d\mu(c, y) \).

CT preferences differ from FT preferences in two ways. First, \( \gamma \) may be zero here, yielding GP’s Dynamic Self-Control preferences where current consumption is the only source of temptation utility. Second, if \( \gamma > 0 \), then the temptation utility of a continuation menu coincides with the normative utility of the menu, as opposed to the FT model where it coincides with the maximum temptation utility achievable in the menu. Note that for CT preferences, if \( \gamma > 0 \), then normative and temptation preferences never disagree when it comes to choosing between continuation menus, and hence, continuation menus do not tempt. In particular, future consumption does not tempt.

Call \( (U, V) \) a representation of the CT preference \( \succsim^{CT} \), where \( U \) and \( V \) have the above functional forms. As before, if \( (U, V) \) represents \( \succsim^{CT} \), say that \( \succsim^{CT} \) generates \( C(\cdot) \) if
\[
C(x) = \arg \max_{\mu \in x} \{U(\mu) + V(\mu)\}.
\]

Theorem 3 is the main result of this section.

**Theorem 3** A choice correspondence \( C(\cdot) \) satisfies Axioms 1-7,8,9 and 10 if and only if there exists a CT preference \( \succsim^{CT} \) that generates \( C(\cdot) \). Furthermore, each such \( C(\cdot) \) is generated by a unique CT preference \( \succsim^{CT} \).

CT preferences reduce to DSC preferences when \( \gamma = 0 \), and to the model of preferences studied by Krussel, Kuruscu and Smith [14] when \( \gamma > 0 \). Theorem 3 tells us that these are models of agents who are tempted only by immediate consumption, and furthermore, in our set-up, these are the only models of this type.

For completeness, we show that the result in Theorem 2 holds for CT preferences as well. The definition of ‘normative preference \( \succsim^{*} \) is elicited from \( C(\cdot, \succsim^{CT}) \)’ is analogous to the counterpart in the previous section.

**Theorem 4** If \( (U, V) \) represents a CT preference \( \succsim^{CT} \), and \( \succsim^{*} \) is the normative preference elicited from \( C(\cdot, \succsim^{CT}) \), then \( U \) is a representation of \( \succsim^{*} \).

### 6 Evidence

The foundations of the FT and CT model reveal that CT agents differ from FT agents when it comes to satisfying Axioms 8 and 9, that is, the two types of agents differ in
their demand for delayed commitment and their $\tau$ functions.\textsuperscript{18} We use this information to provide evidence of temptation by future consumption.

**Demand for Delayed Commitment:** We’ve seen that CT agents do not exhibit a demand for delayed commitment, while FT agents may. The results of Benartzi and Thaler\textsuperscript{[5]} may be interpreted as evidence supporting a demand for delayed commitment. The authors introduce a saving-enhancement plan, called the ‘Save More Tomorrow\textsuperscript{TM} (SMT) plan’. Subjects in a firm are given the opportunity to commit in advance to allocating a portion of their future salary increases towards a defined-contributions plan.\textsuperscript{19} Opting for this plan implies a preference for delayed commitment – subjects would rather not leave the allocation decision for the time when they learn of their salary increase. In particular, they would rather not leave the decision for the time between learning of the increased salary and receiving the increased salary. At such a time, the allocation decision does not affect current consumption, but rather next month’s budget set (menu), and a preference for avoiding making a decision at such a time is consistent with a temptation to under-allocate funds for retirement. The authors implemented the SMT plan in several firms, and found a significant demand for it. Across the implementations, the percentage of employees that opted for the plan ranged between 27\% (216 of 816) and 78\% (162 of 207).

One feature of the SMT plan that is problematic for our purposes is that allocation decisions made by participants are not binding. Thus, SMT falls short of providing commitment. However, the authors point out that only a small proportion of subjects drop out of the plan, and possible reasons for this include inertia, procrastination, etc. If subjects are aware that such psychological factors would stop them from dropping out of the plan in the future, then participating in the plan does serve as a means of commitment.

**Time of Preference Reversals:** A vast amount of experimental evidence in psychology lends support to the so-called Matching Law, which provides a parsimonious summary of subjects’ preferences over delayed rewards (see Ainslie\textsuperscript{[2]} for an overview of the literature). In what follows, we restrict attention to Mazur’s version of the Matching Law\textsuperscript{[17]}, which has received considerable attention.\textsuperscript{20}

Let $s^{+0}$ denote a small immediate reward $s$, and let $l^{+d}$ denote a large reward $l$ available with a delay of $d$ periods. If a subject chooses $s^{+0}$ over $l^{+d}$, then according to the Matching Law, in order to induce a preference reversal, both rewards must be delayed by $\hat{\tau}(s^{+0}, l^{+d})$ periods, where

$$\hat{\tau}(s^{+0}, l^{+d}) = \frac{s(1 + kd) - l}{k(l - s)},$$

\textsuperscript{18}Of course, these two differences do not exhaust all the behavioral differences between the two models.

\textsuperscript{19}A defined-contributions plan provides a vehicle for committing to saving for retirement.

\textsuperscript{20}According to Mazur’s formulation, a subject’s ‘present value’ for a reward $m$ delayed by $d$ periods is given by $\frac{m}{1 + kd}$, where $k$ is a measure of the subject’s sensitivity to delay.
and where $k$ is a parameter that captures the subject’s ‘sensitivity’ to delay. Observe that $\hat{\tau}(s^0, l^d)$ is increasing in $s$ and $d$, and decreasing in $l$. This is intuitive. The greater the value of $s$ or $d$, or the smaller the value of $l$, the more tempting the small reward is, and so the greater the number of periods of delay required in order to induce a preference reversal.

How does the $\tau$ function of CT and FT agents compare to the $\hat{\tau}$ function above? Suppose a preference reversal is observed for a pair of rewards $(\mu, \eta)$.\textsuperscript{21} Then CT agents exhibit,

$$\tau_{CT}(\mu, \eta) = 1,$$

whereas FT agents exhibit,\textsuperscript{22}

$$\tau_{FT}(\mu, \eta) = \frac{\ln U(\mu) - U(\eta)}{\ln \frac{\Delta}{\delta}},$$

where $U(\cdot)$ and $V(\cdot)$ are as in the FT representation. That is, for CT agents, a single period delay suffices to induce a reversal whereas for FT agents, the required delay depends on the rewards and the agent’s discount factors.

First, we inquire whether the $\tau_{FT}$ function shares the same properties as $\hat{\tau}$. To permit a comparison, identify $s^0$ with the consumption stream in $\Delta$ that gives immediate consumption $s$ and some fixed consumption $c$ for all future periods. Similarly, identify $l^d$ with the consumption stream that gives $l$ after $d$ periods and fixed consumption $c$ in all other periods. For expositional simplicity, suppose $u(c) = v(c) = 0$. If $s^0 \succeq l^d$ and a preference reversal is observed, then

$$\tau_{FT}(s^0, l^d) = \frac{\ln \delta^{d} u(l) - u(s)}{\ln \frac{\Delta}{\delta}},$$

If $u(\cdot)$ and $v(\cdot)$ are strictly increasing functions, then the $\tau_{FT}$ function makes the same qualitative predictions as $\hat{\tau}$, that is, $\tau_{FT}$ is increasing in $s$ and $d$ and decreasing in $l$. Furthermore, the role of $k$ in $\hat{\tau}$ is played by $\frac{\Delta}{\delta}$ in $\tau_{FT}$. Presumably, the ‘sensitivity’ to delay corresponds to how fast temptation utility is discounted (relative to normative utility) when rewards are pushed into the future.

Thus, $\tau_{FT}$ shares the same features as $\hat{\tau}$. This supports the idea that agents are tempted by future consumption. On the other hand, the $\tau_{CT}$ function is not rich, and thus does not capture different features of $\hat{\tau}$.

\textsuperscript{21}The necessary and sufficient condition for a preference reversal for $(\mu, \eta)$ is that the agent be overwhelmed by temptation when choosing between the rewards. This is equivalent to the condition $\frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)} \in (0, 1]$, where $U(\cdot)$ and $V(\cdot)$ are as in the CT or FT representation, and where without loss of generality, $U(\mu) \geq U(\eta)$.

\textsuperscript{22}This ignores the fact that time is discrete. The more accurate expression is given in Lemma 10.
We cannot conclude from this, however, that the evidence on preference reversals rejects the CT model. The CT model may be viewed as one where the length of a single period is ‘large’. In such a case, the delay by a single period suffices to induce a preference reversal. But how long must a period be before the CT model is consistent with the evidence on preference reversals? Experiments reveal that a period may have to be longer than a year. For instance, in Kirby and Herrnstein [12, Experiment 3], only a minimum delay of about 3 years would induce a reversal for all subjects.

Whether defining a period to be this long is appropriate or not depends on the particular application one has in mind. However, adjusting the length of a period is not without cost. First, a ‘long’ period decreases the strength of temptation by future consumption, and is tantamount to assuming it away. As a result, one also assumes away various implications of the FT model (including the behavioral and welfare implications discussed in the Introduction). Second, by defining a period to be, say, a year long, one implicitly assumes that the agent finds consumption to be (approximately) perfectly substitutable over a one year stretch of time. However, the evidence on preference reversals makes it very clear that consumption today is much different even from next week’s consumption. That is, most people would not consider even next week as part of what they regard as ‘the current period’.

7 Concluding Remarks

The table below summarizes three basic differences between GP’s and our approach to providing foundations for models of temptation. First, the starting point of GP’s approach is that the conflict between an agent’s normative and temptation preference is reflected in his ranking $\succsim$ of menus in a special period 0, whereas our starting point is that the conflict is reflected in his ranking of delayed rewards. Second, GP hypothesize that self-control problem lead to a preference for commitment, whereas we hypothesize that self-control problems lead to preference reversals. Third, GP’s behavioral definition of normative preference is in terms of the agent’s preferences over singleton menus, that is, his commitment preferences, while our’s is in terms of his ranking of sufficiently delayed alternatives. This paper demonstrates how our approach may be used to address questions that are not best answered by relying on observing a preference for commitment or on identifying a special period 0.
The centrepiece of our analysis is the behavioral definition of normative preference. The definition can be formulated in an abstract setting: Let $\Delta$ be any compact metric space that represents some set of rewards with generic elements $\mu, \eta, \nu$, and take as given a set of preference relations $\{\succeq_t\}_{t=0}^\infty$ on $\Delta$. For each $\mu, \eta$, the preference $\succeq_t$ captures how the agent ranks the rewards $\mu, \eta$ when they are to be received $t$ periods later. The preferences $\{\succeq_t\}_{t=0}^\infty$ express the agent’s view (of delayed rewards) at a fixed point in time. Normative preference $\succeq^\ast$ over $\Delta$ is defined as in Definition 1, that is,

$$\succeq^\ast = \lim_{t \to \infty} \succeq_t,$$

where the definition of the space of preferences and its topology is as in Section 4. Preference reversals suggest that the ranking of alternatives that are delayed by a sufficient number of periods reveals the ranking in the absence of temptation. Normative preference is identified with this temptation-free ranking, since temptation is what drives a wedge between choice and normative preference.

The existence of normative preference can be established under fairly mild conditions. Consider the following axioms on $\{\succeq_t\}$.

**Axiom A1 (Order*)**  \(\succeq_t\) is complete and transitive, for all $t$.

**Axiom A2 (Continuity*)** The sets $\{\eta: \mu \succeq_t \eta\}$ and $\{\eta: \eta \succeq_t \mu\}$ are closed, for all $t$.

**Axiom A3 (Reversal*)** If $\mu \succeq_0 \eta$ and $\mu \succeq_T \eta$ with at least one ranking strict, then $\mu \succ_t \eta$ for all $t > T$.

These axioms provide sufficient conditions for the existence of a normative preference $\succeq^\ast$.

**Theorem 5** If $\{\succeq_t\}$ satisfies A1-3, then $\succeq^\ast$ is well-defined, complete, transitive and continuous.
The intuition is similar to that given in the discussion of Step 2 in Section 4.

We close by noting that, in suitable settings, normative preference can also be interpreted as an agent’s ethics. For instance, in an Ultimatum game, a Proposer may have a sense of distributional fairness, which is captured by an ‘ethical preference’ over the set of possible divisions of a certain sum of money between himself and the Decider. Let this set be denoted by $\Delta$ and imagine that, despite his ethics, the Proposer is tempted to act selfishly. Due to this temptation, his ethics may not be reflected in his choices between elements of $\Delta$. But, to the extent that distant temptations are easier to resist, his ethics may be revealed by how he ranks these elements when the outcome of the game is to be enforced in the future. Thus, our definition of normative preference can serve as definition of ethical preferences. In particular, ethical preferences can be given behavioral foundations.

## Appendix: Topology on $\mathcal{P}$

Let $(\Delta \times \Delta, d)$ be a compact metric space and denote the space of nonempty compact subsets of $\Delta \times \Delta$ by $\mathcal{P}$. For any $A, B \in \mathcal{P}$, let $d(a, B) = \inf_{b \in B} d(a, b)$ and $d(b, A) = \inf_{a \in A} d(b, a)$. The Hausdorff metric $h_d$ induced by $d$ is defined by

$$h_d(A, B) = \max\{\sup_{a \in A} d(a, B), \sup_{b \in B} d(b, A)\},$$

for all $A, B \in \mathcal{P}$. An $\varepsilon$-ball centered at $A$ is defined by

$$B(A, \varepsilon) = \{B : h_d(A, B) < \varepsilon\}.$$

The Hausdorff metric topology on $\mathcal{P}$ is the topology for which the collection of balls $\{B(A, \varepsilon)\}_{A \in \mathcal{P}, \varepsilon \in (0, \infty)}$ is a base.

View the set $\mathcal{P}$ as the space of nonempty and continuous binary relations on $\Delta$ by identifying any such binary relation $B$ on $\Delta$ with $\Gamma(B)$, the graph of $B$:

$$\Gamma(B) = \{(\mu, \eta) \in \Delta \times \Delta : \mu B \eta\}.$$ 

The relation $B$ is continuous if and only if $\Gamma(B)$ is closed. Since $\Delta$ (and hence $\Delta \times \Delta$) is compact metric, it follows that if $B$ is continuous, then $\Gamma(B) \subset \Delta \times \Delta$ is compact. By [4, Thm 3.71(3)], compactness of $\Delta \times \Delta$ implies that $P$ is compact. Also, under compactness of $\Delta \times \Delta$, $\Gamma(B)$ is the Hausdorff metric limit of a sequence $\{\Gamma(B_n)\} \subset \mathcal{P}$ if and only if $\Gamma(B)$ is the ‘closed limit’ of $\{\Gamma(B_n)\}$ [4, Thm 3.79]. To define the closed limit of a sequence $\{\Gamma(B_n)\}$, first define the topological limit superior $Ls\Gamma(B_n)$ and topological limit inferior $Li\Gamma(B_n)$ of the sequence:

$$Ls\Gamma(B_n) = \{a \in \Delta \times \Delta : \text{for every neighborhood } V \text{ of } a, \ V \cap \Gamma(B_n) \neq \phi \text{ for infinitely many } n\}$$

$$Li\Gamma(B_n) = \{a \in \Delta \times \Delta : \text{for every neighborhood } V \text{ of } a, \ V \cap \Gamma(B_n) \neq \phi \text{ for all but a finite number of } n\}.$$

The sequence $\{\Gamma(B_n)\}$ converges to a closed limit $\Gamma(B)$ if $\Gamma(B) = Ls\Gamma(B_n) = Li\Gamma(B_n)$.

26
Appendix: Normative Preference and Proof of Theorem 5

Section 7 defined normative preference $\succeq^*$ and presented Axioms A1-A3 imposed on $\{\preceq_t\}$. Consider also the following axiom. It requires $\Delta$ to be a mixture space.

Axiom A4 (Independence$^*$) $\mu \succ_t \eta \implies \mu \alpha \nu \succ_t \eta \alpha \nu$, for all $t$.

In this Appendix, we prove two results about normative preference that will be used to prove Theorem 1. The first is stated as Theorem 5 in Section 7.

Theorem B.1 If $\{\preceq_t\}$ satisfies A1-3, then $\succeq^*$ is well-defined, complete, transitive and continuous.

Theorem B.2 If $\Delta$ is a mixture space and $\{\preceq_t\}$ also satisfies A4, then $\succeq^*$ also satisfies independence.

B.1 Proof of Theorem B.1

The first Lemma defines a function $\tau : \Delta \times \Delta \to \mathbb{N} \cup \{0\}$ which captures the time at which a reversal takes place in $\{\mu, \eta\}$.

Lemma 1 For each $\mu, \eta$ there exists $\tau(\mu, \eta) \geq 0$ such that if $\tau(\mu, \eta) = 0$, then,

$\mu \not\succeq_0 \eta \iff \mu \succeq_{\tau(\mu, \eta)} \eta$, for all $t \geq 0$,

and if $\tau(\mu, \eta) > 0$, then,

(a) $\mu \succ_0 \eta \implies \mu \succ_t \eta$ for all $t < \tau(\mu, \eta) - 1$; $\mu \succeq_{\tau(\mu, \eta)} \eta$ for $t = \tau(\mu, \eta) - 1$; $\mu \prec_t \eta$ for all $t \geq \tau(\mu, \eta)$.

(b) $\mu \approx_0 \eta \implies \mu \approx_t \eta$ for all $t < \tau(\mu, \eta)$; $\mu \succ_t \eta$ for all $t \geq \tau(\mu, \eta)$ or $\mu \prec_t \eta$ for all $t \geq \tau(\mu, \eta)$.

Proof. Define $\tau : \Delta \times \Delta \to \mathbb{N} \cup \{0\}$ in the following way. Take any $\mu$ and $\eta$. If,

$[\mu \approx_0 \eta \implies \mu \approx_t \eta$ for all $t]$

and $[\mu \succ_0 \eta \implies \mu \succ_t \eta$ for all $t]$, then define $\tau(\mu, \eta) = 0$. If $\mu \not\succeq_0 \eta$ and there exists $T$ such that $\mu \prec_T \eta$, then define

$\tau(\mu, \eta) = \min\{t : \mu \prec_t \eta\}$.

By Reversal, $\mu \prec_t \eta$ for all $t \geq \tau(\mu, \eta)$. Consider two cases.

(a) $\mu >_0 \eta$.

Suppose by way of contradiction that $\mu \approx_t \eta$ for some $t < \tau(\mu, \eta) - 1$. Then by Reversal, $\mu <_t \eta$ for $t' = \tau(\mu, \eta) - 1$, a contradiction. That $\mu \not\succeq_t \eta$ for $t = \tau(\mu, \eta) - 1$ follows from the definition of $\tau(\mu, \eta)$.

(b) $\mu \approx_0 \eta$.

By definition of $\tau(\mu, \eta)$ and the earlier noted fact that $\mu \prec_t \eta$ for all $t \geq \tau(\mu, \eta)$, it must be that $\mu \not\succeq_t \eta$ for all $t < \tau(\mu, \eta)$. Suppose by way of contradiction that $\mu \succ_v \eta$ for any $t' < \tau(\mu, \eta)$. Then by Reversal, $\mu \succ_t \eta$ for all $t > t'$, contradicting the fact that $\mu \prec_t \eta$ for all $t \geq \tau(\mu, \eta)$. ■
Although the preference $\succeq^*$ over $\Delta$ is defined in Definition 1, we abuse notation and let $\succeq^*$ be defined by

$$
\mu \succeq^* \eta \iff \text{there exists a sequence } \{(\mu_n, \eta_n)\} \text{ that converges to } (\mu, \eta)
$$

and $\mu_n \succeq^* (\mu_n, \eta_n)$ for all $n$.

It will later be shown that $\succeq^*$ is a normative preference in the sense of Definition 1.

**Lemma 2** $\mu \succeq^*_\tau (\mu, \eta) \iff \mu \succeq^* \eta$.

**Proof.** This is implied directly by the definition of $\succeq^*$.

**Lemma 3** $\succeq^*$ satisfies order and continuity.

**Proof.** To establish completeness, suppose $\eta \not\succ^* \mu$. By Lemma 2, $\eta \not\succeq^*_\tau (\mu, \eta)$, and by completeness of $\succeq^*_\tau (\mu, \eta)$, $\mu \not\succeq^*_\tau (\mu, \eta)$, and so, again by Lemma 2, $\mu \not\succeq^* \eta$. To establish transitivity, suppose $\mu \succeq^* \eta \succeq^* \nu$. Then by definition of $\succeq^*$, there exist sequences $\{(\mu_n, \eta_n)\}, \{(\eta_n, \nu_n)\}$ such that $\mu_n \rightarrow (\mu, \eta)$ and $\eta_n \rightarrow (\eta, \nu)$, and for each $n$, $\mu_n \succeq^*_\tau (\mu_n, \eta_n)$ and $\eta_n \succeq^*_\tau (\eta_n, \nu_n)$. By Lemma 1, $\mu_n \succeq^*_\tau \eta_n \succeq^*_\tau \nu_n$ for all $t \geq T = \max \{\tau(\mu_n, \eta_n), \tau(\eta_n, \nu_n)\}$. By transitivity of $\succeq^*_\tau$, $\mu_n \succeq^*_\tau \nu_n$ for all $t \geq T$, implying that for each $n$, $\mu_n \succeq^*_\tau (\mu_n, \nu_n)$. It follows that $\{(\mu_n, \nu_n)\}$ is a sequence that converges to $(\mu, \nu)$ and $\mu_n \succeq^*_\tau (\mu_n, \nu_n)$ which means $\mu \succeq^* \nu$, thus establishing transitivity of $\succeq^*$.

To establish continuity, we show that $\{\eta : \eta \succeq^* \mu\}$ is closed; the other case holds by an analogous argument. Take a sequence $\{\nu_n\}$ such that $\nu_n \succeq^* \mu$ for all $n$ and $\nu_n \rightarrow \nu$. Also, take a sequence $\{V_i\}$ where each $V_i \subset \Delta \times \Delta$ is a ball of radius $2^{-i}$ that contains $(\nu, \mu)$. Because $\nu_n \rightarrow \nu$, for every $i$ there exists $n$ such that $(\nu_n, \mu) \in V_i$. Furthermore, $\nu_n \succeq^* \mu$ and the definition of $\succeq^*$ imply the existence of a sequence $\{(\nu'_m, \mu'_m)\}$ such that $(\nu'_m, \mu'_m) \rightarrow (\nu, \mu)$ and $\nu'_m \succeq^*_\tau (\nu'_m, \mu'_m) \mu'_m$ for all $m$. Since $V_i$ is also a neighborhood of $(\nu_n, \mu)$, there exists $m$ such that $(\nu'_m, \mu'_m) \in V_i$. Define

$$(\nu'_m, \mu'_m) \equiv (\nu'_m, \mu'_m),$$

and note that $\nu'_m \succeq^*_\tau (\nu'_m, \mu'_m) \mu'_m$. Furthermore, by construction, $(\nu'_m, \mu'_m) \rightarrow (\nu, \mu)$ as $i \rightarrow \infty$, and so we are done.

The set $\Omega$ of points in $\Delta \times \Delta$ on which $\tau$ is upper semicontinuous will be important. Formally,

$$\Omega = \{(\mu, \eta) : (\mu_n, \eta_n) \rightarrow (\mu, \eta) \Rightarrow \limsup_{n \rightarrow \infty} \tau(\mu_n, \eta_n) \leq \tau(\mu, \eta)\}.$$

Lemma 4 provides a useful characterization of $\succeq^*$.

**Lemma 4** $\mu \prec^* \eta \iff [\mu \succ^*_\tau (\mu, \eta) \text{ and } (\mu, \eta) \in \Omega]$.

**Proof.** $\iff$: Take $\mu$ and $\eta$ such that $\mu \succ^*_\tau (\mu, \eta)$ and $(\mu, \eta) \in \Omega$. Lemma 1 implies $\mu \succ^*_\tau (\mu, \eta) + 1$. Since $\succeq^*_\tau (\mu, \eta) + 1$ is continuous, for every sequence $\{(\mu_n, \eta_n)\}$ that converges to $(\mu, \eta)$, there exists $N$ such that $\mu_n \succ^*_\tau (\mu, \eta) + 1$ for all $n \geq N$. 

28
By hypothesis, \( \limsup_{n \to \infty} \tau(\mu_n, \eta_n) \leq \tau(\mu, \eta) \). Therefore, there exists \( N' \) such that
\[
\tau(\mu, \eta) + 1 > \tau(\mu_n, \eta_n), \text{ for all } n \geq N'.
\]

It follows by definition of \( \tau(\mu_n, \eta_n) \) that
\[
\mu_n > \tau(\mu_n, \eta_n), \text{ for all } n \geq \max\{N, N'\}.
\]

This establishes that for any sequence \( \{(\mu_n, \eta_n)\} \) that converges to \( (\mu, \eta) \), there exists \( M \) such that \( \mu_m > \tau(\mu_n, \eta_n) \eta_n \) for all \( n \geq M \). In particular, there is no sequence \( \{(\mu_n, \eta_n)\} \) that converges to \( (\mu, \eta) \) such that \( \eta_n \succeq \tau(\mu_n, \eta_n) \mu_n \) for all \( n \). Thus \( \eta \npreceq \mu \), as desired.

\( \implies \): Take \( \mu, \eta \) such that \( \mu > \tau(\mu, \eta) \eta \).

Lemma 2 yields
\[
\mu > \tau(\mu, \eta) \eta
\]  
thus establishing the first assertion in the implication. To establish the second assertion, take any sequence \( \{(\mu_n, \eta_n)\} \) that converges to \( (\mu, \eta) \). Since \( \mu > \tau(\mu, \eta) \eta \) and since \( \succeq \tau(\mu, \eta) \) is continuous (Lemma 3), there exists \( N \) such that
\[
\mu_n > \tau(\mu_n, \eta_n) \eta_n, \text{ for all } n \geq N.
\]  
By Lemma 2,
\[
\mu_n > \tau(\mu_n, \eta_n) \eta_n, \text{ for all } n \geq N.
\]  
Without loss of generality, let \( N = 1 \). Suppose by way of contradiction that
\[
\limsup_{n \to \infty} \tau(\mu_n, \eta_n) > \tau(\mu, \eta).
\]

Then, there exists a subsequence \( \{(\mu_{n(m)}, \eta_{n(m)})\} \subset \{(\mu_n, \eta_n)\} \) where for all \( m \),
\[
\tau(\mu_{n(m)}, \eta_{n(m)}) > \tau(\mu, \eta).
\]  
By construction, \( \mu_{n(m)} > \tau(\mu_{n(m)}, \eta_{n(m)}) \eta_{n(m)} \) for all \( m \). Thus, by Lemma 1 and (8),
\[
\eta_{n(m)} \succeq \tau(\mu, \eta) \mu_{n(m)}, \text{ for all } m.
\]

However, since \( \succeq \tau(\mu, \eta) \) is continuous and \( (\mu_{n(m)}, \eta_{n(m)}) \to (\mu, \eta) \), we have \( \eta \succeq \tau(\mu, \eta) \mu \), contradicting (6).

We now prove that \( \succeq \tau(\mu, \eta) \) is a normative preference. Since each \( \succeq \tau(\mu, \eta) \) is a continuous weak order, \( \{\Gamma(\succeq \tau(\mu, \eta))\} \) is a sequence in \( P(\Delta) \). Define:
\[
\Gamma(\succeq \tau(\mu, \eta)) = \{ (\mu, \eta) \in \Delta \times \Delta : \exists T \text{ such that } (\mu, \eta) \in \Gamma(\succeq \tau(\mu, \eta)) \text{ for all } t \geq T \},
\]
and note that \( \Gamma(\succeq \tau(\mu, \eta)) = \overline{\Gamma(\succeq \tau(\mu, \eta))} \).

**Lemma 5** \( \Gamma(\succeq \tau(\mu, \eta)) = \lim_{t \to \infty} \Gamma(\succeq \tau(\mu, \eta)) \).

**Proof.** To establish the existence of a closed limit, it suffices to show that \( Ls\Gamma(\succeq \tau(\mu, \eta)) \subset Li\Gamma(\succeq \tau(\mu, \eta)) \), since \( Li\Gamma(\succeq \tau(\mu, \eta)) \subset Ls\Gamma(\succeq \tau(\mu, \eta)) \) always holds.

Step 1: \( Ls\Gamma(\succeq \tau(\mu, \eta)) \subset \Gamma(\succeq \tau(\mu, \eta)) \).

\( ^{23} \Gamma(\succeq \tau(\mu, \eta)) \) is the graph of \( \succeq \tau(\mu, \eta) \), as defined in Appendix A.
Let \((\mu, \eta) \in Ls \Gamma(\overline{\Gamma}_t)\). By Lemma 1, there exists \(T^* < \infty\) such that either \((\mu, \eta) \in \Gamma(\overline{\Gamma}_t)\) for all \(t \geq T^*\) or \((\mu, \eta) \not\in \Gamma(\overline{\Gamma}_t)\) for all \(t \geq T^*\). If \((\mu, \eta) \in \Gamma(\overline{\Gamma}_t)\) for all \(t \geq T^*\), then \((\mu, \eta) \in \Gamma(\overline{\Gamma}_\tau) \subset \Gamma(\overline{\Gamma}^*)\) and we are done. Therefore, suppose \((\mu, \eta) \not\in \Gamma(\overline{\Gamma}_t)\) for all \(t \geq T^*\). By Lemma 4, it suffices to show \((\mu, \eta) \not\in \Omega\). So suppose by way of contradiction that \((\mu, \eta) \in \Omega\).

By definition (for instance, see [6, p 110] ), \((\mu, \eta) \in Ls \Gamma(\overline{\Gamma}_t)\) implies that there is a subsequence \(\{\Gamma(\overline{\Gamma}_{\tau(n)})(\mu, \eta)\}\) and a sequence \(\{(\mu_n, \eta_n)\}\) that converges to \((\mu, \eta)\) such that \((\mu_n, \eta_n) \in \Gamma(\overline{\Gamma}_{\tau(n)})\) for each \(n\). By assumption, \((\mu, \eta) \in \Omega\), and so,

\[
\limsup_{n \to \infty} \tau(\mu_n, \eta_n) \leq T^*.
\]

It follows that there exists \(M\) such that \(\tau(\mu_n, \eta_n) < T^* + 1\) for all \(n \geq M\). Without loss of generality, assume \(M = 1\). Also, since \(\Gamma(\overline{\Gamma}_{\tau(T^* + 1)})\) is closed and \((\mu, \eta) \not\in \Gamma(\overline{\Gamma}_{\tau(T^* + 1)})\), there exists \(N\) such that \((\mu_n, \eta_n) \not\in \Gamma(\overline{\Gamma}_{\tau(T^* + 1)})\) for all \(n \geq N\). By Lemma 1, \(\tau(\mu_n, \eta_n) < T^* + 1\) and \((\mu_n, \eta_n) \not\in \Gamma(\overline{\Gamma}_{\tau(T^* + 1)})\) for all \(n \geq N\) implies that for all \(n \geq N\) and \(t \geq T^* + 1\),

\[
(\mu_n, \eta_n) \not\in \Gamma(\overline{\Gamma}_t).
\]

However, by construction of the sequence \(\{(\mu_n, \eta_n)\}\), we have \((\mu_n, \eta_n) \in \Gamma(\overline{\Gamma}_{\tau(n)})\) for all \(n\). Let \(N'\) be such that \(t(N') \geq T^* + 1\). It follows that for all \(n \geq \max\{N, N'\}\),

\[
(\mu_n, \eta_n) \in \Gamma(\overline{\Gamma}_{\tau(n)}),
\]

and so there exists \(n \geq N\) and \(t \geq T^* + 1\) that contradicts (9).

**Step 2:** \(\Gamma(\overline{\Gamma}^*) \subset Ls \Gamma(\overline{\Gamma}_t)\).

First show that \(\Gamma(\overline{\Gamma}^*) \subset Ls \Gamma(\overline{\Gamma}_t)\). Observe that if \((\mu, \eta) \in \Gamma(\overline{\Gamma}_\tau)\), then there exists \(T < \infty\) such that \((\mu, \eta) \in \Gamma(\overline{\Gamma}_T)\) for all \(t \geq T\). Hence for every neighborhood \(V\) of \((\mu, \eta)\),

\[
V \cap \Gamma(\overline{\Gamma}_T) \neq \emptyset
\]

for all but a finite number of \(t\).

It follows that \((\mu, \eta) \in Ls \Gamma(\overline{\Gamma}_t)\), thus establishing that \(\Gamma(\overline{\Gamma}_\tau) \subset Ls \Gamma(\overline{\Gamma}_t)\), as desired. To complete the proof of Step 2, note that since \(Ls \Gamma(\overline{\Gamma}_t)\) is closed [4, Lemma 3.67], it follows that \(\overline{\Gamma(\overline{\Gamma}_\tau)} \subset Ls \Gamma(\overline{\Gamma}_t)\). But \(\overline{\Gamma(\overline{\Gamma}^*)} = \Gamma(\overline{\Gamma}_\tau)\). The assertion follows.

By Steps 1 and 2, \(Ls \Gamma(\overline{\Gamma}_t) \subset \Gamma(\overline{\Gamma}^*) \subset Ls \Gamma(\overline{\Gamma}_t)\). Hence,

\[
Ls \Gamma(\overline{\Gamma}_t) = Ls \Gamma(\overline{\Gamma}_t) = \Gamma(\overline{\Gamma}^*).
\]

This completes the proof. ■

**B.2 Proof of Theorem B.2**

This is proved in 5 steps.

**Step 1:** \(\mu >_t \eta \iff \mu \alpha \nu >_t \eta \alpha \nu\), for all \(t\).

Axioms A1, A2 and A4 together imply this stronger version of Independence*.

**Step 2:** \(\tau(\mu, \eta) = \tau(\mu \alpha \nu, \eta \alpha \nu)\) and \(\mu >_{\tau(\mu, \eta)} \eta \iff \mu \alpha \nu >_{\tau(\mu \alpha \nu, \eta \alpha \nu)} \eta \alpha \nu\).

This follows from Step 1.

**Step 3:** \((\mu, \eta) \not\in \Omega \implies (\mu \alpha \nu, \eta \alpha \nu) \not\in \Omega\).

If \(\{(\mu_n, \eta_n)\}\) is a sequence that converges to \((\mu, \eta)\) and

\[
\limsup_{n \to \infty} \tau(\mu_n, \eta_n) > \tau(\mu, \eta),
\]
then \( \{(\mu_n \alpha, \eta_n \alpha)\} \) is a sequence that converges to \((\mu \alpha, \eta \alpha)\) and, by the first assertion in Step 2, \[
\limsup_{n \to \infty} \tau(\mu_n \alpha, \eta_n \alpha) > \tau(\mu \alpha, \eta \alpha).
\]
Thus, \((\mu \alpha, \eta \alpha) \not\in \Omega \).

Step 4: \(\mu \sim^* \eta \implies \mu \alpha \sim^* \eta \alpha\).

Suppose \(\mu \sim^* \eta\). Then
\[
\mu \sim^* \eta \implies \mu \approx \tau(\mu, \eta) \eta \text{ or } (\mu, \eta) \not\in \Omega \text{ by Lemma 4}
\]
\[
\implies \mu \alpha \approx \tau(\mu \alpha, \eta \alpha) \eta \alpha \text{ or } (\mu \alpha, \eta \alpha) \not\in \Omega \text{ by Steps 2 and 3}
\]
\[
\implies \mu \alpha \sim^* \eta \alpha, \text{ as desired.}
\]

Step 5: \(\mu \succ^* \eta \implies \mu \alpha \succ^* \eta \alpha\).

By the main theorem in Herstein and Milnor [10], under order and continuity of \(\gg^*\), Step 4 implies the result.

C Appendix: Proof of Theorem 1 (Necessity)

Define the function \(\phi : \Delta(C \times Z) \to R\) by \(\phi(\mu) = U(\mu) + V(\mu)\) for all \(\mu \in \Delta\) and consider the correspondence \(\mathcal{C} : Z \rightharpoonup \Delta(C \times Z)\) defined by
\[
\mathcal{C}(x) = \arg \max_{\mu \in x} \phi(\mu).
\]
Clearly, \(\mathcal{C}(x) \subset x\). Since \(\phi\) is continuous, the Maximum Theorem [4, Thm 16.31] yields that \(\mathcal{C}\) is nonempty, compact-valued (in particular closed-valued) and upper hemicontinuous. Hence \(\phi\) generates a closed-valued choice correspondence \(\mathcal{C}\) that satisfies Axiom 2. That \(\mathcal{C}\) satisfies Axiom 1 can be checked easily. We need to show that \(\mathcal{C}\) satisfies Axioms 3-10. Proofs for Independence, Separability and Indifference to Timing are omitted. Let \(\gg_{t}^*\) be the the binary relation that is represented by \(\phi\). For all \(t \geq 1\), define \(\gg_{t}^*\) over \(Z\) by
\[
x \gg_{t}^* y \iff x^{+t} \gg_{0} y^{+t}.
\]

Lemma 6 \(\mathcal{C}\) satisfies Set-Betweenness.

Proof. Note that
\[
\phi(x^{+t}) = \delta^t W(x) + \gamma^t V(y) + \text{constant} = \delta^t \max_{\mu \in x} (U(\mu) + V(\mu) - \max_{\eta \in x} V(\eta)) + \gamma^t V(x) + \text{constant} = \delta^t \max_{\mu \in x} U(\mu) + V(\mu) - (1 - \frac{\gamma^t}{\delta^t}) \max_{\eta \in x} V(\eta) + \text{constant}.
\]
Hence, \(\gg_{t}^*\) is represented by
\[
\phi'(x) = \max_{\mu \in x} U(\mu) + V(\mu) - (1 - \frac{\gamma^t}{\delta^t}) \max_{\eta \in x} V(\eta).
\]

Defining \(V'(\mu) = (1 - \frac{\gamma^t}{\delta^t}) V(\mu)\) and \(U'(\mu) = U(\mu) + \frac{\gamma^t}{\delta^t} V(\mu)\), we have
\[
\phi'(x) = \max_{\mu \in x} U'(\mu) + V'(\mu) - \max_{\eta \in x} V'(\eta),
\]
and so $\phi'$ represents a Self-Control preference [7]. Therefore, $C$ satisfies Set-Betweenness.

**Lemma 7** $C$ satisfies Sophistication.

**Proof.** In the proof of the previous lemma, we showed that $\preceq_t$ is represented by

$$\phi'(x) = \max_{\mu \in x} U(\mu) + V(\mu) - (1 - \frac{\gamma_t}{\delta_t}) \max_{\eta \in x} V(\eta),$$

which can be re-written as

$$\phi'(x) = \max_{\mu \in x} U'(\mu) + V'(\mu) - \max_{\eta \in x} V'(\eta),$$

with $U' + V' = U + V$. By [18, Lemma 1], for any $\mu, \eta, t$ such that $\mu \not\approx_t \eta$,

$$\{\mu\} \preceq_t \{\mu, \eta\} \lor \{\eta\} \lor \{\mu, \eta\} \approx_t \{\mu\} \iff U(\mu) + V(\mu) \geq U(\eta) + V(\eta).$$

But $C(x) = \arg\max_{\mu \in x} (U(\mu) + V(\mu))$, and so, $U(\mu) + V(\mu) \geq U(\eta) + V(\eta)$ is equivalent to $\mu \in C(\{\mu, \eta\})$, which in turn is equivalent to $\mu \not\preceq \eta$. ■

**Lemma 8** $C$ satisfies Menus Can Tempt.

**Proof.** Note that $\preceq_t$ is represented by $W(\cdot) + \frac{\gamma_t}{\delta_t} V(\cdot)$. Therefore, it suffices to show that

$$W(z + \frac{\gamma_t}{\delta_t} \nabla(x)) > W(z + \frac{\gamma_t}{\delta_t} \nabla(x + y)) \implies W(\{z, x\}) + \frac{\gamma_t}{\delta_t} V(\{z, x\}) > W(\{z, x, y\}) + \frac{\gamma_t}{\delta_t} V(\{z, x, y\}).$$

Suppose $W(x) + \frac{\gamma_t}{\delta_t} V(x) > W(x + y) + \frac{\gamma_t}{\delta_t} V(x + y)$.

Step 1: $U(c, x) + \frac{\gamma_t}{\delta_t} V(c, x) > U(c, x + y) + \frac{\gamma_t}{\delta_t} V(c, x + y)$

By hypothesis, $W(x) + \frac{\gamma_t}{\delta_t} V(x) > W(x + y) + \frac{\gamma_t}{\delta_t} V(x + y)$. Then, $V(x + y) \geq V(x)$ implies

$$W(x) > W(x + y),$$

which in turn implies

$$V(x + y) = V(y) > V(x).$$

Given that $\frac{\gamma_t}{\delta_t} < 1$, it follows from the hypothesis that

$$W(x) + \frac{\gamma_t+1}{\delta_t+1} V(x) > W(x + y) + \frac{\gamma_t+1}{\delta_t+1} V(x + y).$$

But adding $u(c) + \frac{\gamma_t}{\delta_t} v(c)$ to both sides of (12) yields

$$U(c, x) + \frac{\gamma_t}{\delta_t} V(c, x) > U(c, x + y) + \frac{\gamma_t}{\delta_t} V(c, x + y).$$

Step 2: $U(c, x) + \frac{\gamma_t}{\delta_t} V(c, x) > U(c, y) + \frac{\gamma_t}{\delta_t} V(c, x + y)$

---

24The nondegeneracy part of Set-Betweenness follows from nondegeneracy of the FT preference $\preceq$ and Lemma 11 below.
By (18, Thm 3),

**Proof.**

There exists Lemma 9. All we need to show is that there exists \( \mu, \eta \) such that

\[
\text{Step 3: } \nabla(\{(c,x),(c,y)\}) = V(c,x) > V(c,y) > V(c,x)
\]

By (10), \( \nabla(y) > \nabla(x) \), which implies

\[
V(c,y) > V(c,x).
\]  \hspace{1cm} (13)

Furthermore, (11) and (13) imply

\[
\nabla(\{(c,x),(c,y)\}) = V(c,y) = V(c,x) \cup y.
\]  \hspace{1cm} (14)

**Step 4:** The result.

Consider two possibilities:

a) \( W(c,x) > W(\{(c,x),(c,y)\}) = W(c,y) \)

Then,

\[
W(\{(c,x),(c,y)\}) = U(c,y).
\]  \hspace{1cm} (15)

Now,

\[
W(\{(c,x)\}) + \frac{\gamma}{\delta} \nabla(\{(c,x)\}) \\
= U(c,x) + \frac{\gamma}{\delta} V(c,x) \\
> U(c,y) + \frac{\gamma}{\delta} V(c,x) \cup y \quad \text{by Step 2} \\
= W(\{(c,x),(c,y)\}) + \frac{\gamma}{\delta} \nabla(\{(c,x),(c,y)\}), \quad \text{by (15) and (14)}
\]
as desired.

b) \( W(\{(c,x)\}) > W(\{(c,x),(c,y)\}) > W(\{(c,y)\}) \).

Then, \( W(\{(c,x),(c,y)\}) = U(c,x) + (V(c,x) - V(c,y)) \). Hence,

\[
W(\{(c,x)\}) + \frac{\gamma}{\delta} \nabla(\{(c,x)\}) - \frac{\gamma}{\delta} \nabla(\{(c,x),(c,y)\}) \\
= U(c,x) + \frac{\gamma}{\delta} (V(c,x) - V(c,y)) \quad \text{by (14)} \\
> U(c,x) + (V(c,x) - V(c,y)) \quad \text{since } \frac{\gamma}{\delta} < 1 \text{ and (13).} \\
= W(\{(c,x),(c,y)\}).
\]

That is, \( W(\{(c,x)\}) + \frac{\gamma}{\delta} \nabla(\{(c,x)\}) > W(\{(c,x),(c,y)\}) + \frac{\gamma}{\delta} \nabla(\{(c,x),(c,y)\}), \) as desired \( \blacksquare \)

The next Lemma establishes that \( \mathcal{C} \) satisfies the second part of Reversal.

**Lemma 9** There exists \( \mu, \eta \) and \( T \) such that \( \mu \succsim \eta \) and \( \mu^{+T} \succ \eta^{+T} \).

**Proof.** By [18, Thm 3], \( \gamma < \delta \) implies that \( \succsim \) exhibits preference reversals, in the sense of [18, Section 4.1]. All we need to show is that there exists \( \mu, \eta \) such that

\[
\{\mu\} > \{\mu, \eta\} \sim \{\eta\}.
\]  \hspace{1cm} (16)

Since \( \succsim \) is nondegenerate, there exists \( x, y \) such that \( x \succ x \cup y \succ y \). This implies that \( U \) and \( V \) are nonconstant, and furthermore, \( U \) is not an affine transformation of \( V \) (see proof of [18, Thm 2]).
Thus, $U$ is not a positive affine transformation of $U + V$. It follows that there exist $\mu, \eta$ such that either $[U(\mu) > U(\eta)$ and $U(\mu) + V(\mu) \leq U(\eta) + V(\eta)]$, or $[U(\mu) \geq U(\eta)$ and $U(\mu) + V(\mu) < U(\eta) + V(\eta)]$. Except when $U(\mu) = U(\eta)$ and $U(\mu) + V(\mu) < U(\eta) + V(\eta)$, [18, Lemma 1(b)] implies (16) in each of these cases. However, if $U(\mu) = U(\eta)$ and $U(\mu) + V(\mu) < U(\eta) + V(\eta)$, nonconstancy of $U$ and linearity of $U$ and $U + V$ imply the existence of $\mu'$ and $\eta'$ close to $\mu$ and $\eta$, respectively, such that $U(\mu') > U(\eta')$ and $U(\mu') + V(\mu') < U(\eta') + V(\eta')$. Hence, we are still in one of the earlier cases. The assertion follows.

The proof of Lemma 10 establishes that $C$ satisfies the first part of Reversal.

**Lemma 10** Suppose $U(\mu) \geq U(\eta)$. Then,

$$
\frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)} \in (0, 1] \implies \tau(\mu, \eta) = \min\{k \in \mathbb{N} : k > \frac{\ln \frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)}}{\ln \frac{t}{\delta}}\} > 0
$$

$$
\frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)} \not\in (0, 1] \implies \tau(\mu, \eta) = 0.
$$

**Proof.** By the representation, for $t \geq 0$,

$$
\mu^{+t} \gtrless \eta^{+t} \iff U(\mu) + \frac{\gamma^t}{\delta} V(\mu) \geq U(\eta) + \frac{\gamma^t}{\delta} V(\eta).
$$

Suppose $\frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)} \in (0, 1]$. Since $U(\mu) \geq U(\eta)$, we have $U(\mu) > U(\eta)$ and $V(\mu) < V(\eta)$. Furthermore, $\frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)} \leq 1$ implies $U(\mu) + V(\mu) \leq U(\eta) + V(\eta)$. Since $\frac{\gamma}{\delta} < 1$ and $U(\mu) > U(\eta)$, there exists $\tau(\mu, \eta)$ such that

$$
\forall t < \tau(\mu, \eta), \ U(\mu) + \frac{\gamma^t}{\delta} V(\mu) \leq U(\eta) + \frac{\gamma^t}{\delta} V(\eta)
$$

$$
\forall t \geq \tau(\mu, \eta), \ U(\mu) + \frac{\gamma^t}{\delta} V(\mu) > U(\eta) + \frac{\gamma^t}{\delta} V(\eta). \tag{17}
$$

To find $\tau(\mu, \eta)$, first find the $t^*$ that solves

$$
U(\mu) + \frac{\gamma^t}{\delta} V(\mu) = U(\eta) + \frac{\gamma^t}{\delta} V(\eta).
$$

The solution is $t^* = \frac{\ln \frac{U(\mu) - U(\eta)}{V(\mu) - V(\eta)}}{\ln \frac{\gamma}{\delta}}$. Then $\tau(\mu, \eta)$ is the smallest integer greater than $\frac{\ln \frac{U(\mu) - U(\eta)}{V(\mu) - V(\eta)}}{\ln \frac{\gamma}{\delta}}$, that is,\(^{25}\)

$$
\tau(\mu, \eta) = \min\{k \in \mathbb{N} : k > \frac{\ln \frac{U(\mu) - U(\eta)}{V(\mu) - V(\eta)}}{\ln \frac{\gamma}{\delta}}\}.
$$

If $V(\eta) = V(\mu)$ or if $V(\eta) \not= V(\mu)$ and $\frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)} = 0$, then it is straightforward to establish that

$$
\mu \gtrless \eta \iff \mu^{+t} \gtrless \eta^{+t},
$$

\(^{25}\)Recall that, as defined in Lemma 1, the function $\tau$ gives the smallest time $t$ for which the post-reversal preference is strict.
and so \( \tau(\mu, \eta) = 0 \). Suppose \( V(\eta) \neq V(\mu) \) and \( \frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)} < 0 \). Then, \( V(\mu) > V(\eta) \). By hypothesis, \( U(\mu) \geq U(\eta) \), and so \( \frac{1}{2} < 1 \) implies \( \tau(\mu, \eta) = 0 \). Finally, suppose \( V(\eta) \neq V(\mu) \) and \( \frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)} > 1 \). Then \( V(\mu) < V(\eta) \) and \( U(\mu) + V(\mu) > U(\eta) + V(\eta) \). Again it follows that \( \tau(\mu, \eta) = 0 \). Hence, \( \tau(\mu, \eta) = 0 \) if \( \frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)} \notin (0, 1) \) and \( V(\eta) \neq V(\mu) \) or if \( V(\eta) = V(\mu) \). In particular, \( \tau(\mu, \eta) = 0 \) if \( \frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)} \notin (0, 1) \).

**Lemma 11** If \( U(\mu) > U(\eta) \) then \( \mu > \eta \) for all \( t \geq \tau(\mu, \eta) \). In particular, \( U(\mu) > U(\eta) \) and \( \mu > \eta \) imply \( \tau(\mu, \eta) = 0 \).

**Proof.** The first assertion is a corollary of Lemma 10. The second assertion follows from the first.

Define \( \Omega = \{(\mu, \eta) : \text{for any sequence } \{(\mu_n, \eta_n)\} \text{ that converges to } (\mu, \eta), \]
\[ \lim_{n \to \infty} \tau(\mu_n, \eta_n) \leq \tau(\mu, \eta) \}. \]

**Lemma 12** If \( U(\mu) \neq U(\eta) \) then \( (\mu, \eta) \in \Omega \).

**Proof.** Without loss of generality, \( U(\mu) > U(\eta) \). Take a sequence \( \{(\mu_n, \eta_n)\} \) that converges to \( (\mu, \eta) \). Consider the following possibilities:

(a) \( V(\mu) > V(\eta) \).

Then \( \frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)} > 1 \). Since \( \frac{U(\mu_n) - U(\eta_n)}{V(\eta_n) - V(\mu_n)} \to \frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)} \), there exists \( N \) such that \( \frac{U(\mu_n) - U(\eta_n)}{V(\eta_n) - V(\mu_n)} \geq 1 \) and \( V(\mu_n) > V(\eta_n) \) for all \( n \geq N \). Lemma 10 implies \( \tau(\mu_n, \eta_n) = 0 \) for all \( n \geq N \). Hence \( \limsup_{n \to \infty} \tau(\mu_n, \eta_n) = 0 \leq \tau(\mu, \eta) \), implying \( (\mu, \eta) \in \Omega \).

(b) \( V(\mu) < V(\eta) \).

Consider three possibilities.

(i) \( \frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)} \notin (0, 1) \)

Note that \( U(\mu) > U(\eta) \) and \( V(\mu) < V(\eta) \) implies \( \frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)} \neq 0 \). Hence, \( \frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)} \notin [0, 1] \). Now argue as in (a).

(ii) \( \frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)} \in (0, 1) \)

Then, there exists \( N \) such that for all \( n \geq N \), \( \frac{U(\mu_n) - U(\eta_n)}{V(\eta_n) - V(\mu_n)} \in (0, 1) \). Without loss of generality, let \( N = 1 \). By Lemma 10, \( \tau(\mu_n, \eta_n) = \min\{k \in N : k > \frac{\ln \left( \frac{U(\mu_n) - U(\eta_n)}{V(\eta_n) - V(\mu_n)} \right)}{\ln \frac{1}{2}} \} \) for each \( n \). Since \( \frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)} \to \frac{U(\mu_n) - U(\eta_n)}{V(\eta_n) - V(\mu_n)} \), we have \( \lim_{n \to \infty} \tau(\mu_n, \eta_n) = \tau(\mu, \eta) \), which establishes the result.

(iii) \( \frac{U(\mu) - U(\eta)}{V(\eta) - V(\mu)} = 1 \)

Note that there exists \( M \) such that \( V(\mu_n) < V(\eta_n) \) for all \( n \geq M \). Without loss of generality, \( M = 1 \). If there exists \( N \) such that \( \frac{U(\mu_n) - U(\eta_n)}{V(\eta_n) - V(\mu_n)} > 1 \) for all \( n \geq N \), then by Lemma 10, \( \tau(\mu_n, \eta_n) = 0 \leq \tau(\mu, \eta) \) for all \( n \geq N \) and so \( (\mu, \eta) \in \Omega \). If there is no such \( N \), then construct a subsequence \( \{(\mu_{n(m)}, \eta_{n(m)})\} \) by deleting all \( (\mu_n, \eta_n) \) in \( \{(\mu_n, \eta_n)\} \) such that \( \frac{U(\mu_n) - U(\eta_n)}{V(\eta_n) - V(\mu_n)} \notin (0, 1) \). The subsequence \( \{(\mu_{n(m)}, \eta_{n(m)})\} \) converges to \( (\mu, \eta) \) and for all \( m \), \( \frac{U(\mu_{n(m)}) - U(\eta_{n(m)})}{V(\eta_{n(m)}) - V(\mu_{n(m)})} \in (0, 1) \). Note that \( \tau(\mu_{n(m)}, \eta_{n(m)}) = 0 \) for all these discarded \( (\mu_n, \eta_n) \). If we show that \( \limsup_{n \to \infty} \tau(\mu_{n(m)}, \eta_{n(m)}) \leq \tau(\mu, \eta) \), then that establishes \( \limsup_{n \to \infty} \tau(\mu_n, \eta_n) \leq \tau(\mu, \eta) \).

But the assertion is already proved in (ii).

(c) \( V(\mu) = V(\eta) \).

35
Lemma 10 that $\tau$ can FT preference that generates $\succsim$.

The proof is divided into three sections. The first studies $\preccurlyeq$, the second studies the normative preference $\succsim^*$ derived from $\succsim$. The third defines a candidate preference $\succeq$ in terms of $\succsim^*$ and verifies that $\succeq$ is indeed an FT preference that generates $C$. Uniqueness is proved in Appendix E.

By continuity of $U, V$,

$$\lim_{n \to \infty} U(\mu_n) - U(\eta_n) > 0$$
$$\lim_{n \to \infty} V(\eta_n) - V(\mu_n) = 0.$$ 

Thus, there exists $N$ such that for each $n \geq N,$

$$U(\mu_n) - U(\eta_n) > V(\eta_n) - V(\mu_n),$$

Then, for each $n \geq N,$ either $U(\mu_n) - U(\eta_n) \notin (0, 1]$ and $V(\eta_n) \neq V(\mu_n)$, or $V(\eta_n) = V(\mu_n)$. It follows from Lemma 10 that $\tau(\mu_n, \eta_n) = 0$ for all $n \geq N$. Hence, $\lim_{n \to \infty} \tau(\mu_n, \eta_n) = 0.$

The last lemma verifies that $C$ satisfies Commitment is Normative. Note that $\tau(\mu, \eta) = 0$ and $(\mu, \eta) \in \Omega$ imply that that $\tau$ is continuous at $(\mu, \eta)$.

**Lemma 13**

(a) $x >_t x \cup y \implies \tau((x+y)^t, (x \cup y)^t) = 0$ and $(x+t, (x \cup y)^t) \in \Omega$.

(b) $x >_t x \cup y >_t y \implies \tau((x \cup y)^t, y^t) = 0$ and $((x \cup y)^t, y^t) \in \Omega$.

**Proof.** (a) Let $x >_t x \cup y$, that is,$$W(x) + \frac{\gamma}{\delta^t} \nabla(x) > W(x \cup y) + \frac{\gamma}{\delta^t} \nabla(x \cup y).$$

From $x < x \cup y$ it follows that $\nabla(x \cup y) \geq \nabla(x)$. Hence the displayed inequality implies $W(x) > W(x \cup y)$, which in turn implies $U(x^t) > U((x \cup y)^t)$. By Lemmas 11 and 12, $\tau(x^t, (x \cup y)^t) = 0$ and $(x+t, (x \cup y)^t) \in \Omega$.

(b) Note that by [18, Lemma 1(a)], $W(x) > W(x \cup y)$ implies 
$$\nabla(x \cup y) = \nabla(y) > \nabla(x).$$

Hence, $x \cup y >_t y$ implies
$$W(x \cup y) + \frac{\gamma}{\delta^t} \nabla(x \cup y) > W(y) + \frac{\gamma}{\delta^t} \nabla(y).$$

It follows that $W(x \cup y) > W(y)$, which implies $U((x \cup y)^t) > U(y^t)$. By Lemmas 11 and 12, $\tau((x \cup y)^t, y^t) = 0$ and $((x \cup y)^t, y^t) \in \Omega.$

**D Appendix: Proof of Theorem 1 (Sufficiency)**

The proof is divided into three sections. The first studies $\succeq$ and the second studies the normative preference $\succsim^*$ derived from $\succeq$. The third defines a candidate preference $\succeq$ in terms of $\succsim^*$ and verifies that $\succeq$ is indeed an FT preference that generates $C$. Uniqueness is proved in Appendix E.
D.1 Properties of $\succeq$

Define the choice correspondence $C^\ast(\cdot, \succeq)$ by
\[
C^\ast(x, \succeq) \equiv \{ \mu \in x : \mu \succeq \eta \text{ for all } \eta \in x \}.
\]

Say that $\succeq$ rationalizes $C(\cdot)$ if $C(x) = C^\ast(x, \succeq)$ for all $x$.

Lemma 14 $\succeq$ is the unique preference relation that rationalizes $C(\cdot)$. Furthermore $\succeq$ satisfies the vNM axioms.

Proof. Step 1: $\succeq$ is continuous.

We want to show that $\{ \eta : \eta \succeq \mu \}$ and $\{ \eta : \mu \succeq \eta \}$ are closed for all $\mu$. To show that the former is closed, take $\{ \eta_n \}$ such that $\eta_n \succeq \mu$ for all $n$ and $\eta_n \to \eta$. Consider the sequence of menus $\{ \{\eta_n, \mu\} \}$. Since $Z$ is endowed with the Hausdorff metric, $\eta_n \to \eta$ implies $\{\eta_n, \mu\} \to \{\eta, \mu\}$. By definition of $\succeq$, $\eta_n \in C(\{\eta_n, \mu\})$ for all $n$. By Continuity (upper hemicontinuity of $C$) and by [4, Thm 16.20], $\eta \in C(\{\eta, \mu\})$, that is, $\eta \succeq \mu$, as desired. This establishes that $\{ \eta : \eta \succeq \mu \}$ is closed. To see that $\{ \eta : \mu \succeq \eta \}$ is closed, take $\{ \eta_n \}$ such that $\mu \succeq \eta_n$ for all $n$ and $\eta_n \to \eta$, and consider the sequence of menus $\{ \{\eta_n, \mu\} \}$. Since $\{\eta_n, \mu\} \to \{\eta, \mu\}$ and $\mu \in C(\{\eta_n, \mu\})$ for all $n$, Continuity implies $\mu \in C(\{\eta, \mu\})$, that is, $\mu \succeq \eta$ as desired. Hence continuity of $\succeq$ is established.

For the next steps, say that a binary relation is a weak order if it is complete and transitive and define the revealed preference relation $\succeq'$ (with domain $\Delta$) by
\[
\mu \succeq' \eta \iff \exists x \text{ such that } \mu, \eta \in x \text{ and } \mu \in C(x).
\]

Step 2: $\succeq' = \succeq$.

Suppose $\mu \succeq \eta$. Then, by definition, there exists $x$ (namely $\{\mu, \eta\}$) such that $\mu, \eta \in x$ and $\mu \in C(x)$; so $\mu \succeq' \eta$. Hence, $\mu \succeq \eta \implies \mu \succeq' \eta$. Conversely, if $\mu \succeq' \eta$, then $\exists x$ such that $\mu, \eta \in x$ and $\mu \in C(x)$. Nonemptiness of $C$ and WARP imply $\mu \in C(\{\mu, \eta\})$. Hence $\mu \succeq' \eta \implies \mu \succeq \eta$.

Step 3: $C^\ast(\cdot, \succeq)$ is nonempty.

The domain $Z$ consists of compact menus, $\succeq$ is continuous (Step 1) and thus, $C^\ast(\cdot, \succeq) \neq \emptyset$ by [4, Thm 2.41].

Step 4: $\succeq$ is the unique weak order that rationalizes $C(\cdot)$.

The result follows from Steps 2 and 3, and [16, Prop 1.D.2].

Step 5: $\succeq$ satisfies the vNM axioms.

This follows from Steps 1 and 4, and by Independence.  ■

For $t > 0$, define $\succeq_t$ over $Z$ by
\[
x \succeq_t y \iff x^{+t} \succeq y^{-t}.
\]

We establish that each $\succeq_t$ satisfies the following axioms.\[26\]

B1 (Order**) $\succeq_t$ is complete and transitive.

\[26\]These axioms are discussed in [18].
B2 (Continuity**) The sets \( \{ y : x \succeq_t y \} \) and \( \{ y : y \succeq_t x \} \) are closed.

B3 (Independence**) \( x >_t y \implies ax + (1 - \alpha)z >_t ay + (1 - \alpha)z \).

B4 (Set-Betweenness**) \( x \succeq_t y \implies x \succeq_t \{ x \cup y \} \).

B5 (Separability**) If \( \mu^1 = \pi^1, \mu^2 = \pi^2, \eta^1 = \nu^1 \) and \( \eta^2 = \nu^2 \), then,

\[ \{ \mu, \eta \} \approx_t \{ \pi, \nu \} \] .

Let \( \Delta_s \subset \Delta \) be the set of lotteries on \( C \times Z \) with finite support and \( \Delta_s(Z) \) the set of lotteries on \( Z \) with finite support. Let \( \delta_x \) denote the lottery degenerate at menu \( x \). Define \( \varphi : \Delta_s(Z) \to Z \) by

\[ \varphi(\sum p(x)\delta_x) = \sum p(x)x. \]

B6 (Indifference to Timing**) For all \( \mu, \eta, \pi, \nu \in \Delta_s \), if \( \mu_1 = \pi_1, \eta_1 = \nu_1, \varphi(\mu_2) = \varphi(\pi_2) \) and \( \varphi(\eta_2) = \varphi(\nu_2) \), then,

\[ \{ \mu, \eta \} \approx_t \{ \pi, \nu \} . \]

Start by showing that \( \succeq_t \) satisfies Order*, Continuity*, Independence* and Set-Betweenness*, that is, \( \succeq_t \) is a Self-Control preference [7].

**Lemma 15** \( \succeq_t \) satisfies Order**, Continuity**, Independence** and Set-Betweenness**.

**Proof.** It is clear that by Lemma 14, \( \succeq_t \) satisfies Order** and Continuity**. By Set-Betweenness, \( \succeq_t \) satisfies Set-Betweenness**. To see that \( \succeq_t \) satisfies Independence**, observe that by Independence, for any \( x, y, z \),

\[ x^+ > y^+ \implies ax^+ + (1 - \alpha)z^+ > ay^+ + (1 - \alpha)z^+ , \]

and that by Indifference to Timing,

\[ ax^+ + (1 - \alpha)z^+ \approx (ax + (1 - \alpha)z)^+ , \]

\[ ay^+ + (1 - \alpha)z^+ \approx (ay + (1 - \alpha)z)^+ . \]

Therefore, by Order** and definition of \( \succeq_t \),

\[ x >_t y \implies ax + (1 - \alpha)z > ay + (1 - \alpha)z , \]

that is, \( \succeq_t \) satisfies Independence**. \( \blacksquare \)

By Lemma 15 and [7, Theorem 1], each \( \succeq_t \) is represented by \( W_t : Z \to R \) such that

\[ W_t(x) = \max_{\mu \in x}(U_t(\mu) + V_t(\mu)) - \max_{\eta \in Z} V_t(\eta) , \tag{18} \]

where \( U_t, V_t : \Delta \to R \) are linear and continuous.

**Lemma 16** \( \mu \succeq \eta \implies U_t(\mu) + V_t(\mu) \geq U_t(\eta) + V_t(\eta) . \)
Proof. By [18, Lemma 1] and Sophistication, the desired equivalence holds for all $\mu, \eta$ such that $\{\mu\} \nsucc_T \{\eta\}$. We need to establish the result for all $\mu, \eta$. Fix any $t > 0$. We show that the set

$$S = \{(\mu, \eta) : \{\mu\} \nsucc_T \{\eta\}\}$$

is a dense subset of $\Delta \times \Delta$. Since $\succsim^*$ and the preference represented by $U_t(\cdot) + V_t(\cdot)$ agree on the ranking of each pair $(\mu, \eta) \in S$, the desired result then follows by the continuity of both $\succsim^*$ and the preference represented by $U_t(\cdot) + V_t(\cdot)$.

By the nondegeneracy condition in Set-Betweenness, there exists $x, y, t$ such that $x^{+t} > (x \cup y)^{+t}$. Let $\rho = x^{+t}$ and $\nu = (x \cup y)^{+t}$, so that $\rho > \nu$. By Commitment is Normative, $\rho >_t \nu$. Since $\succsim^*_t$ restricted to $\Delta$ (the set of singletons in $Z$) satisfies the vNM axioms, it follows that for any $\mu, \eta$ such that $\{\mu\} \nsucc_T \{\eta\}$,

$$\{\mu_\alpha \rho\} >_t \{\eta_\alpha \nu\} \text{ for all } \alpha \in (0, 1).$$

Conclude that for any $(\mu, \eta) \in (\Delta \times \Delta) - S$, there is a sequence in $S$ that converges to $(\mu, \eta)$. □

Lemma 17 $\succsim^*_t$ satisfies Separability**.

Proof. Step 1: Show that $U_t(\mu) = \int_{C \times Z} \left( u_t(c) + \hat{W}_t(y) \right) d\mu(c, y)$.

Take $\mu, \eta$ such that $\mu = \frac{1}{2}(c, x) + \frac{1}{2}(c, \bar{x})$ and $\eta = \frac{1}{2}(c, x) + \frac{1}{2}(c, \bar{x})$. By Separability,

$$\left\{ \frac{1}{2}(c, x) + \frac{1}{2}(c, \bar{x}) \right\} \nsucc_T \left\{ \frac{1}{2}(c, \bar{x}) + \frac{1}{2}(c, x) \right\}.$$  

It follows that $\succsim^*_t$ satisfies GP’s version of Separability, and so, by [8, Lemma 9(1)], $U$ is additively separable, thus establishing Step 1.

Step 2: Show that

$$U_t(\mu) + V_t(\mu) = \int_{C \times Z} \left( u_t(c) + \hat{W}_t(y) + v_t(c) + \hat{V}_t(y) \right) d\mu(c, y)$$

and $V_t(\mu) = \int_{C \times Z} \left( v_t(c) + \hat{V}_t(y) \right) d\mu(c, y)$.

Since $V$ is linear and continuous, there exists a continuous function $\varpi_t : C \times Z \rightarrow R$ such that for all $\mu \in \Delta$,

$$V_t(\mu) = \int \varpi_t(c, x) d\mu.$$  

By Separability, $\frac{1}{2}(c, x) + \frac{1}{2}(c, \bar{x}) \approx \frac{1}{2}(c, x) + \frac{1}{2}(c, \bar{x})$. Hence, by Lemma 16,

$$U_t\left( \frac{1}{2}(c, x) + \frac{1}{2}(c, \bar{x}) \right) + V_t\left( \frac{1}{2}(c, x) + \frac{1}{2}(c, \bar{x}) \right)$$

$$= U_t\left( \frac{1}{2}(c, x) + \frac{1}{2}(c, \bar{x}) \right) + V_t\left( \frac{1}{2}(c, x) + \frac{1}{2}(c, \bar{x}) \right).$$

By Step 1, $U_t\left( \frac{1}{2}(c, x) + \frac{1}{2}(c, \bar{x}) \right) = U_t\left( \frac{1}{2}(c, x) + \frac{1}{2}(c, \bar{x}) \right)$. Therefore,

$$U_t\left( \frac{1}{2}(c, x) + \frac{1}{2}(c, \bar{x}) \right) + V_t\left( \frac{1}{2}(c, x) + \frac{1}{2}(c, \bar{x}) \right) = U_t\left( \frac{1}{2}(c, x) + \frac{1}{2}(c, \bar{x}) \right) + V_t\left( \frac{1}{2}(c, x) + \frac{1}{2}(c, \bar{x}) \right)$$

$$\Rightarrow V_t\left( \frac{1}{2}(c, x) + \frac{1}{2}(c, \bar{x}) \right) = V_t\left( \frac{1}{2}(c, x) + \frac{1}{2}(c, \bar{x}) \right)$$

$$\Rightarrow V_t(c, x) + V_t(c, \bar{x}) = V_t(c, x) + V_t(c, \bar{x})$$

$$\Rightarrow \varpi_t(c, x) + \varpi_t(c, \bar{x}) = \varpi_t(c, x) + \varpi_t(c, \bar{x})$$

$$= \varpi_t(c, x) + \varpi_t(c, \bar{x}) = \varpi_t(c, x) + \varpi_t(c, \bar{x})$$

39
Hence by (18),

\[
\{ a \text{ set of preference relations defined over } \Delta \}
\]

Since \(C\) to establish the result.

Lemma 17, use Indifference to Timing to show that \(\hat{\tau}\) satisfies the conditions in Theorem B.1. Thus, there is a well-defined normative preference \(\succeq\).

The proof follows steps that are similar to those in Lemma 17: for the \(\bar{W}\) and \(\hat{W}\) in the proof of Lemma 17, use Indifference to Timing to show that \(\hat{W}\) is linear, then show that \(\bar{W}\) is linear, and then use (18) to establish the result.

**Lemma 18** \(\succeq\) satisfies Indifference to Timing**.

**Proof.** The proof follows steps that are similar to those in Lemma 17: for the \(\bar{W}\) and \(\hat{W}\) in the proof of Lemma 17, use Indifference to Timing to show that \(\hat{W}\) is linear, then show that \(\bar{W}\) is linear, and then use (18) to establish the result.

**D.2 Normative Preference \(\succeq\)**

For each \(t > 0\), denote the restriction of \(\succeq_t\) to \(\Delta \subset Z\) by \(\succeq_t\), and define \(\succeq_{0} = \succeq\). Then \(\{\succeq_t\}_{t=0}^{\infty}\) is a set of preference relations defined over \(\Delta\). Since \(C(\cdot)\) satisfies WARP, Continuity and Reversal, \(\{\succeq_t\}_{t=0}^{\infty}\) satisfies the conditions in Theorem B.1. Thus, there is a well-defined normative preference \(\succeq^*\) defined over \(\Delta\) and a well-defined function \(\tau : \Delta \times \Delta \rightarrow R\) as in Lemma 1. Let \(\Omega\) be the subset of \(\Delta \times \Delta\) on which \(\tau\) is upper semicontinuous, that is,

\[
\Omega = \{ (\mu, \eta) \in \Delta \times \Delta : (\mu_n, \eta_n) \to (\mu, \eta) \implies \lim_{n \to \infty} \tau(\mu_n, \eta_n) \leq \tau(\mu, \eta) \}.
\]

Lemma 19 states some results from Appendix B.

**Lemma 19**

(a) \(\mu \succ^* \eta \iff [\mu \succ_{(\mu, \eta)} \eta\text{ and } (\mu, \eta) \in \Omega] \)

(b) \(\mu \succeq^* \eta \iff \text{there exists a sequence } \{(\mu_n, \eta_n)\} \text{ that converges to } (\mu, \eta) \text{ and } \mu_n \succeq^*_{(\mu_n, \eta_n)} \eta_n \text{ for all } n. \)

(c) \(\mu \succeq^*_{(\mu, \eta)} \eta \implies \mu \succeq^* \eta. \)

**Lemma 20** \(\succeq^*\) satisfies order, continuity and independence
Proof. Theorem B.1 establishes that \( \succeq^* \) is complete, transitive and continuous. To prove independence, it suffices by Theorem B.2 to show that each \( \succeq_{\Delta|\Delta} \) satisfies independence. Recall that \( \succeq_{\Delta|\Delta} \) is the restriction of \( \succeq_{\Delta|\Delta} \) to \( \Delta \), and apply Lemma 15. ■

In the remainder of the subsection we establish a stationarity property of \( \succeq^* \).

Lemma 21 \( (c, \mu) \succeq^* (c, \eta) \iff (c', \mu) \succeq^* (c', \eta) \)

Proof. Step 1: \( (c, x) \frac{1}{2} (c', x') \sim^* (c, x') \frac{1}{2} (c', x) \).

By Separability, for all \( t \),
\[
\frac{1}{2} (c, x) + \frac{1}{2} (c', x') + t = \frac{1}{2} (c, x') + \frac{1}{2} (c', x) + t.
\]

Hence, \( \tau((c, x) \frac{1}{2} (c', x'), (c, x') \frac{1}{2} (c', x)) = 0 \) and \( (c, x) \frac{1}{2} (c', x') \approx (c, x') \frac{1}{2} (c', x) \). Apply Lemma 19(c) to obtain
\[
(c, x) \frac{1}{2} (c', x') \sim^* (c, x') \frac{1}{2} (c', x).
\]

Step 2: The result.

Suppose by way of contradiction that \( (c, \mu) \succeq^* (c, \eta) \) and \( (c', \mu) \sim^* (c', \eta) \). Since \( \succeq^* \) satisfies the vNM axioms,
\[
(c, \mu) \frac{1}{2} (c', \eta) \sim^* (c, \eta) \frac{1}{2} (c', \mu).
\] But by Step 1,
\[
(c, \eta) \frac{1}{2} (c', \mu) \sim^* (c, \mu) \frac{1}{2} (c', \eta) \text{ and } (c, \mu) \frac{1}{2} (c', \eta) \sim^* (c, \eta) \frac{1}{2} (c', \mu),
\]
and so \( (c, \eta) \frac{1}{2} (c', \mu) \sim^* (c, \mu) \frac{1}{2} (c', \eta) \), contradicting (19). ■

Lemma 22 \( \mu^{+1} \alpha \eta^{+1} \sim^* (\mu \alpha \eta)^{+1} \)

Proof. By Indifference to Timing, \( \mu^{+1} \alpha \eta^{+1} \approx (\mu \alpha \eta)^{+1} \), and by Lemma 18, for all \( t \geq 1 \), \( \{\mu^{+1} \alpha \eta^{+1}\}^{+t} \approx \{\mu \alpha \eta)^{+1}\}^{+t} \). Hence, \( \tau(\mu^{+1} \alpha \eta^{+1}, (\mu \alpha \eta)^{+1}) = 0 \) and \( \mu^{+1} \alpha \eta^{+1} \approx (\mu \alpha \eta)^{+1} \). Apply Lemma 19(c) to obtain \( \mu^{+1} \alpha \eta^{+1} \sim^* (\mu \alpha \eta)^{+1} \). ■

Lemma 23 \( \mu \succeq^* \eta \implies (c, \mu) \succeq^* (c, \eta) \)

Proof. Step 1: \( \mu \succeq_{\tau(\mu, \eta)} \eta \iff \mu^{+1} \succeq_{\tau(\mu^{+1}, \eta^{+1})} \eta^{+1} \).

First show that if \( \tau(\mu, \eta) > 0 \), then
\[
\tau(\mu^{+1}, \eta^{+1}) = \tau(\mu, \eta) - 1.
\]
For this purpose, suppose without loss of generality that \( \mu \leq \eta \). If \( \tau(\mu, \eta) > 0 \), then by Lemma 1, \( \mu^{+t} > \eta^{+t} \) for all \( t \geq \tau(\mu, \eta) \) and \( \mu^{+t} \leq \eta^{+t} \) for all \( t < \tau(\mu, \eta) \). It follows that
\[
\mu^{+t+1} > \eta^{+t+1}, \text{ for all } t \geq \tau(\mu, \eta) - 1.
\]
and
\[
\mu^{+t+1} \leq \eta^{+t+1}, \text{ for all } t < \tau(\mu, \eta) - 1.
\]
The assertion follows. Now prove the result. It follows by definition of \( \tau \) if \( \tau(x, y) = 0 \) (in which case \( \tau(x^{+1}, y^{+1}) = 0 \) as well). When \( \tau(\mu, \eta) > 0 \), then note that \( \mu^{+\tau(\mu, \eta)} = \tau(\mu, \eta)^{+\tau(\mu, \eta) - 1} \) and \( \eta^{+\tau(\mu, \eta)} = \tau(\eta, \eta)^{+\tau(\mu, \eta) - 1} \). The result follows from (20).
Step 2: \( \mu \succ^* \eta \implies \mu^{+1} \succ^* \eta^{+1}. \)

If \( \mu \succ^* \eta \), then by Lemma 19(b), there exists a sequence \( \{(\mu_n, \eta_n)\} \) such that \((\mu_n, \eta_n) \to (\mu, \eta)\) and \(\mu_n \succ^*\tau(\mu_n, \eta_n) \eta_n\) for all \(n\). It follows by Step 1 that \(\{(\mu^{n+1}_n, \eta^{n+1}_n)\}\) is a sequence such that \((\mu^{n+1}_n, \eta^{n+1}_n) \to (\mu^{+1}, \eta^{+1})\) and \(\mu^{n+1}_n \succ^*\tau(\mu^{n+1}_n, \eta^{n+1}_n) \eta^{n+1}_n\) for all \(n\). But then by Lemma 19(b), \(\mu^{+1} \succ^* \eta^{+1}\). The result follows by Lemma 21.

**Lemma 24** \( \mu \succ^* \eta \iff (c, \mu) \succ^* (c, \eta) \).

**Proof.** By Lemma 21, it suffices to show \( \mu \succ^* \eta \iff (\tau, \mu) \succ^* (\tau, \eta) \). Define a binary relation \(\succ^*\) over \(\Delta\) by \( \mu \succ^* \eta \iff (\tau, \mu) \succ^* (\tau, \eta) \). We need to show \( \mu \succ^* \mu' \iff \mu \succ^* \mu' \). However, this follows from the following three observations:

(a) By Lemma 23, \( \mu \succ^* \eta \implies \mu \succ^* \eta \).

(b) The preference \(\succ^*\) is non-trivial, that is, there exist \(\rho, \nu \in \Delta\) such that \(\rho \succ^{**} \nu\): By the non-degeneracy condition in Set-Betweenness, there exists \(x, y\) such that \(x^{+t} \succ (x \cup y)^{+t}\). By Commitment is Normative, we can assume that \(t > 1\). Commitment is Normative also implies \(\tau(x^{+t}, (x \cup y)^{+t}) = 0\) and \(\tau\) is continuous at \((x^{+t}, (x \cup y)^{+t})\), and so \((x^{+t}, (x \cup y)^{+t}) \in \Omega\). By Lemma 19(a), \(x^{+t} \succ^* (x \cup y)^{+t}\). It follows that \(x^{+(t-1)} \succ^* (x \cup y)^{+(t-1)}\), that is, \(\succ^*\) is non-trivial.

(c) The preference \(\succ^*\) satisfies the von Neumann-Morgenstern (vNM) axioms: Since \(\succ^*\) is a continuous weak order, so is \(\succ^{**}\). To show independence, note that

\[
\mu \succ^{**} \eta \\
\implies (\tau, \mu) \succ^* (\tau, \eta) \\
\implies \alpha(\tau, \mu) + (1 - \alpha)(\tau, \nu) \succ^* \alpha(\tau, \eta) + (1 - \alpha)(\tau, \nu) \\
\implies (\tau, \alpha \mu + (1 - \alpha) \nu) \succ^* (\tau, \alpha \eta + (1 - \alpha) \nu) \\
\implies \alpha \mu + (1 - \alpha) \nu \succ^{**} \alpha \eta + (1 - \alpha) \nu. \]

**D.3 FT Preference \(\succeq\)**

The candidate for the FT preference \(\succeq\) over \(Z\) that generates \(C(\cdot)\) is defined by

\[ x \succeq y \iff (c, x) \succ^* (c, y), \]

for some \(c \in C\). By Lemma 21, the preference \(\succeq\) is invariant to the choice of \(c\). We verify that \(\succeq\) is an FT preference by checking that it satisfies the conditions in [18, Theorems 1 and 3].

**Lemma 25** For any \(x, y\) and \(t \geq 1\),

(a) \(x \succ_t x \cup y \implies x \succ x \cup y\).

(b) \(x \succ_t x \cup y \implies x \succ x \cup y \implies y\).

(c) \(x \succ_t x \cup y \succeq_t y \implies x \succ x \cup y \sim y\).

**Proof.** Suppose \(x \succ_t x \cup y\), or equivalently, \(x^{+t} \succ (x \cup y)^{+t}\). By Commitment is Normative, \(\tau(x^{+t}, (x \cup y)^{+t}) = 0\) and \(\tau\) is continuous at \((x^{+t}, (x \cup y)^{+t})\), and so \((x^{+t}, (x \cup y)^{+t}) \in \Omega\). Hence, by Lemma 19(a), \(x^{+t} \succ^* (x \cup y)^{+t}\), and repeated application of Lemma 24 establishes \(x^{+1} \succ^* (x \cup y)^{+1}\), which in turn establishes (a). A similar argument establishes (b).
Turn to (c). In what follows we prove that \( x \cup y \approx y \) for all \( t' \geq t \), since then \((x \cup y)^{t'} \approx (x \cup y)^{t+1}\) \( y^{t+1} \), and so by Lemma 19(c), \((x \cup y)^{t'} \sim^{*} y^{t+1}, \) that is, \( x \cup y \sim y \).

As before, \( x \sim t \cup y \) implies \( x \sim_{t} x \cup y \) for all \( t' \geq t \). It follows by Set-Betweenness that \( x \sim_{t} y \) for all \( t' \geq t \). By [18, Lemma 1(b)], for all \( t' \geq t \),

\[
x \cup y \approx y \Leftrightarrow \max_{\mu \in y}(U') + V') \geq \max_{\eta \in x}(U) + V).
\]

where \( U' \) and \( V' \) are as in (18). By Lemma 16, for all \( t' \geq t \),

\[
\max_{\mu \in y}(U') + V') \geq \max_{\eta \in x}(U) + V) \iff \max_{\mu \in y}(U) + V) \geq \max_{\eta \in x}(U) + V).
\]

By hypothesis, \( x \cup y \approx y \). Thus, by (21) and (22), \( x \cup y \approx y \) for all \( t' \geq t \), as desired. ■

In the remainder of the proof we verify that \( \approx \) is a nondegenerate FT preference and that it generates \( C(\cdot) \). In Appendix D.1 we stated some axioms for the preference \( \approx \). The same names will be used for axioms that impose similar restrictions on \( \approx \).

**Lemma 26.** \( \approx \) satisfies Order**, Continuity**. Moreover, \( \approx \) satisfies Independence**: for all \( \alpha \in (0,1) \),

\[
\{ \rho \} \triangleright \{ \eta \} \implies \{ \alpha \rho + (1 - \alpha) \eta \} \triangleright \{ \alpha \eta + (1 - \alpha) \rho \}.
\]

**Proof.** Follows from Lemmas 20 and 24, and definition of \( \approx \). ■

**Lemma 27.** \( \approx \) satisfies Stationarity**: \( z \approx z' \iff \{(c, z)\} \approx \{(c, z')\} \).

**Proof.** This follows from Lemmas 24 and 21, and by definition of \( \approx \). ■

**Lemma 28.** \( \approx \) satisfies Set-Betweenness**: \( x \approx y \implies x \approx x \cup y \approx y \).

**Proof.** Begin by establishing that if \( \mu > \tau(\mu, \eta) \eta \) then

\[
(\mu^{+1}, \eta^{+1}) \in \Omega \implies (\mu, \eta) \in \Omega.
\]

Suppose \((\mu, \eta) \not\in \Omega \), so that there exists a sequence \( \{(\mu_n, \eta_n)\} \) that converges to \((\mu, \eta) \) and \( \lim \sup_n \tau(\mu, \eta_n) \geq \tau(\mu, \eta) \). Without loss of generality, \( \mu \eta \) for all \( n \). Suppose by way of contradiction that \( \lim \sup_n \tau(\mu, \eta_n) \geq T < \infty \). Thus, there exists \( N \) such that for all \( n \geq N, T + 1 > \tau(\mu, \eta_n) \) Also, for large enough \( n \), \( \mu_n \tau(\mu, \eta_n) \eta_n \), and since

\[
T + 1 > \tau(\mu, \eta_n) \geq \tau(\mu, \eta),
\]

it follows that for all large enough \( n, \eta_n \approx T + 1 \mu_n \). By continuity of \( \approx T + 1 \mu \). But since \( T + 1 > \tau(\mu, \eta) \), this contradicts the hypothesis that \( \mu > \tau(\mu, \eta) \eta \). Therefore, \( \lim \sup_n \tau(\mu, \eta_n) \geq T < \infty \).

To complete the argument, observe that \( \{(\mu_n^{+1}, \eta_n^{+1})\} \) is a sequence that converges to \((\mu^{+1}, \eta^{+1}) \), and by (20), \( \lim \sup_n \tau(\mu_n^{+1}, \eta_n^{+1}) = \infty \).

It follows that \((\mu^{+1}, \eta^{+1}) \not\in \Omega \), thus proving (23).

To prove Set-Betweenness**, we need to show

\[
x'^{+} \approx y'^{+} \implies x^{+} \approx x \cup y \approx x^{+} \).
\]

Note that we had assumed \( \tau(\mu, \eta) \eta \) for all \( n \), and so \( \tau(\mu, \eta_n) > 0 \) for all \( n \), as required by (20).
Denote $\tau(x^{+1},y^{+1})$ by $\tau$ and consider two cases.

Case (a): $x^{+1} \succ y^{+1}$

Then by Lemma 19(a), $x^{+1} \succ y^{+1}$. By Set-Betweenness, $x^{+1} \succeq_{t} (x \cup y)^{+1} \succeq_{t} y^{+1}$ for all $t \geq \tau$. Hence

$$x^{+1} \succeq_{\tau(x^{+1},(y^{+1})^{+1}} (x \cup y)^{+1} \succeq_{\tau((y^{+1})^{+1},y^{+1})} y^{+1},$$

and by Lemma 19(c), $x^{+1} \succ^* (x \cup y)^{+1} \succ^* y^{+1}$.

Case (b): $x^{+1} \prec^* y^{+1}$

By Lemma 24, $x^{+\tau} \prec^* y^{+\tau}$. Note that $\tau(x^{+\tau},y^{+\tau}) = 0$. By Lemma 19(a), either $x^{+\tau} \approx y^{+\tau}$ or $[x^{+\tau} \neq y^{+\tau}$ and $(x^{+\tau},y^{+\tau}) \notin \Omega]$ holds.

(i) Suppose first that $x^{+\tau} \approx y^{+\tau}$. Then by Set-Betweenness, $x^{+\tau} \approx_{t} (x \cup y)^{+\tau} \approx_{t} y^{+\tau}$ for all $t$. Thus, $\tau(x^{+\tau},(x \cup y)^{+\tau}) = \tau((x \cup y)^{+\tau},y^{+\tau}) = 0$ and $x^{+\tau} \approx (x \cup y)^{+\tau} \approx y^{+\tau}$. By Lemma 19(c), $x^{+\tau} \prec^* (x \cup y)^{+\tau} \prec^* y^{+\tau}$. By repeated application of Lemma 24, $x^{+1} \prec^* (x \cup y)^{+1} \prec^* y^{+1}$, as desired.

(ii) Now suppose $[x^{+\tau} \neq y^{+\tau}$ and $(x^{+\tau},y^{+\tau}) \notin \Omega]$. Since, $\tau(x^{+\tau},y^{+\tau}) = 0$, we have $x^{+\tau} \succ y^{+\tau}$ for all $t$. By Set-Betweenness, $x^{+\tau} \succeq_{t} (x \cup y)^{+\tau}$ for all $t$. If $x^{+\tau} \succ_{t} (x \cup y)^{+\tau}$ for some $t$, then Commitment is Normative implies that $\tau$ is continuous at $(x^{+\tau},y^{+\tau}+\tau)$, which implies $(x^{+\tau}+\tau, y^{+\tau}+\tau) \notin \Omega$. By repeated application of (23), $(x^{+\tau},y^{+\tau}) \in \Omega$, a contradiction. Therefore, $x^{+\tau} \approx_{t} (x \cup y)^{+\tau}$ for all $t$, and so, by an argument similar to the one in (i), Lemma 19(c) and repeated application of Lemma 24 implies $x^{+1} \prec^* (x \cup y)^{+1}$. By hypothesis $x^{+1} \prec^* y^{+1}$ and we have seen that $x^{+1} \prec^* (x \cup y)^{+1}$. Hence by transitivity of $\succ^*$, $(x \cup y)^{+1} \succ^* y^{+1}$. Put together, $x^{+1} \prec^* (x \cup y)^{+1} \prec^* y^{+1}$, as desired. ■

**Lemma 29** $\succeq$ satisfies Separability**: if $\mu^{1} = \pi^{1}, \mu^{2} = \pi^{2}, \eta^{1} = \nu^{1}$ and $\eta^{2} = \nu^{2}$, then $\{\mu, \eta\} \prec \{\pi, \nu\}$.**

**Proof.** By Lemma 17, $\{\mu, \eta\} \approx_{t} \{\pi, \nu\}$ for all $t \geq 1$. Thus, $\tau(\{\mu, \eta\}^{+1}, \{\pi, \nu\}^{+1}) = 0$ and $\{\mu, \eta\}^{+1} \approx_{t} \{\pi, \nu\}^{+1}$. From Lemma 19(c), we see that $\{\mu, \eta\}^{+1} \prec^* \{\pi, \nu\}^{+1}$, and so $\{\mu, \eta\} \prec \{\pi, \nu\}$. ■

**Lemma 30** $\succeq$ satisfies Indifference to Timing**: for any $\mu, \eta, \pi, \nu \in \Delta_{s}$, if $\mu^{1} = \pi^{1}, \eta^{1} = \nu^{1}, \varphi(\pi^{2}) = \varphi(\eta^{2}) = \varphi(\mu^{2})$, then $\{\mu, \eta\} \prec \{\pi, \nu\}$.**

**Proof.** Observe that by Lemma 18, $\{\mu, \eta\} \approx_{t} \{\pi, \nu\}$ for all $t \geq 1$, and then argue as in Lemma 29. ■

**Lemma 31** $\succeq$ satisfies Temptation Stationarity**: $x \succ x \cup y \iff \{(c,x)\} \succ \{(c,x),(c,y)\}$.**

**Proof.** By Menus Can Tempt, $\tau(x^{+1},(x \cup y)^{+1}) = \tau(\{(c,x)\}^{+1}, \{(c,x),(c,y)\}^{+1})$. Denote these by $\tau$. Observe that

$$x \succ x \cup y \iff x^{+1} \succ^* (x \cup y)^{+1} \iff (x^{+1})^{+\tau} > (x \cup y)^{+\tau} \text{ and } ((x^{+1})^{+\tau}, ((x \cup y)^{+\tau})^{+\tau}) \in \Omega \text{ by Lemma 19(a)}$$

and $\{(c,x)\}^{+1} \succ^* \{(c,x),(c,y)\}^{+1}$ by Lemma 19(a) and $\{(c,x)\}^{+1} \succ\ast \{(c,x),(c,y)\}^{+1}$ by Lemma 19(a) $

\iff \{(c,x)\} \succ \{(c,x),(c,y)\}$, as desired. ■
Lemma 32 \( \succcurlyeq \) is nondegenerate.

**Proof.** Use the nondegeneracy condition in Set-Betweenness and Lemma 25(b). □

The above Lemmas establish that \( \succcurlyeq \) satisfy the conditions of [18, Thm 1], and so there exist \( \delta, \gamma \in (0,1) \), continuous functions \( u, v : C \to R \) and continuous linear functions \( U, V : \Delta(C \times Z) \to R \) and \( W, \overline{V} : Z \to R \) such that for all \( x \in Z \),

\[
W(x) = \max_{\mu \in x} (U(\mu) + V(\mu) - \max_{\eta \in x} V(\eta)),
\]

\[
U(\mu) = \int_{C \times Z} (u(c) + \delta W(y)) d\mu(c, y),
\]

\[
V(\mu) = \int_{C \times Z} (v(c) + \gamma \overline{V}(y)) d\mu(c, y), \text{ where } \overline{V}(x) = \max_{\eta \in x} V(\eta)
\]

and furthermore, \( W \) represents \( \succcurlyeq \). It remains to show that \( \gamma < \delta \) and that \( \succcurlyeq \) generates \( C(\cdot) \).

**Lemma 33** \( \succcurlyeq \) generates \( C(\cdot) \).

**Proof.** By nondegeneracy of \( \succcurlyeq \), \( U \) is not an affine transformation of \( V \) (see proof of [18, Theorem 2]). Lemmas 25(a) and 25(b) establish that for all \( t \), \( \succcurlyeq \) has more preference for commitment than \( \succeq_t \) and \( \succcurlyeq \) has more self-control than \( \succeq_t \); see GP for definitions of these terms. By [7, Theorem 8],

\[
U_t = \alpha U + (1 - \alpha) V \quad \text{and} \quad V_t = \alpha' U + (1 - \alpha') V,
\]

for \( \alpha, \alpha' \in [0, 1] \), which implies

\[
U_t + V_t = (\alpha + \alpha') U + (2 - \alpha - \alpha') V. \tag{24}
\]

Furthermore, by [7, Theorem 9],

\[
U_t + V_t = \beta U + V \quad \text{for } \beta \in [0, 1]. \tag{25}
\]

Together, (24) and (25) imply \((\alpha + \alpha') U + (2 - \alpha - \alpha') V = \beta U + V \). Since \( U \) is not an affine transformation of \( V \), conclude that \( U_t + V_t = U + V \). Hence, by Lemma 16,

\[
\mu \succcurlyeq \eta \iff U(\mu) + V(\mu) \geq U(\eta) + V(\eta).
\]

By Lemma 14, \( \succcurlyeq \) rationalizes \( C(\cdot) \). Therefore the above displayed equivalence implies that for all \( x \in Z \),

\[
C(x) = \arg \max_{\mu \in x} \{U(\mu) + V(\mu)\},
\]

as desired. □

**Lemma 34** \( \gamma < \delta \).

**Proof.** Since \( \succcurlyeq \) generates \( C \),

\[
\mu \succcurlyeq \eta \iff U(\mu) + V(\mu) \geq U(\eta) + V(\eta).
\]

By Set-Betweenness, there is \( x, y \) and \( t \) such that \( x \subset y \) and \( x^{+t} > y^{+t} \). Thus,
\[
x^t > y^t
\]
\[
\iff U(x^t) + V(x^t) > U(y^t) + V(y^t)
\]
\[
\iff \sum_{i=0}^{t-1} \delta_i u(x) + \gamma_i v(x) + \delta_i W(x) + \gamma_i^I V(x) > \sum_{i=0}^{t-1} \delta_i u(y) + \gamma_i v(y) + \delta_i W(y) + \gamma_i^I V(y)
\]
\[
\iff W(x) + \frac{\gamma^I}{\delta} V(x) > W(y) + \frac{\gamma^I}{\delta} V(y).
\]
Since \( V(y) \geq V(x) \) and \( \frac{\gamma^I}{\delta} > 0 \), conclude that \( W(x) > W(y) \), and hence also that \( V(y) > V(x) \).
Suppose by way of contradiction that \( \frac{\gamma^I}{\delta} > 1 \). Then the preceding implies \( y^{t+T} \geq x^{t+T} \) for a large enough \( T > t \), that is, \( \tau(x^{t+T}, y^{t+T}) > 0 \). But, by Commitment is Normative, \( \tau(x^{t+T}, y^{t+T}) = 0 \), a contradiction. Thus, \( \frac{\gamma^I}{\delta} \leq 1 \).

Suppose by way of contradiction that \( \frac{\gamma^I}{\delta} = 1 \). Then, for all \( \mu, \eta, t > 0 \),
\[
\mu^t = \eta^t
\]
\[
\iff U(\mu^t) + V(\mu^t) = U(\eta^t) + V(\eta^t)
\]
\[
\iff W(\mu) + \frac{\gamma^I}{\delta} V(\mu) = W(\eta) + \frac{\gamma^I}{\delta} V(\eta)
\]
\[
\iff W(\mu) + \frac{\gamma^I}{\delta} V(\mu) \geq W(\eta) + \frac{\gamma^I}{\delta} V(\eta)
\]
\[
\iff U(\mu) + V(\mu) \geq U(\eta) + V(\eta)
\]
\[
\iff \mu \geq \eta.
\]
That is, \( \tau(\mu, \eta) = 0 \) for all \( \mu, \eta \). This contradicts the latter part of Reversal. Hence, \( \frac{\gamma^I}{\delta} < 1 \), and the assertion follows. \( \blacksquare \)

E Appendix: Proof of Theorem 1 (Uniqueness)

Lemma 35 If \( U(\mu) = U(\eta) \) and \( \mu > \eta \), then \((\mu, \eta) \notin \Omega\).

Proof. By the definition of \( \geq \), \( U(\mu) = U(\eta) \) and \( \mu < \eta \) imply \( V(\mu) < V(\eta) \). By nondegeneracy, there exists \( \nu, \rho \) such that \( U(\nu) > U(\rho) \) and \( V(\nu) < V(\rho) \). Consider the sequence \( \{(\mu_0, \nu), (\rho, \eta)\} \) that converges to \((\mu, \eta)\). Since \( U, V \) are linear, for each \( n \), \( U(\mu_0, \nu) > U(\rho, \eta) \) and \( V(\mu_0, \nu) < V(\rho, \eta) \) and by Lemma 11,
\[
\mu_0, \nu >_1 \rho, \eta \text{ for all } n \geq 1 \geq \mu(\rho_0, \nu, \eta).
\] (26)

The hypothesis \((U(\mu) = U(\eta) \) and \( \mu < \eta \) implies that for all \( t \),
\[
\mu <_t \eta.
\] (27)

Suppose by way of contradiction that \( \limsup_{n \to \infty} \tau(\mu_0, \nu, \rho, \eta) = T \leq \tau(\mu, \eta) \). Then there exists \( N \) such that
\[
\tau(\mu_0, \nu, \rho, \eta) < T + 1, \text{ for all } n \geq N.
\] (28)

However, \( \mu <_T + 1 \eta \) by (26) and so there exists \( N' \) such that
\[
\mu_0, \nu <_T + 1 \rho, \eta, \text{ for all } n \geq N'.
\]
But by (26) this implies \( \tau(\mu_0, \nu, \rho, \eta) > T + 1 \) for all \( n \geq \max\{N, N'\} \), a contradicting (28). \( \blacksquare \)

Lemma 36 If \( x > y \), then there exists \( T \) such that \((c, x) >_1 (c, y) \) for all \( t \geq T \).

Proof. The hypothesis implies \( U(c, x) > U(c, y) \), and the result follows from Lemma 11. \( \blacksquare \)
Lemma 37 If $x \sim y$, then $\tau(x^{+1}, y^{+1}) = 0$.

Proof. Since, for any $t$, 

$$(c, x) \gtrless_t (c, y) \iff W(x) + \frac{\gamma t}{\delta} V(x) \geq W(y) + \frac{\gamma t}{\delta} V(y),$$

the hypothesis $x \sim y$ implies $(c, x) \gtrless_t (c, y) \iff \nabla(x) \geq \nabla(y)$. It follows that for all $t, t'$, $(c, x) \gtrless_t (c, y) \iff (c, x) \gtrless_{t'} (c, y)$, that is, $\tau(x^{+1}, y^{+1}) = 0$. ■

Lemma 38 $C(\cdot)$ is generated by a unique FT preference $\succeq$.

Proof. Suppose, by way of contradiction, that $\succeq$ and $\succeq'$ are two different FT preferences that generate $C(\cdot)$. Then, there exist $x$ and $y$ such that $x \succ y$ and $y \succeq' x$. Let $(U, V)$ and $(U', V')$ be representations of $\succeq$ and $\succeq'$, respectively. Consider three possibilities:

(a) $x \succ y$ and $y \succ' x$.
   By Lemma 36, $\succeq$ and $\succeq'$ do not generate the same choice correspondence, a contradiction.

(b) $x \succ y$, $\tau(x^{+1}, y^{+1}) > 0$, and $y \sim' x$.
   By Lemma 37, $y \sim' x$ implies $\tau(x^{+1}, y^{+1}) = 0$, and so $\succeq$ and $\succeq'$ do not generate the same choice correspondence, a contradiction.

(c) $x \succ y$, $\tau(x^{+1}, y^{+1}) = 0$, and $y \sim' x$.
   By the representation, $x \succ y$ implies $\{c, x\} \succ \{c, y\}$, and so, $U(c, x) > U(c, y)$. By Lemma 12, $(x^{+1}, y^{+1}) \notin \Omega$. We show that $(x^{+1}, y^{+1}) \notin \Omega$ also holds, thereby establishing the desired contradiction. Lemma 36, $x \succ y$ and $\tau(x^{+1}, y^{+1}) = 0$ imply $(c, x) > (c, y)$. However, $y \sim' x$ and $(c, x) > (c, y)$ imply $U'(c, x) = U'(c, y)$ and $V'(c, x) > V'(c, y)$, and so, by Lemma 35, $(x^{+1}, y^{+1}) \notin \Omega$. ■

F Appendix: Proof of Theorem 2

First prove the Theorem for a representation $(U, V)$ of a nondegenerate FT preference $\succeq$ for which $V \geq 0$. Let $\succeq$ be the preference relation that is represented by $\varphi : \Delta(C \times Z) \rightarrow R$ where for all $\mu \in \Delta$,

$$\varphi(\mu) = U(\mu) + V(\mu).$$

For each $t > 0$, define $\succeq_t$ on $\Delta$ by

$$\mu \succeq_t \eta \iff \mu^{+t} \succeq \eta^{+t}.$$ 

It is straightforward to establish that $\succeq_t$ is represented by $\varphi_t : \Delta(C \times Z) \rightarrow R$ where for all $\mu \in \Delta$,

$$\varphi_t(\mu) = U(\mu) + \left(\frac{\gamma}{\delta}\right)^t V(\mu).$$

Lemma 39 The sequence $\{\varphi_t\}$ uniformly converges to $U$. 

47
Proof. The sequence \( \{ \varphi_t \} \) is a sequence of continuous real functions defined on a compact space \( \Delta \). Since \( V \geq 0 \) and \( \frac{4}{\alpha} < 1 \), the sequence is monotone decreasing and \( \varphi_t \) converges pointwise to the continuous function \( U \). Therefore, by Dini’s Theorem [4, Theorem 2.62], the convergence is uniform. ■

Since \( \succsim \) is nondegenerate, there is \( x, y \) such that \( x \succsim y \). By the representation, \( U(c, x) > U(c, y) \). Thus, there exists \( \rho, \nu \in \Delta \) such that \( U(\rho) > U(\nu) \). By linearity of \( U \),

\[
U(\mu) \geq U(\eta) \implies U(\mu \rho) > U(\eta \nu), \text{ for all } \alpha \in (0, 1).
\]  

(29)

This observation will be used in the next Lemma. Let \( \succsim_U \) be the preference relation represented by \( U \). As in Appendix A, identify any binary relation \( B \) \( \succsim \) \( \Delta \). The sequence \( \Gamma(\succsim_U) \).

Proof. First establish \( Ls \Gamma(\succsim_U) \subset \Gamma(\succsim_U) \). If \( (\mu, \eta) \in Ls \Gamma(\succsim_U) \) then there is a subsequence \( \{ \Gamma(\succsim_U(n)) \} \) and a sequence \( \{(\mu_n, \eta_n)\} \) that converges to \( (\mu, \eta) \) such that \( (\mu_n, \eta_n) \in \Gamma(\succsim_U(n)) \) for each \( n \). Therefore, for each \( n \),

\[
\varphi_{\tau(n)}(\mu_n) \geq \varphi_{\tau(n)}(\eta_n).
\]

Since \( \varphi_{\tau(n)} \) converges to \( U \) uniformly, it follows that \( U(\mu) \geq U(\eta) \). Hence \( (\mu, \eta) \in \Gamma(\succsim_U) \), as desired.

Next establish \( \Gamma(\succsim_U) \subset Ll \Gamma(\succsim_U) \). Let \( (\mu, \eta) \in \Gamma(\succsim_U) \) and take any neighborhood \( V \) of \( (\mu, \eta) \). By (29), there exists \( T < \infty \) such that \( \varphi_{\tau(n)}(\mu \rho) > \varphi_{\tau(n)}(\eta \nu) \) for all \( t \geq T \), that is, \( (\mu \rho, \eta \nu) \in \Gamma(\succsim_U) \) for all \( t \geq T \). Hence,

\[
\Gamma(\succsim_U) \neq \emptyset \text{ for all but a finite number of } t,
\]

that is, \( (\mu, \eta) \in Ll \Gamma(\succsim_U) \). This completes the proof. ■

By Lemma 40 and Theorem B.1, \( \succsim_U = \succsim^* \), that is, \( U \) is a representation of normative preference \( \succsim^* \), as desired.

This proves the Theorem for a representation \( (U, V) \) for which \( V \geq 0 \). To complete the proof, let \( (U, V) \) be any representation of \( \succsim \). Given nondegeneracy of \( \succsim \), [7, Theorem 4] implies that for any \( \alpha \) such that \( V + \alpha \geq 0 \), \( (U, V + \alpha) \) is also a representation of \( \succsim \). Hence, it follows from the preceding that \( U \) is a representation of normative preference \( \succsim^* \).

G Appendix: Proof of Theorem 3

Necessity of each axiom is either trivial or as in Appendix C. We prove sufficiency. Define \( \{ \tilde{\succsim}_t \} \) over \( Z \) as in Appendix D.1 and note that by Preferences Reverse Tomorrow, \( \tilde{\succsim}_t = \tilde{\succsim}_{t'} \) for all \( t, t' > 0 \). Let \( \tilde{\succsim}' = \tilde{\succsim}_1 \), and take \( \tilde{\succsim}' \) is the candidate CT preference. It is readily verified that \( \tilde{\succsim}' \) is nondegenerate and satisfies Order*\(^*\), Independence*, Continuity*, Set-Betweenness*, Separability* and Indifference to Timing*\(^*\). Preferences Reverse Tomorrow implies that \( \tilde{\succsim}' \) satisfies Stationarity*, and by Lemma 16, Sophistication’ implies that \( \tilde{\succsim}' \) generates \( C(\cdot) \).

Given the restrictions on \( \tilde{\succsim}' \), it can be shown that there is a representation \( W' \) such that

\[
W'(x) = \max_{\nu \in \mathbb{E}} \int_{C \times \mathbb{Z}} \left( u(c) + \delta W'(y) + v(c) + \hat{V}(y) \right) d\mu(c, y) - \max_{\eta \in \mathbb{E}} \int_{C \times \mathbb{Z}} \left( v(c) + \hat{V}(y) \right) d\eta(c, y),
\]

48
Lemma 41 \{ (c, x) \} \succ \{ (c, y) \} \implies \{ (c, x) \} \sim \{ (c, x), (c, y) \}

Proof. Follows from Menus Do Not Tempt. □

Lemma 42 There exists \( \gamma \geq 0 \) and \( \theta \) such that for all \( x \),
\[
\hat{V}(x) = \gamma W'(x) + \theta.
\]

Proof. If \( \hat{V} \) is constant (equal to some \( \theta \)), then the result follows with \( \gamma = 0 \). Therefore suppose that \( \hat{V} \) is nonconstant. Suppose by way of contradiction that \( \hat{V} \) is not ordinally equivalent to \( W' \). That is, there exists \( x \) and \( y \) such that
\[
\hat{V}(x) \geq \hat{V}(y) \text{ and } W'(x) < W'(y),
\]
or
\[
\hat{V}(x) > \hat{V}(y) \text{ and } W'(x) \leq W'(y).
\]

We show that in either case, nonconstancy of \( \hat{V} \) and \( W' \) implies the existence of menus for which both inequalities are strict. We prove this for the case that \( \hat{V}(x) = \hat{V}(y) \) and \( W'(x) < W'(y) \). The same argument can be applied to the other case, that is, \( \hat{V}(x) > \hat{V}(y) \) and \( W'(x) = W'(y) \). So suppose \( \hat{V}(x) = \hat{V}(y) \) and \( W'(x) < W'(y) \). There are two possibilities to consider. First, there is \( z \) such that \( \hat{V}(x) > \hat{V}(z) \). If \( W'(x) < W'(z) \), there is nothing to prove. If \( W'(z) \leq W'(x) \), then \( W'(z) < W'(y) \), and since \( \hat{V}(y) > \hat{V}(z) \), the assertion is proved for this case as well. Second, there is \( z \) such that \( \hat{V}(z) > \hat{V}(x) \). The argument is similar. Since \( \hat{V} \) is nonconstant, one of the two possibilities must be true, and hence we are done.

Therefore we can assume without loss of generality that
\[
W'(x) > W'(y) \text{ and } \hat{V}(x) < \hat{V}(y).
\]

Observe that
\[
W'(\{(c, x)\}) = u(c) + \delta W'(x) > u(c) + \delta W'(y) = W'(\{(c, y)\}).
\]

If \( \delta W'(x) + \hat{V}(x) > \delta W'(y) + \hat{V}(y) \), then
\[
W'(\{(c, x)\}) = u(c) + \delta W'(x) > u(c) + \delta W'(x) + \hat{V}(x) - \hat{V}(y) = W'(\{(c, x), (c, y)\}).
\]

This contradicts Lemma 41. On the other hand, if \( \delta W'(x) + \hat{V}(x) \leq \delta W'(y) + \hat{V}(y) \), then
\[
W'(\{(c, x)\}) = u(c) + \delta W'(x) > u(c) + \delta W'(y) = W'(\{(c, x), (c, y)\}),
\]

again contradicting Lemma 41. Thus we establish that \( \hat{V} \) is ordinally equivalent to \( W' \). Since \( \hat{V} \) and \( W' \) are also linear, it follows that they must be cardinaly equivalent. Thus, there exists \( \gamma \geq 0 \) and \( \theta \) such that for all \( x \),
\[
\hat{V}(x) = \gamma W'(x) + \theta,
\]
as was to be shown. □

\[\text{28 A } \hat{V}-\text{best and worst menu exists since } \hat{V} \text{ is continuous and } Z \text{ is compact.}\]
Without loss of generality, we can set $\theta = 0$. Therefore, $\succsim'$ is represented by the function defined by:

$$W'(z) = \max_{\mu \in z} \int_{C \times Z} \left( u(c) + \delta W'(x) + v(c) + \gamma W'(x) \right) d\mu(c, x)$$

$$- \max_{\eta \in z} \int_{C \times Z} \left( v(c) + \gamma W'(y) \right) d\eta(c, y).$$

It remains to show that $\succsim'$ is the unique CT preference that generates $C(\cdot)$. Note that for any $\mu, \eta, (c, x) \in C(\{(c, x), (c, y)\})$

$$\iff u(c) + \delta W'(x) + v(c) + \gamma W'(x) \geq u(c) + \delta W'(y) + v(c) + \gamma W'(y)$$

$$\iff W'(x) \geq W'(y)$$

$$\iff x \succsim' y.$$ 

Therefore, if two CT preferences generate the same choice correspondence, then they must coincide.

H Appendix: Proof of Theorem 4

The argument used in the proof of Theorem 2 goes through. The only modification is that each $\succsim_t$ defined there is now represented by $\varphi^*_t : \Delta(C \times Z) \to \mathbb{R}$ where for all $\mu \in \Delta$, $\varphi^*_t(\mu) = U(\mu)$.

References


51