Temptation, Welfare and Revealed Preference[*]

Jawwad Noor

Department of Economics,

Boston University,

270 Bay State Road, Boston MA 02215.

Email: jnoor@bu.edu.

September 28, 2005

---

[*] I am greatly indebted to Larry Epstein for his guidance and encouragement throughout the project. I have benefitted from comments from Paulo Barelli, Barton Lipman, Josef Perktold, Alvin Roth, Alvaro Sandroni, William Thomson, Gábor Virág, seminar participants at Rochester, Yale, Iowa and Boston University, and participants at the Canadian Economic Theory Conference (May 2004), the Risk, Uncertainty and Decisions Conference (June 2004), and the Midwest Economic Theory Meeting (Nov 2004). Any errors are my responsibility.

Abstract

Choice may be determined both by a consideration of one's welfare (normative preference) and by desires (temptation preference). To provide foundations for such a theory, Gul and Pesendorfer [10, 11] adopt a preference over choice problems as a primitive and hypothesize that temptation creates a preference for commitment. This paper argues that temptation may in fact create the *absence* of a preference for commitment, and that the primitive may not be empirically meaningful since it requires us to observe behavior in the absence of temptation. An alternative approach to providing foundations is introduced. Motivated by the 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, and temptation preference is inferred from discrepancies between normative preference and choice. With a choice correspondence as the primitive, agents who are 'tempted not to commit' are modeled. The foundations of the model are used to identify evidence supporting such temptation.

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

JEL classification number: D11, D60.

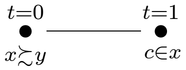
1. Introduction

An agent may be breaking his diet, taking drugs, making an expensive purchase, etc. while telling himself that he really should not. Such instances suggest that choice is determined not by one, but two preference orderings: a *temptation preference* that captures the agent's desires, and a *normative preference* that captures his view of what choices he *should* make, his view of what is best for his welfare. Choice behavior is the outcome of an aggregation of temptation preference and normative preference. The agent is said to experience temptation when his desires conflict with his judgment regarding the best course of action, that is, when his temptation preference conflicts with his normative preference. He is said to have self-control problems when he cannot always resist the desires he judges to be 'bad', that is, when his choices do not necessarily respect his normative preference.

Normative preference embodies the criteria used by the agent to judge his own welfare, and therefore, arguably, it is the appropriate guide for welfare policy. This is in contrast with the traditional view that emphasizes choice as the appropriate guide for welfare policy. Thus, the notions of normative preference and temptation are of interest, and it is worthwhile to ask: what are 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.

1.1. The Commitment Approach

Gul and Pesendorfer (henceforth GP) are the first to provide foundations for a model of temptation and self-control [10, 11]. The primitive of their model is a preference $\succsim$ over Z, the space of menus (choice problems). 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:



The preference $\succsim$ dictates the agent's choice of menu in period 0. After choosing a menu x, he subsequently makes some 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 period 0 behavior $\succsim$ contains information about both the agent's normative and temptation preferences over C.[1]

Specifically, GP hypothesize that *temptation creates a preference for commitment* in period 0.[2] To illustrate this 'commitment approach', consider an dieter who normatively prefers having a salad s for dinner, but anticipates being tempted to have a burger b if it is available. GP suggest that, in order to avoid temptation, this dieter will commit to having s for dinner by, say, going to a restaurant that serves only s rather than one that also serves b:

$$\{s\} \succ \{s, b\}.$$

That is, period 1 temptation leads to a period 0 preference for commitment. This choice in turn reveals to an observer the dieter's normative preference for s and his temptation preference for b.

This story rests on an assumption that we shall argue is problematic: the commitment approach assumes 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" [11, p 129]. This assumption is crucial. To see why, imagine that although the dieter understands the value of commitment, he is nevertheless *tempted by the restaurant that serves burgers* – in period 0 he is tempted to choose $\{s, b\}$ over $\{s\}$. This could be because the idea of having a burger b at dinner time tempts the

---

[1] What 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 [11, p 120].

[2] Also see Strotz [23], Ainslie [2], Laibson [15] and O'Donoghue and Rabin [20].

dieter now, so that he is tempted by the opportunity to indulge temptations later. Indeed, if this temptation is strong enough, the dieter does not choose commitment:

$$\{s, b\} \succsim \{s\}.$$

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. In particular, the primitive $\succsim$ of the approach is not just any ranking of menus, but rather *an agent's ranking of menus in the absence of temptation.*[3] However, this immediately raises several questions: Is $\succsim$ observable? Is it possible to distinguish between a ranking of menus that is in the absence of temptation and one that is not? Even if we could distinguish the two, what if we only observe rankings that are subject to temptation? If it is indeed plausible that menus may tempt (that is, that a dieter may find a restaurant tempting), observing $\succsim$ would require us to deduce how agents would rank menus *if* they did not experience temptation. But, can this be done?

To the extent that the preference $\succsim$ is not observable, adopting it as the primitive of a model of temptation is problematic. The purpose of behavioral foundations 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 for the characterizations to be in terms of observable behavior.

1.2. The Preference Reversal Approach

This paper suggests 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. Let C

---

[3] This is equivalent to saying that GP's primitive is a *normative preference over menus.* This observation will be exploited in this paper.

be a consumption set. For any $c \in C$, let c^{+t} represent an alternative that provides c in t periods, and let the ranking of such alternatives be captured by a preference ordering $\gtrapprox$, the asymmetric and symmetric part of which is $>$ and $\approx$, respectively.

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:

$$\begin{aligned} 20^{+0} &> 30^{+1} \\ 20^{+1} &< 30^{+2}, \end{aligned}$$

where a unit of time is, say, 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 [21] and Rachlin and Green [22] 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 reveal that *delayed temptations are easier to resist than immediate temptations.*

This suggests a way of deriving an agent's normative preference from his choices. From the agent's ranking $\gtrapprox$ of delayed consumption alternatives, derive a set of preference relations $\{\gtrapprox_t\}_{t=0}^{\infty}$ over C as follows: for each $c, \widehat{c} \in C$ and $t \geq 0$,

$$c \gtrapprox_t \widehat{c} \Longleftrightarrow c^{+t} \gtrapprox \widehat{c}^{+t}$$

Thus, each $\gtrapprox_t$ 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 $\gtrapprox_t$ of alternatives diminishes. That is, as t grows, the temptation component underlying $\gtrapprox_t$ becomes less significant, and so, each $\gtrapprox_t$ provides an increasingly better approximation of the agent's underlying normative preference. For this reason, it is intuitive to identify normative preference with the (appropriately defined) limit:

$$\lim_{t\to\infty} \gtrapprox_t .$$

This serves as a behavioral definition 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 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 constitutes choices that may well be subject to temptation. In contrast, the commitment approach 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). Consider the possibility that agents may be tempted also by future consumption, that is, they may be tempted also by opportunities of consuming tempting items in the future (refer to such agents as Future Temptation or FT agents).

One reason to be interested in FT agents is as follows: CT agents always take advantage of commitment opportunities in order to deal with their self-control problems. They do not care for temptations that lie in the future, and therefore have no reason not to commit. However, as we saw in Section 1.1, temptation by future consumption can induce agents with self-control problems to avoid taking advantage of commitment opportunities. When tomorrow's drugs tempt an addict today, he may be tempted not to commit to abstinence, that is, he may be tempted to retain the possibility of consuming drugs later. This is reminiscent of the problem of non-compliance with commitment-based treatment procedures among addicts: disulfiram-based treatments for alcoholics and naltexrone-based treatments for heroin and morphine addicts are known to be of limited effectiveness, primarily because patients do not comply with the disulfirum or naltexrone regimen despite exhibiting a desire for the treatment [2, 8, 9].

The fact that self-control problems may lead CT agents to seek commitment and FT agents to avoid it suggests that the implications of temptation by future consumption may be worth studying. But a prior question, which is the focus here, is whether there exists any evidence of such temptation. We use the ideas in Section 1.2 to characterize CT and FT agents in terms of observable behavior. By contrasting the axioms that characterize 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 consumption. The result of our analysis is that supporting evidence does indeed exist (Section 6 presents the evidence). The peculiarities of temptation by future consumption include its implications for preference reversals and the demand for mechanisms that ensure commitment in the future. Evidence supporting such behavior comes from the preference reversals literature and from experiments on saving behavior.

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 $\succsim$ over menus as the primitive, we adopt as the primitive a choice correspondence $\mathcal{C}$ that describes choices *from* menus. These choices are possibly subject to temptation, and arguably are observable in principle. We find restrictions on the choice correspondence $\mathcal{C}$ that are necessary and sufficient for it to be 'generated' or 'rationalized' by a GP-style preference $\succsim$, in a sense made precise in Section 2. This $\succsim$ 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 $\succsim$ by the commitment approach is related to the normative preference derived from $\mathcal{C}$ by our approach.

The remainder of the paper is organized as follows. Section 2 provides formal details of our model of Future Temptation, and Section 3 presents axioms and the representation result. Section 4 outlines the proof of the representation theorem and shows that in the model, GP's and our approach yields the same normative preference. Thus, our approach, in a sense, our approach provides a characterization of normative preference that is dual to GP's, and which is based on observable behavior. A model of Current Temptation is axiomatized in Section 5. Evidence supporting temptation by future consumption is discussed in Section 6. Section 7 concludes. All proofs are collected in appendices.

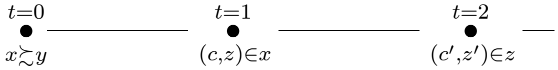
2. A Model of Future Temptation

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' that arises from a special period 0 and conclude by describing a way to avoid it.

For any compact metric space X, $\Delta(X)$ denotes the set of all probability measures on the Borel σ-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 a compact metric space C. 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 [11] for the formal definition of Z. We often let Δ denote $\Delta(C \times Z)$. The reader should take note of this.

Adopt a binary relation $\succsim$ over Z as the primitive, and consider the following time-line.



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.[4] 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 $W : Z \longrightarrow \mathbb{R}$ of the following form: there exist δ and γ, $0 < \gamma < \delta < 1$, continuous functions $u, v : C \longrightarrow \mathbb{R}$, and continuous linear functions $U, V : \Delta(C \times Z) \longrightarrow \mathbb{R}$ and $\overline{V} : Z \longrightarrow \mathbb{R}$ such that for all $x \in Z$,

$$W(x) = \max_{\mu \in x}\{U(\mu) + \Big(V(\mu) - \max_{\eta \in x} V(\eta)\Big)\}, \tag{2.1}$$

$$\begin{aligned} \text{where } U(\mu) &= \int_{C \times Z} (u(c) + \delta W(y))\, d\mu(c, y), \\ V(\mu) &= \int_{C \times Z} \big(v(c) + \gamma \overline{V}(y)\big)\, d\mu(c, y), \\ \overline{V}(x) &= \max_{\eta \in x} V(\eta). \end{aligned}$$

---

[4] More generally, if the chosen alternative is a nondegenerate lottery $\mu \in \Delta$, 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.

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.[5] 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, (2.1) states that the utility $W(x)$ of a menu x is the maximum value of normative utility net of self-control cost.

Though $W(\cdot)$ represents the agent's choice *of* menu in some period 0, it is suggestive of how the agent makes his choice *from* a menu: it suggests that he maximizes normative utility net of self-control costs. Since the term '$\max_{\eta \in x} V(\eta)$' in (2.1) is a constant when x is given, maximizing over μ in x essentially maximizes

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

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

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 δ. Temptation utility $V(\cdot)$ depends on current consumption and the temptation value $\overline{V}(\cdot)$ of a continuation menu discounted by γ. As the form of $\overline{V}(\cdot)$ shows, a continuation menu is 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 [11], which exhibit temptation by immediate consumption only. DSC preferences are not a special case of FT preferences since the latter require $\gamma > 0$.

---

[5] GP call it 'commitment utility'.

To summarize, the agent's period 0 ranking of menus is represented by $W(\cdot)$ and his choice from menus in each subsequent period $t > 0$ is captured by the choice correspondence $\mathcal{C}(\cdot, \succsim)$ over Z defined by

$$\mathcal{C}(z, \succsim) = \arg\max_{z \in x} \{U(\mu) + V(\mu)\}, \tag{2.2}$$

for all $z \in Z$.[6] Each alternative μ in a menu z yields normative and temptation utility, and the agent's choice from the menu tries to balance the two. For a simple illustration, consider a menu $\{(c, x), (c, y)\}$ where there is no choice of immediate consumption, the continuation menu x gives the opportunity to indulge temptations later, and the continuation menu y offers commitment (imagine that y corresponds to entering rehabilitation and that x corresponds to not entering rehabilitation). Since x contains more tempting items than y, the functional form of $V(\cdot)$ implies that the temptation utility of (c, x) is higher than that of (c, y). On the other hand, the functional form of $U(\cdot)$ implies that the normative utility of (c, y) is higher than that of (c, x) – normative utility $U(\cdot)$ evaluates continuation menus according to $W(\cdot)$, and the presence of temptation in a menu reduces its value according to $W(\cdot)$.[7] Thus, normative and temptation preferences disagree over the choice from $\{(c, x), (c, y)\}$, that is, the agent experiences temptation (by (c, x)). The eventual choice lies in $\mathcal{C}(\{(c, x), (c, y)\}, \succsim)$.

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. Begin by recalling that the period 0

---

[6] Though $\mathcal{C}(\cdot, \succsim)$ is defined in terms of one representation of $\succsim$, it does not depend on the particular representation. To see this, call (U, V) a representation of $\succsim$. If $\succsim$ exhibits a preference for commitment ($x \succ x \cup y$ for some x, y) then by [10, Thm 4], (U', V') is another representation of $\succsim$ if and only if there exists $\alpha > 0$ and β_U, β_V such that $U' = \alpha U + \beta_U$ and $V' = \alpha V + \beta_V$. If $\succsim$ does not exhibit a preference for commitment, then without loss of generality, $V = 0$, and U is unique up to an affine transformation. In either case, any representation of $\succsim$ would give rise to the same $\mathcal{C}(\cdot, \succsim)$.

[7] Note that temptation utility enters $W(\cdot)$ only in the form of a cost.

ranking $\succsim$ of menus is represented by

$$W(\cdot). \tag{2.3}$$

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 $\mathcal{C}(\{(c, x), (c, y)\}, \succsim)$, that is, by solving

$$\max_{\{(c,x),(c,y)\}} \{u(\cdot) + \delta W(\cdot) + v(\cdot) + \gamma \overline{V}(\cdot)\}.$$

Therefore, the period $t > 0$ ranking of menus is represented by

$$W(\cdot) + \frac{\gamma}{\delta} \overline{V}(\cdot). \tag{2.4}$$

Compare (2.3) and (2.4) and conclude that, in general, the ranking of menus in period 0 is different from that in any period $t > 0$.[8] Since $\overline{V}(\cdot)$ captures temptation utility from menus, it is evident that the agent's ranking of menus in period 0 is not subject to temptation (that is, period 0 is special), whereas temptation affects his ranking in all subsequent periods.

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? In fact, can we even 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

---

[8] Though $\overline{V}(\cdot)$ has a specific functional form in this example, the argument reveals that such a discrepancy arises generically in extensions of DSC preferences that permit temptation by menus.

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 $t > 0$ choice *from* menus, instead of period 0 choice *of* menus. These period $t > 0$ choices are potentially 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 $\mathcal{C}(\cdot)$ over Z. For any choice correspondence $\mathcal{C}(\cdot)$ and any FT preference $\succsim$, say that $\succsim$ *generates* $\mathcal{C}(\cdot)$ if,

$$\mathcal{C}(\cdot) = \mathcal{C}(\cdot, \succsim),$$

where $\mathcal{C}(\cdot, \succsim)$ is defined by (2.2). Our problem is to find restrictions on $\mathcal{C}(\cdot)$ that imply the existence of an FT preference $\succsim$ that generates $\mathcal{C}(\cdot)$. That is, under what conditions on $\mathcal{C}(\cdot)$ 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? These restrictions on $\mathcal{C}(\cdot)$ allow us to obtain the FT preference $\succsim$ and the special period 0 as parts of a representation *result*. In particular, we need *not assume* the existence of such a period nor take such a preference as the primitive.

It is also noteworthy that such restrictions on $\mathcal{C}(\cdot)$ 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 Δ, and for each preference $\gtrapprox$ in this class we define a choice correspondence $\mathcal{C}^*(\cdot, \gtrapprox)$ over $\mathcal{K}(\Delta)$ by

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

---

[9] GP note that DSC preferences do not suffer from this problem: when only immediate consumption is tempting ($\gamma = 0$), problems (2.3) and (2.4) are identical.

for each $x \in \mathcal{K}(\Delta)$. 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 $\mathcal{C}(\cdot)$ in order for there to exist a preference $\gtrapprox$ over Δ such that

$$\mathcal{C}(\cdot) = \mathcal{C}^*(\cdot, \gtrapprox).$$

In an analogous fashion, we are starting with a class of preferences (namely, FT preferences), defining an appropriate choice correspondence $\mathcal{C}(\cdot, \succsim)$, 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 $\mathcal{K}(\Delta)$ rather than over the space of alternatives Δ.

The primitive of the model is a choice correspondence $\mathcal{C}(\cdot)$ 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 in Section 4, such a derivation is achieved by exploiting the ideas contained in Section 1.2. In particular, we obtain the preference $\succsim$ in the form of *a normative preference over menus.*

3. Axioms and Representation Result

Generic elements of Z are x, y, z whereas generic elements of Δ are μ, η, ν. 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 σ-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 $\mathcal{C} : Z \rightsquigarrow \Delta(C \times Z)$ where, for all $x \in Z$, $\mathcal{C}(x) \neq \phi$ and $\mathcal{C}(x) \subset x$. 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 $x^{+1} \equiv (\bar{c}, x)$ and inductively for $t > 1$, $x^{+t} = (\bar{c}, x^{+(t-1)})$. Then $x^{+t} \in \Delta$ 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 μ^{+t} and identify μ^{+0} with μ. 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 Δ.

- 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 $\gtrapprox$ denote the revealed preference relation (defined on Δ) that is generated by choices from binary menus, that is,

$$\mu \gtrapprox \eta \Longleftrightarrow \mu \in \mathcal{C}(\{\mu, \eta\}). \tag{3.1}$$

The indifference relation $\approx$ and the strict preference relation $>$ are derived from $\gtrapprox$ in the usual way.

Consider the following axioms on $\mathcal{C}(\cdot)$. 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 $\mu, \eta \in x \cap y$, $\mu \in \mathcal{C}(x)$ and $\eta \in \mathcal{C}(y)$, then $\mu \in \mathcal{C}(y)$.*

This is the familiar Weak Axiom of Revealed Preference. It is a minimal consistency requirement on choices. The axiom states that for μ and η that are both contained in menus x and y, if μ is revealed preferred to η (that is, if it is chosen from x) then η cannot be revealed preferred to μ (that is, if it is chosen from y then μ must have also been chosen). 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 salad, b a burger, and B a large burger. The agent normatively prefers s to b and b to B, but finds B more tempting than b, and b more tempting than s. For simplicity, assume that $\mathcal{C}(\cdot)$ is single-valued. The following choices violate WARP:

$$\{s\} = \mathcal{C}(\{s, b\}) \text{ and } \{b\} = \mathcal{C}(\{s, b, B\}).$$

Yet, the choices are not unreasonable. Since b is not so tempting, he is able to apply self-control and choose s out of $\{s, b\}$. But when faced with $\{s, b, B\}$, the presence of B whets his appetite for a burger. In order to compromise between his craving for the large burger B and his normative preference for s, he settles for b.

This example is one where self-control is menu-dependent – the presence of B affects the agent's ability to resist b. 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 $\{s, b, B\}$ was different from b, that is, if either $\{B\} = \mathcal{C}(\{s, b, B\})$ or $\{s\} = \mathcal{C}(\{s, b, B\})$. These cases capture the following stories: In the first case, B is so tempting that it cannot be resisted, and in the second case, B is not that much more tempting than b, and thus can be resisted along with b.

Axiom 2 (Continuity). *$\mathcal{C}(\cdot)$ is upper hemicontinuous.*

Upper hemicontinuity of $\mathcal{C}(\cdot)$ 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.

Axiom 3 (Independence). $\mu > \eta \Longrightarrow \mu\alpha\nu > \eta\alpha\nu$.

This is the familiar Independence axiom.

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

$$(\frac{1}{2}(c, x) + \frac{1}{2}(c', x'))^{+t} \approx (\frac{1}{2}(c, x') + \frac{1}{2}(c', x))^{+t}.$$

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.

---

[10] Formally, upper hemicontinuity implies that if $\{x_n\}$ is a sequence of menus converging to x, and $\mu_n \in \mathcal{C}(\{x_n\})$ for each n, then the sequence $\{\mu_n\}$ has a limit point in $\mathcal{C}(\{x\})$.

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 α 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 $t > 0$,*

$$x^{+t} \gtrapprox y^{+t} \Longrightarrow x^{+t} \gtrapprox (x \cup y)^{+t} \gtrapprox y^{+t}.$$

Moreover, there exists x, y such that $x^{+1} > (x \cup y)^{+1} > y^{+1}$.

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 μ reveals that η 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 η and so, the agent succumbs to temptation. On the other hand, the ranking $\{\mu, \eta\}^{+t} > \{\eta\}^{+t}$ would suggest that μ 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 μ. The story is that the reward μ 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 μ. The indifference between $\{\mu\}^{+t}$ and

$\{\mu, \eta\}^{+t}$ is another expression of the overwhelming temptation by μ – he foresees that he will choose μ 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 are menus x and y such that the agent anticipates experiencing temptation and exerting self-control in $x \cup y$.

Axiom 7 (Sophistication). *If* $\{\mu\}^{+t} > \{\eta\}^{+t}$*, then*

$$\{\mu, \eta\}^{+t} > \{\eta\}^{+t} \Longleftrightarrow \mu > \eta.$$

As the name suggests, this axiom connects the agent's expectation of his future choices with his actual choices.[11] The left-hand-side ranking reveals that the agent anticipates choosing μ if, after t periods, he faces the menu $\{\mu, \eta\}$. Since $\mathcal{C}(\cdot)$ is time-invariant, the actual choice from $\{\mu, \eta\}$ after t periods is given by $\mathcal{C}(\cdot)$. Thus, $\mu > \eta$ says that after t periods, the agent's actual choice from $\{\mu, \eta\}$ is μ. That is, the axiom says that the agent's anticipated choice and actual choice coincide.

To see that the hypothesis that $\{\mu\}^{+t} > \{\eta\}^{+t}$ is required in order to to interpret the axiom as one capturing sophistication, suppose that it is dropped. Then $\mu \approx \eta$ is consistent with the possibility that $\{\eta\}^{+t} > \{\mu, \eta\}^{+t} \approx \{\mu\}^{+t}$, which does not preclude an expectation to strictly prefer μ when choosing from $\{\mu, \eta\}$ after t periods.

Axiom 8 (Menus Can Tempt).

$$x^{+t} > (x \cup y)^{+t} \text{ for some } t \Longleftrightarrow \{(c, x)\}^{+t'} > \{(c, x), (c, y)\}^{+t'} \text{ for some } t'.$$

The axiom embodies the idea that future consumption tempts the agent in the form of a temptation by menus – the presence of tempting alternatives in a menu

---

[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.

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 we observe y being tempting if and only if we also observe it containing tempting alternatives.

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

This axiom imposes the structure of preference reversals on $\mathcal{C}(\cdot)$. 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.

As a simple consequence of Reversal we obtain a function $\tau : \Delta \times \Delta \to \mathbb{R}$ such that for any pair of rewards μ, η, $\tau(\mu, \eta)$ is the number of periods that μ and η need to be delayed before a preference reversal is observed; if no reversal is observed, then $\tau(\mu, \eta) = 0$. For instance, if $\mu > \eta$, $\mu^{+1} \gtrapprox \eta^{+1}$ and $\mu^{+t} < \eta^{+t}$ for all $t > 1$, then $\tau(\mu, \eta) = 2$. For a precise definition of the function τ, see Appendix B.

Before stating the final axiom, we make an observation that will assist us in its interpretation. Suppose that an agent exhibits $\mu > \eta$ for some μ, η. What can we expect to observe if this ranking respects normative preference? We suggest that the function τ will satisfy the following property $(*)$ at (μ, η):

$$\tau(\mu, \eta) = 0 \text{ and } \tau \text{ is continuous at } (\mu, \eta). \tag{$*$}$$

To understand why, first note that the ranking $\mu > \eta$ respects normative preference only if the choice between μ and η is not subject to overwhelming temptation, that is, only if either there is no temptation, or η is tempting but not overwhelmingly so. Also recall that a preference reversal occurs when the choice between two rewards is overwhelmed by temptation – we observe a preference reversal when pushing rewards into the future makes resistible an otherwise overwhelming temptation. If the choice between μ and η is not overwhelmed by temptation, then pushing μ and η into the future should not affect their ranking. That is, if the choice between μ and η reflects normative preference, we should observe no preference reversal, $\tau(\mu, \eta) = 0$. Furthermore, if temptation does not overwhelm the choice between the pair of rewards (μ, η), then under the presumed continuity of underlying temptation and normative preference, it should not overwhelm the choice between *neighboring* pairs of rewards either. That is, we should expect to observe no preference reversal for any pairs of rewards in some neighborhood of (μ, η).[12] This condition is equivalent to the statement that τ is continuous at (μ, η).

Now we state the final axiom.

Axiom 10 (Commitment is Normative). *If* $x^{+t} > (x \cup y)^{+t}$*, then* τ *satisfies property (*∗*) at* $(x^{+t}, (x \cup y)^{+t})$*. Moreover, if* $x^{+t} > (x \cup y)^{+t} > y^{+t}$*, then* τ *satisfies property (*∗*) at* $((x \cup y)^{+t}, y^{+t})$.

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 $(x \cup y)^{+t} \gtrapprox x^{+t}$. Hence, a t period delay makes the temptation by $x \cup y$ resistible. Consequently, the choice between x^{+t} and $(x \cup y)^{+t}$ is not overwhelmed by temptation, and thus, their ranking reflects normative preference. Given the above discussion, this in turn should imply the noted restrictions on the function τ at the point $(x^{+t}, (x \cup y)^{+t})$. To understand the second part of the axiom, observe that the ranking $x^{+t} > (x \cup y)^{+t}$ also 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

---

[12] In other words, for any sequence $\{(\mu_n, \eta_n)\}$ converging to (μ, η), there is some N such that $n \geq N$ implies $\tau(\mu_n, \eta_n) = 0$.

$(x \cup y)^{+t}$ and y^{+t} is not overwhelmed by temptation, and thus, the ranking of the rewards reflects normative preference as well.

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[13]

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

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

$$\mathcal{C}(\cdot) = \mathcal{C}(\cdot, \succsim),$$

where $\mathcal{C}(\cdot, \succsim)$ is defined by (2.2). Theorem 3.1 is the main result of this paper.

Theorem 3.1. *A choice correspondence* $\mathcal{C}(\cdot)$ *satisfies Axioms 1-10 if and only if there exists an FT preference* $\succsim$ *that generates it. Furthermore, each such* $\mathcal{C}(\cdot)$ *is generated by a unique FT preference* $\succsim$*.*

Theorem 3.1 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. An equivalent restatement of the Theorem is: a choice correspondence $\mathcal{C}(\cdot)$ satisfies Axioms 1-10 if and only if there exist functions $U(\cdot)$ and $V(\cdot)$ as in the representation of an FT preference such that for any menu $x \in Z$,

$$\mathcal{C}(x) = \arg\max_{\mu \in x} \{U(\mu) + V(\mu)\}.$$

Note that $U(\cdot) + V(\cdot)$ represents the order $\gtrapprox$ defined by (3.1).

The functions $U(\cdot)$ and $V(\cdot)$ constitute a decomposition of the agent's choices into its respective normative and temptation component; the Theorem assures us that

---

[13] In terms of the representation, this is equivalent to the condition that $U(\cdot)$ and $V(\cdot)$ are nonconstant and $U(\cdot)$ is not an affine transformation of $V(\cdot)$, where $U(\cdot)$ and $V(\cdot)$ are as in the representation of FT preferences. See Lemma C.2.

the normative and temptation preference delivered by the decomposition are unique, since two distinct FT agents cannot share the same choice correspondence.[14] The next section outlines the proof of the Theorem and justifies our interpretation of $U(\cdot)$ as normative utility. This in turn justifies our interpretation of $V(\cdot)$ as temptation utility: since choice maximizes $U(\cdot)+V(\cdot)$, the function $V(\cdot)$ captures the component of choice that, in some sense, is not 'explained' by normative preference.

We have provided foundations for the FT model that are an alternative to those in [18], where a preference over menus was taken as the primitive. 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 3.1. The construction is based on the behavioral definition of normative preference presented in the Introduction. This section also shows that this definition allows us to interpret $U(\cdot)$ as normative utility.

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

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

$$\mu \gtrapprox_t \eta \Longleftrightarrow \mu^{+t} \in \mathcal{C}(\{\mu^{+t}, \eta^{+t}\}).$$

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

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

---

[14] Also see the uniqueness properties of the FT representation in footnote 6.

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

To be formal about the definition of $\succsim^*$, say that a binary relation B on Δ is nonempty if $\mu B \eta$ for some $\mu, \eta \in \Delta$. Following Hildenbrand [12], identify any nonempty continuous binary relation on Δ with its graph, a nonempty compact subset of $\Delta \times \Delta$. Thus, the space of nonempty continuous preferences on Δ 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. Think of $\{\gtrapprox_t\}$ as a sequence in $\mathcal{P}$.

Definition 4.1. *The normative preference* $\succsim^*$ *over* Δ *is the limit of* $\{\gtrapprox_t\}$.

The existence of normative preference obtains under WARP, Continuity and Reversal alone. Reversal requires that for every $\mu, \eta \in \Delta$, the post-preference reversal preferences agree on the ranking of μ and η. This implies that for every $\mu, \eta \in \Delta$, there exists T such that all preferences $\gtrapprox_t$, $t \geq T$, agree on the ranking of μ and η. In a sense, the difference between $\gtrapprox_t$ and $\gtrapprox_{t+1}$ decreases as t grows, since $\gtrapprox_t$ and $\gtrapprox_{t+1}$ agree on the ranking of more and more pairs of rewards. This provides intuition for why normative preference exists.

Step 3: *Derive* $\succsim$ *over* Z.

The preference $\succsim$ in Theorem 3.1 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 $\succsim$ is in fact the agent's normative preference over menus. That is, $\succsim$ is simply the ranking over Z induced by $\succsim^*$:

$$x \succsim y \Longleftrightarrow x^{+1} \succsim^* y^{+1},$$

for all $x, y \in Z$. Thus, we say $x \succsim y$ if and only if normative preference ranks the x^{+1} higher than y^{+1}. Recall that x^{+1} and y^{+1} are degenerate lotteries in Δ that yield some common immediate consumption $\bar{c}$ and the respective menus x and y in the next period.

Theorem 3.1 establishes that under Axioms 1-10 imposed on $\mathcal{C}(\cdot)$, the preference $\succsim$ is a well-defined FT preference. The latter is verified by checking that $\succsim$ satisfies the axioms in [18]. Commitment is Normative and Sophistication play key roles in establishing that $\succsim$ generates $\mathcal{C}(\cdot)$. To see this, recall from the discussion following Set Betweenness that the choices an agent anticipates making from menus is revealed by a ranking of menus. That is, a ranking of menus generates an 'anticipated choice correspondence'. Like $\succsim^*$, each $\gtrapprox_t$ induces a ranking over Z, and thus generates a choice correspondence. Sophistication ensures that each $\gtrapprox_t$ generates the same choice correspondence, namely, $\mathcal{C}(\cdot)$. Commitment is Normative ensures that $\succsim$ generates the same choice correspondence as each $\gtrapprox_t$. The assertion follows.

This completes the derivation of $\succsim$. Now turn to the justification of our interpretation of the function U as normative utility.

For any FT preference $\succsim$, a normative preference $\succsim^*$ is said to be elicited from $\mathcal{C}(\cdot, \succsim)$ if it is derived as follows: First, a choice correspondence $\mathcal{C}(\cdot, \succsim)$ is derived from $\succsim$ by defining it as in (2.2). Then, a set of preference relations $\{\gtrapprox_t\}_{t=0}^{\infty}$ defined over Δ is obtained from $\mathcal{C}(\cdot, \succsim)$ as in Step 1 above, and finally, $\succsim^*$ is defined as in Step 2.

Call (U, V) a representation of the FT preference $\succsim$ if U and V are as in the FT representation (2.1). Theorem 4.2 establishes that U represents the normative preference $\succsim^*$ elicited from $\mathcal{C}(\cdot, \succsim)$.

Theorem 4.2. *If (U, V) represents an FT preference $\succsim$, and $\succsim^*$ is the normative preference elicited from $\mathcal{C}(\cdot, \succsim)$, then U represents $\succsim^*$.*

Put differently, if we start with a choice correspondence $\mathcal{C}(\cdot)$ and derive a normative preference $\succsim^*$ from it, and if this choice correspondence is generated by an

FT preference $\succsim$ with representation (U, V), then U represents $\succsim^*$. This justifies our referring to U as normative utility. However, Theorem 4.2 is of interest also for other reasons. Recall from Section 2 that for any $\mu \in \Delta$, $U(\mu) = W(\{\mu\})$, where W is defined by (2.1). Thus, for any $\mu, \eta \in \Delta$,

$$U(\mu) \geq U(\eta) \Longleftrightarrow \{\mu\} \succsim \{\eta\}.$$

That is, U represents the agent's preference under commitment in a special period 0. GP [11] informally interpret U as capturing the agent's view of his long-run self-interest – they identify normative preference with commitment preference. In terms of the example of the dieter in the Introduction, the dieter is said to have a normative preference for a salad s according to GP's definition if he would commit to s rather than to b in a special period 0. Theorem 4.2 provides formal justification for such a definition of normative preference, albeit in the context of the FT model.

Theorem 4.2 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 1.1 and 2 about the non-observability of the period 0 preference $\succsim$ implies that GP's definition of normative preference may be based on unobservables. Theorem 4.2 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 $\mathcal{C}(\cdot)$ as before. The axioms imposed on $\mathcal{C}(\cdot)$ coincide with those of the FT model, 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,

$$\{(c,x)\}^{+t} > \{(c,x),(c,y)\}^{+t},$$

for some x, y and t. Such a ranking expresses a *demand for delayed commitment*: Under $\{(c,x),(c,y)\}^{+t}$, the agent has the opportunity to decide after t periods whether to face x in the $(t+1)^{th}$ period or y. Under $\{(c,x)\}^{+t}$, 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: y contains tempting alternatives. Since future temptations affect the agent today, he is tempted by the menu y, and thus he would rather avoid having to choose between x and y. The following axiom rules out a demand for delayed commitment.

Axiom 11 (Menus Do Not Tempt). *For all* c, x, y *and* $t > 0$,

$$\{(c,x),(c,y)\}^{+t} \gtrsim \{(c,x)\}^{+t}.$$

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 12 (Preferences Reverse Tomorrow). *For all* x, y *and* $t > 0$,

$$x^{+1} \gtrsim y^{+1} \Longleftrightarrow x^{+t} \gtrsim y^{+t}.$$

The axiom strengthens (the first part of) Reversal by imposing a stationarity property on the ranking of *any* pair of delayed alternatives.[15] Thus, if $x^{+1} \gtrsim y^{+1}$,

---

[15] Given Preferences Reverse Tomorrow, Axioms 5-7 and 11 can be weakened by restricting their statements to hold only for $t = 1$, and Axiom 10 can be dropped altogether. In order to facilitate comparison with the FT model, we retain the stronger statements.

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 μ, η, where τ 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 μ is overwhelmingly tempting. When both μ and η are delayed by one period, they become future rewards, and since the agent is tempted only by immediate rewards, he experiences no temptation by μ^{+1} and so reverses his preferences: $\mu^{+1} < \eta^{+1}$.

Turn to the representation. Let $\succsim^{CT}$ be a binary relation over Z that is represented by some function $W^{CT} : Z \longrightarrow \mathbb{R}$. Say that $\succsim^{CT}$ 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 $u, v : C \longrightarrow \mathbb{R}$ such that for all $x \in Z$,

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

$$\begin{aligned} \text{where } U(\mu) &= \int_{C\times Z} \big(u(c) + \delta W^{CT}(y)\big)\, d\mu(c, y) \\ V(\mu) &= \int_{C\times Z} \big(v(c) + \gamma W^{CT}(y)\big)\, d\mu(c, y). \end{aligned}$$

CT preferences differ from FT preferences in two ways. First, γ 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* $\mathcal{C}(\cdot)$ if

$$\mathcal{C}(x) = \arg\max_{\mu \in x}\{U(\mu) + V(\mu)\}.$$

Theorem 5.1 is the main result of this section.

Theorem 5.1. *A choice correspondence $\mathcal{C}(\cdot)$ satisfies Axioms 1-7 and 10-12 if and only if there exists a CT preference $\succsim^{CT}$ that generates $\mathcal{C}(\cdot)$. Furthermore, each such $\mathcal{C}(\cdot)$ 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, Kurusçu and Smith [14] when $\gamma > 0$. Theorem 5.1 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 4.2 holds for CT preferences as well. The definition of 'normative preference $\succsim^*$ is elicited from $\mathcal{C}(\cdot, \succsim^{CT})$' is analogous to the counterpart in the previous section.

Theorem 5.2. *If (U, V) represents a CT preference $\succsim^{CT}$, and $\succsim^*$ is the normative preference elicited from $\mathcal{C}(\cdot, \succsim^{CT})$, then U is a representation of $\succsim^*$.*

6. Evidence

The foundations of the FT and CT model reveal that CT agents must differ from FT agents when it comes to satisfying Axioms 8 and 9. That is, the two types of agents differ at least in their demand for delayed commitment and their time-of-reversal functions τ. We use this information to provide evidence of temptation by future consumption.

Time of Preference Reversals: A large amount of experimental evidence in psychology supports the claim that agents discount future rewards by using a hyperbolic discount function.[16] In what follows, we restrict attention to Mazur's [17] version of the hyperbolic discounting functional form, which has fit experimental evidence particularly well. According to his formulation, a reward that is delayed by d periods is discounted by

$$\frac{1}{1+kd},$$

where k parameterizes the subject's sensitivity to delay.

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 the hyperbolic discount function implies that in order to induce a preference reversal, both rewards must be delayed by $\widehat{\tau}(s^{+0}, l^{+d})$ periods, where

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

and where k is a parameter that captures the subject's 'sensitivity' to delay. Thus, $\widehat{\tau}(s^{+0}, l^{+d})$ is the time-of-reversal function implied by hyperbolic discounting. Observe that $\widehat{\tau}(s^{+0}, l^{+d})$ is increasing in s and d, and decreasing in l. The empirical evidence on hyperbolic discounting serves as indirect empirical evidence in favor of these properties of $\widehat{\tau}$. Some direct evidence may be found in Ainslie and Haendel [3].

How does the τ function of CT and FT agents compare to the $\widehat{\tau}$ function above? Suppose a preference reversal is observed for a pair of rewards (μ, η).[17] Then CT agents exhibit,

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

---

[16] See Ainslie [2] for an overview of the literature.

[17] The necessary and sufficient condition for a preference reversal to be observed for (μ, η) 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, $U(\mu) \geq U(\eta)$.

whereas FT agents exhibit,[18]

$$\tau^{FT}(\mu, \eta) = \frac{\ln \frac{U(\mu)-U(\eta)}{V(\eta)-V(\mu)}}{\ln \frac{\gamma}{\delta}},$$

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 on the agent's discount factors.

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

$$\tau^{FT}(s^{+0}, l^{+d}) = \frac{\ln \frac{\delta^d u(l)-u(s)}{v(s)-\gamma^d v(l)}}{\ln \frac{\gamma}{\delta}}.$$

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

Thus, τ^{FT} shares the same features as $\widehat{\tau}$, while τ^{CT} clearly does not. This supports the idea that agents are tempted by future consumption.

It is worth noting that the evidence reveals that temptation by future consumption is not restricted to a short time horizon. For instance, Ainslie and Haendel [3] and

---

[18] Strictly speaking, since time is discrete, τ^{FT} should be rounded off to the next integer. The same is true for $\widehat{\tau}$.

Kirby and Hernnstein [13, Experiment 3] find that the time of a reversal may be in years – the latter uncovers a preference reversal at almost 3.5 years delay. That is, it may take a delay of 3.5 years before the temptation of a reward is resistible, suggesting that rewards that are that far in the future may tempt.

Demand for Delayed Commitment: CT agents do not exhibit a demand for delayed commitment, while FT agents may. That is, for FT agents there may exist c, x, y and t such that

$$\{(c,x)\}^{+t} > \{(c,x),(c,y)\}^{+t}.$$

The results of Benartzi and Thaler [5] may be interpreted as evidence supporting a demand for delayed commitment. The authors introduce a saving-enhancement plan, called the 'Save More Tomorrow$^{\mathrm{TM}}$ (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. Any withdrawals one makes from a defined-contributions plan before the age of $59\frac{1}{2}$ is at a cost. Thus, such plans provide a means of committing funds for retirement.

A desire to opt for this plan suggests a preference for delayed commitment since subjects would rather not leave the allocation decision for the time *prior* to receiving the increased salary. For concreteness, imagine that one expects to learn of a salary increase in, say, December. The actual increase would take place in the following month. In December, the allocation decision does not affect current consumption, but rather next month's budget set (menu), and a preference for avoiding making a decision in December is consistent with a temptation to under-allocate next month's funds for retirement.[19] The authors implemented the SMT plan in several firms, and find 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).[20]

---

[19] If this temptation did not exist at such a time, the SMT plan would be of no use to the agent: he would make the optimal allocation decision at such a time anyway.

[20] One feature of the SMT plan that is problematic for our purposes is that the allocation decisions

However, while Benartzi and Thaler's results are suggestive of a demand for delayed commitment, they are not demonstrative. One can adjust the length of a period in a way that the results could be explained by the CT model as well. For instance, if a period is defined to be two months long, then in December, January's income is part of the current period. In such a case, the allocation decision in December involves current consumption, and the CT model can account for why agents prefer to make an allocation decision prior to December. Thus, whether Benartzi and Thaler's results are taken as evidence of a demand for delayed commitment depends on the stand one takes regarding the length of a period. Note that the earlier discussion on the time of reversals is not sensitive to the definition of a period – the CT model cannot account for the evidence on the time of reversals.

We should add that, while the evidence on the time of reversals supports temptation by future consumption, it does not necessarily reject the CT model. Recall that delaying rewards by about 3.5 years would induce a preference reversal for all the subjects in Kirby and Hernnstein [13, Experiment 3]. Thus, if one defines a period to be over 3.5 years long, the subjects begin to look like CT agents: a one period delay induces a preference reversal for all subjects. The upshot of this observation is that the appropriate model to use in applications depends on the specific application one has in mind. In applications that study choices that are made frequently, it may be natural to assume that a period is 'short', and so the FT model may be more appropriate than the CT model. On the other hand, in applications that involve choices made less frequently, the natural length of a period would be longer, and the longer it is, the more appropriate it may be to use the CT model.

---

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.

7. Concluding Remarks

We summarize some basic differences between GP's and our approach to providing foundations for models of temptation.

- The starting point of GP's approach is the idea that temptation creates a preference for commitment, whereas our starting point is the idea that delayed temptations are easier to resist than immediate temptations. Thus, GP look for a reflection of the conflict between an agent's normative and temptation preference in his ranking of menus, whereas we look for a reflection in his ranking of delayed rewards.

- GP's primitive is not just any ranking of menus, but rather one that captures the agent's choices between menus in a special period 0 where no temptation is experienced. This is a restrictive feature of GP's approach; the observability of the primitive ranking is not ensured, for instance, when modeling FT agents who are tempted by menus that contain tempting items. In contrast, our primitive – a ranking of delayed rewards – is one that captures choices that may well be subject to temptation. Thus, its observability is not in question.

- Traditional revealed preference theory identifies choice and welfare. This identification becomes invalid when studying decision-makers who have self-control problems since temptation drives a wedge between choice and welfare. But GP hypothesize that the identification may nevertheless be valid at the level of choice between menus. It is on the basis of this hypothesis that GP are able to provide foundations for a model of temptation: an agent's choices between menus can be used to infer which alternatives he regards as best for his welfare and which he finds tempting. We argued in this paper that temptation can exist even at the level of choices between menus, and thus, a reliance on the traditional identification of choice and welfare may not be justified. In contrast with GP, our approach does not rely on the identification at any level.

A. Appendix: Topology on $\mathcal{P}$

Since is Δ is compact and metrizable, $\Delta \times \Delta$ is compact and metrizable under the product topology. Let d be a metric that generates the topology on $\Delta \times \Delta$. 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

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

for all $A, B \in \mathcal{P}$. An ε-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 Δ by identifying any such binary relation B on Δ with $\Gamma(B)$, the graph of B:

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

If B is a weak order (complete and transitive binary relation) then $\Gamma(B)$ is nonempty. If B is also continuous then $\Gamma(B)$ is closed, and hence compact.[21] Thus, the set of continuous weak orders on Δ is a subset of $\mathcal{P}$.

By [4, Thm 3.71(3)], compactness of $\Delta \times \Delta$ implies that $\mathcal{P}$ 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:

$Ls\Gamma(B_n) = \{a \in \Delta \times \Delta :$ for every neighborhood V of a,

$$V \cap \Gamma(B_n) \neq \phi \text{ for infinitely many } n\}$$

$Li\Gamma(B_n) = \{a \in \Delta \times \Delta :$ for every neighborhood V of a,

---

[21] The space of lotteries Δ is connected, and moreover, it is separable since it is compact metric. To show that $\Gamma(B)$ is closed if B is a continuous weak order, use [7, Lemma 5.1 and Exercise 3.16].

$$V \cap \Gamma(B_n) \neq \phi \text{ for all but a finite number of } n\}.$$

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

B. Appendix: Normative Preference

This appendix collects some results from [19]. Take as given a set of preference relations $\{\gtrapprox_t\}_{t=0}^{\infty}$ on Δ, a set of lotteries defined over any compact metric space and endowed with the weak convergence topology. For each μ, η, the preference $\gtrapprox_t$ captures how the agent ranks the rewards μ, η when they are to be received t periods later. Normative preference $\succsim^*$ over Δ is defined as in Definition 4.1, that is,

$$\succsim^* = \lim_{t \to \infty} \gtrapprox_t .$$

Consider the following axioms on $\{\gtrapprox_t\}$.

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

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

Axiom A3 (Reversal*) If $\mu \lessapprox_0 \eta$ and $\mu \gtrapprox_T \eta$ with at least one ranking strict, then $\mu >_t \eta$ for all $t > T$.

Axiom A4 (Independence*) $\mu >_t \eta \Longrightarrow \mu\alpha\nu >_t \eta\alpha\nu$, for all t.

Define the function $\tau : \Delta \times \Delta \to \mathbb{R}$ which captures the time at which a reversal takes place for each (μ, η) in the following way: First take any $(\mu, \eta) \in \Delta \times \Delta$ such that $\mu \gtrapprox_0 \eta$. If

$$[\mu \approx_0 \eta \Longrightarrow \mu \approx_t \eta \text{ for all } t],$$
$$\text{and } [\mu >_0 \eta \Longrightarrow \mu >_t \eta \text{ for all } t],$$

then define $\tau(\mu, \eta) = 0$, and if there exists T such that $\mu <_T \eta$, then define

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

Finally, to cover all the remaining cases, let $\tau(\mu, \eta) = \tau(\eta, \mu)$ for all μ, η. Note that by Reversal, if $\mu >_0 \eta$ and $\mu \approx_{t'} \eta$ for some t' then there exists $T > t'$ such that $\mu <_T \eta$ and also that Reversal rules out post-reversal indifference.

The following results were proved in [19] and will be used later.

Lemma B.1. *Suppose* $\{\gtrapprox_t\}_{t=0}^{\infty}$ *satisfies A1 and A3. Then, for any* μ, η, *if* $\tau(\mu,\eta) = 0$ *then,*

$$\mu \gtrapprox_0 \eta \Longleftrightarrow \mu \gtrapprox_t \eta, \quad \textit{for all } t \geq 0,$$

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

(a) $\mu >_0 \eta \Longrightarrow \mu >_t \eta$ *for all* $t < \tau(\mu,\eta) - 1$; $\mu \gtrapprox_t \eta$ *for* $t = \tau(\mu,\eta) - 1$; $\mu <_t \eta$ *for all* $t \geq \tau(\mu,\eta)$.

(b) $\mu \approx_0 \eta \Longrightarrow \mu \approx_t \eta$ *for all* $t < \tau(\mu,\eta)$; $\mu >_t \eta$ *for all* $t \geq \tau(\mu,\eta)$ *or* $\mu <_t \eta$ *for all* $t \geq \tau(\mu,\eta)$.

Lemma B.2. *If* $\{\gtrapprox_t\}$ *satisfies A1-3, then* $\succsim^*$ *is well-defined, complete, transitive and continuous.*

Lemma B.3. *If* $\{\gtrapprox_t\}$ *satisfies A1-4, then* $\succsim^*$ *also satisfies independence.*

The last Lemma provides two characterizations of normative preference $\succsim^*$. Part (c) of the Lemma is implied by (a). Let Ω be the subset of $\Delta \times \Delta$ on which τ is upper semicontinuous, that is,

$$\Omega = \{(\mu,\eta) \in \Delta \times \Delta : (\mu_n,\eta_n) \to (\mu,\eta) \Longrightarrow \limsup_{n\to\infty} \tau(\mu_n,\eta_n) \leq \tau(\mu,\eta)\}.$$

Lemma B.4. *(a)* $\mu \succ^* \eta \Longleftrightarrow [\mu >_{\tau(\mu,\eta)} \eta$ *and* $(\mu,\eta) \in \Omega]$

(b) $\mu \succsim^* \eta \Longleftrightarrow \exists$ *a sequence* $\{(\mu_n,\eta_n)\}$ *that converges to* (μ,η) *and* $\mu_n \gtrapprox_{\tau(\mu_n,\eta_n)} \eta_n$ *for all* n.

(c) $\mu \gtrapprox_{\tau(\mu,\eta)} \eta \Longrightarrow \mu \succsim^* \eta$.

C. Appendix: Nondegeneracy

Let $\succsim$ be any binary relation that is represented by $W : Z \to \mathbb{R}$ such that

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

where $U, V : \Delta \to \mathbb{R}$ are linear and continuous. Say that (U,V) is a representation of $\succsim$, and that $\succsim$ is nondegenerate if there exists $x, y \in Z$ such that

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

The following Lemma will be useful throughout.

Lemma C.1. *For all x, y,*

(a) $x \succ x \cup y \Longleftrightarrow \max_y V > \max_x V$ *and* $W(x) > W(y)$.

(b) $x \cup y \succ y \Longleftrightarrow \max_x U + V > \max_y U + V$ *and* $W(x) > W(y)$.

(c) $x \succ x \cup y \succ y \Longleftrightarrow \max_x U + V > \max_y U + V$ *and* $\max_y V > \max_x V$.

Proof. See [18]. ■

Lemma C.2. *Let (U, V) be a representation of $\succsim$, where U and V are as in (2.1). Then U and V are nonconstant and U is not an affine transformation of V if and only if $\succsim$ is nondegenerate.*

Proof. $\Longleftarrow$: Suppose there exists x, y such that $x \succ x \cup y \succ y$. By Lemma C.1(a), it is clear that V is not constant. Neither can U be constant, for if $U(\mu) = k$ for all μ, then a simple calculation shows that $W(z) = k$ for all z. Moreover, by Lemma C.1(c), U cannot be a positive affine transformation of V. Suppose by way of contradiction that U is a negative affine transformation of V. Then, $U = -\alpha V + \beta$ for some $\alpha > 0$, and so, by the functional forms of U and V,

$$W(z) = -\alpha \overline{V}(z) + \xi(c), \tag{C.1}$$

where $\xi(c)$ is some function of c. Observe that by definition of $\overline{V}$, for any $z, w \in Z$,

$$\overline{V}(z) \geq \overline{V}(w) \Longrightarrow \overline{V}(z) = \overline{V}(z \cup w) \geq \overline{V}(w). \tag{C.2}$$

But by hypothesis, there is x, y such that $W(x) > W(x \cup y) > W(y)$. It follows by (C.1) that

$$\overline{V}(y) > \overline{V}(x) \text{ and } \overline{V}(y) > \overline{V}(x \cup y) > \overline{V}(x),$$

contradicting (C.2).

$\Longrightarrow$: Suppose that U and V are nonconstant and that U is not an affine transformation of V. Thus, $U + V$ is nonconstant, and V is not a positive affine transformation of $U + V$. It follows that there exist μ, η such that either $[V(\mu) > V(\eta)$ and $U(\mu) + V(\mu) \leq U(\eta) + V(\eta)]$, or $[V(\mu) \geq V(\eta)$ and $U(\mu) + V(\mu) < U(\eta) + V(\eta)]$. If a case where all inequalities are strict obtains, then Lemma C.1(c) implies

$$\{\mu\} \succ \{\mu, \eta\} \succ \{\eta\},$$

thus yielding the result. If, however, we have $V(\mu) > V(\eta)$ and $U(\mu) + V(\mu) = U(\eta) + V(\eta)$, then nonconstancy of $U + V$ and linearity of V and $U + V$ imply the existence of μ' and η' close to μ and η, respectively, such that $U(\mu') > U(\eta')$ and $U(\mu') + V(\mu') > U(\eta') + V(\eta')$. Applying Lemma C.1(c) yields the desired result. Similarly, the result follows in the case where $V(\mu) = V(\eta)$ and

$U(\mu) + V(\mu) < U(\eta) + V(\eta)$ by using the nonconstancy of V, the linearity of V and $U + V$ and Lemma C.1(c). ■

The following corollary of the above Lemma will be used later.

Lemma C.3. *If $\succsim$ is nondegenerate and has a representation (U, V) where U and V are as in (2.1), then there exist $\rho, \nu \in \Delta$ such that*

$$\{\rho\} \succ \{\rho, \nu\} \succ \{\nu\}.$$

D. Appendix: Proof of Theorem 3.1 (Necessity)

By hypothesis, there exists an FT preference $\succsim$ with a representation (U, V) that generates the choice correspondence $\mathcal{C}$. Define the function $\phi : \Delta(C \times Z) \rightarrow \mathbb{R}$ by $\phi(\mu) = U(\mu) + V(\mu)$ for all $\mu \in \Delta$. Thus,

$$\mathcal{C}(x) = \arg\max_{\mu \in x} \phi(\mu).$$

Clearly, $\mathcal{C}(x) \subset x$. Since ϕ 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 ϕ 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 $\gtrapprox$ be the binary relation that is represented by ϕ. For all $t \geq 1$, define $\gtrapprox_t$ over Z by

$$x \gtrapprox_t y \Longleftrightarrow x^{+t} \gtrapprox y^{+t}.$$

Lemma D.1. *$\mathcal{C}$ satisfies Set-Betweenness.*

Proof. Note that

$$\begin{aligned}
\phi(x^{+t}) &= \delta^t W(x) + \gamma^t \overline{V}(y) + \text{constant} \\
&= \delta^t (\max_{\mu \in x} U(\mu) + V(\mu) - \max_{\eta \in x} V(\eta)) + \gamma^t \overline{V}(x) + \text{constant} \\
&= \delta^t (\max_{\mu \in x} U(\mu) + V(\mu) - (1 - \frac{\gamma^t}{\delta^t}) \max_{\eta \in x} V(\eta)) + \text{constant}.
\end{aligned}$$

Hence, $\gtrapprox_t$ 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 ϕ' represents a Self-Control preference [10]. Therefore, $\mathcal{C}$ satisfies the first part of Set-Betweenness.

To establish the nondegeneracy part of Set-Betweenness, observe that by nondegeneracy of the FT preference $\succsim$ and by Lemma C.3, there exists μ, η such that $\{\mu\} \succ \{\mu, \eta\} \succ \{\eta\}$.We show $\{\mu\}^{+1} > \{\mu, \eta\}^{+1} > \{\eta\}^{+1}$. Recall that $\gtrapprox$ is represented by ϕ. By Lemma C.1, $\{\mu\} \succ \{\mu, \eta\} \succ \{\eta\}$ implies $U(\mu) > U(\eta)$, $\phi(\mu) > \phi(\eta)$ and $V(\mu) < V(\eta)$. The latter implies $\overline{V}(\{\eta\}) = \overline{V}(\{\mu, \eta\})$. Now compute that $V(\mu) < V(\eta) \iff (\gamma - \delta)V(\mu) > (\gamma - \delta)V(\eta) \iff \delta U(\mu) + \gamma V(\mu) > \delta(U(\mu) + V(\mu) - V(\eta)) + \gamma V(\eta) \iff \delta W(\{\mu\}) + \gamma \overline{V}(\{\mu\}) > \delta W(\{\mu, \eta\}) + \gamma \overline{V}(\{\mu, \eta\}) \iff \phi(\{\mu\}^{+1}) > \phi(\{\mu, \eta\}^{+1})$. Furthermore, $\{\mu, \eta\} \succ \{\eta\} \iff \delta W(\{\mu, \eta\}) > \delta W(\{\eta\}) \iff \delta W(\{\mu, \eta\}) + \gamma \overline{V}(\{\mu, \eta\}) > \delta W(\{\eta\}) + \gamma \overline{V}(\{\eta\}) \iff \phi(\{\mu, \eta\}^{+1}) > \phi(\{\eta\}^{+1})$. Put together, $\phi(\{\mu\}^{+1}) > \phi(\{\mu, \eta\}^{+1}) > \phi(\{\eta\}^{+1})$, as desired. ■

Lemma D.2. *$\mathcal{C}$ satisfies Sophistication.*

Proof. In the proof of the previous lemma, we showed that $\gtrapprox_t$ 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 Lemma C.1(b), for any μ, η, t such that $\{\mu\} >_t \{\eta\}$,

$$\{\mu, \eta\} >_t \{\eta\} \iff U(\mu) + V(\mu) > U(\eta) + V(\eta).$$

But $U + V$ represents $\gtrapprox$ and thus $\{\mu, \eta\} >_t \{\eta\}$ is equivalent to $\mu > \eta$. ■

Lemma D.3. *$\mathcal{C}$ satisfies Menus Can Tempt.*

Proof. Note that $\gtrapprox_t$ is represented by $W(\cdot) + \frac{\gamma^t}{\delta^t}\overline{V}(\cdot)$. By [18, Theorem 3.1], the FT preference $\succsim$ satisfies the following condition:

$$x \succ x \cup y \iff \{(c, x)\} \succ \{(c, x), (c, y)\},$$

and so $W(x) > W(x \cup y) \iff W(\{(c, x)\}) > W(\{(c, x), (c, y)\})$.

Suppose that $x >_t x \cup y$, that is, $W(x) + \frac{\gamma^t}{\delta^t}\overline{V}(x) > W(x \cup y) + \frac{\gamma^t}{\delta^t}\overline{V}(x \cup y)$ for some t. Then, $\overline{V}(x \cup y) \geq \overline{V}(x)$ implies

$$W(x) > W(x \cup y) \ , \tag{D.1}$$

which, by Lemma C.1(a), implies

$$\overline{V}(y) > \overline{V}(x). \tag{D.2}$$

By the earlier observation, (D.1) implies $W(\{(c,x)\}) > W(\{(c,x),(c,y)\}$. Furthermore, (D.2) implies $\overline{V}(\{(c,x),(c,y)\}) > \overline{V}(\{(c,x)\})$ since $\overline{V}(\{(c,x),(c,y)\}) = v(c) + \max_{\{x,y\}} \overline{V} = V(c,y) > V(c,x) = \overline{V}(\{(c,x)\})$. Since $\frac{\gamma}{\delta} < 1$, it follows that for large enough t',

$$W(\{(c,x)\}) + \frac{\gamma^{t'}}{\delta^{t'}}\overline{V}(\{(c,x)\}) > W(\{(c,x),(c,y)\}) + \frac{\gamma^{t'}}{\delta^{t'}}\overline{V}(\{(c,x),(c,y)\}),$$

that is, $\{(c,x)\} >_{t'} \{(c,x),(c,y)\}$ for some t', as desired. The converse can be established in an analogous manner. Note that $\overline{V}(\{(c,x),(c,y)\}) > \overline{V}(\{(c,x)\})$ implies $\overline{V}(y) > \overline{V}(x)$, which in turn implies $\overline{V}(x \cup y) > \overline{V}(x)$. ■

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

Lemma D.4. *There exists μ, η and T such that $\mu \lessapprox \eta$ and $\mu^{+T} > \eta^{+T}$.*

Proof. By [18, Theorem 5.3], $\gamma < \delta$ implies that $\succsim$ exhibits preference reversals, in the sense of [18, Section 5.1]. All we need to show is that there exists μ, η such that

$$\{\mu\} \succ \{\mu, \eta\} \sim \{\eta\}.$$

The proof of this closely follows the sufficiency part of the proof of Lemma C.2. Since $\succsim$ is nondegenerate, Lemma C.2 implies that U and V are nonconstant, and furthermore, U is not an affine transformation of V. Thus, $U+V$ is nonconstant and U is not a positive affine transformation of $U+V$. It follows that there exist μ, η such that either $[U(\mu) > U(\eta)$ and $U(\mu)+V(\mu) \leq U(\eta)+V(\eta)]$, or $[U(\mu) \geq U(\eta)$ and $U(\mu)+V(\mu) < U(\eta)+V(\eta)]$. Use Lemma E.3(a) and (b) and follow the proof of Lemma C.2 to establish the result. ■

The proof of Lemma D.5 establishes that $\mathcal{C}$ satisfies the first part of Reversal.Recall the time-of-reversal function τ defined in Appendix B.

Lemma D.5. *Suppose* $U(\mu) \geq U(\eta)$. *Then,*

$$\begin{aligned} \frac{U(\mu)-U(\eta)}{V(\eta)-V(\mu)} &\in (0,1] \Longrightarrow \tau(\mu,\eta) = \min\{k \in \mathbb{N} : k > \frac{\ln \frac{U(\mu)-U(\eta)}{V(\eta)-V(\mu)}}{\ln \frac{\gamma}{\delta}}\} > 0 \\ \frac{U(\mu)-U(\eta)}{V(\eta)-V(\mu)} &\notin (0,1] \Longrightarrow \tau(\mu,\eta) = 0. \end{aligned}$$

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

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

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

$$\begin{aligned} \forall t &< \tau(\mu,\eta),\ U(\mu) + \frac{\gamma^t}{\delta^t} V(\mu) \leq U(\eta) + \frac{\gamma^t}{\delta^t} V(\eta) \\ \forall t &\geq \tau(\mu,\eta),\ U(\mu) + \frac{\gamma^t}{\delta^t} V(\mu) > U(\eta) + \frac{\gamma^t}{\delta^t} V(\eta). \end{aligned} \tag{D.3}$$

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

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

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

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

If $V(\eta) = V(\mu)$ or if $V(\eta) \neq V(\mu)$ and $\frac{U(\mu)-U(\eta)}{V(\eta)-V(\mu)} = 0$, then it is straightforward to establish that

$$\mu \gtrapprox \eta \Longleftrightarrow \mu^{+t} \gtrapprox \eta^{+t},$$

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 $\frac{\gamma}{\delta} < 1$ 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 D.6. *If* $U(\mu) > U(\eta)$ *then* $\mu >_t \eta$ *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 D.5. The second assertion follows from the first. ■

Define $\Omega = \{(\mu, \eta)$: for any sequence $\{(\mu_n, \eta_n)\}$ that converges to (μ, η),

$$\limsup_{n\to\infty} \tau(\mu_n, \eta_n) \leq \tau(\mu, \eta)\}.$$

Lemma D.7. *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 (μ, η). Consider the following possibilities:

(a) $V(\mu) \neq V(\eta)$.

Consider three possibilities.

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

We show the result for the case where $\frac{U(\mu)-U(\eta)}{V(\eta)-V(\mu)} > 1$; the other case follows by an analogous argument. Note that $U(\mu) > U(\eta)$ and $\frac{U(\mu)-U(\eta)}{V(\eta)-V(\mu)} > 1$ implies $V(\mu) < V(\eta)$. By continuity of V there exists M such that $V(\mu_n) > V(\eta_n)$ for all $n \geq M$. Without loss of generality, $M = 1$. Note that $\frac{U(\mu_n)-U(\eta_n)}{V(\eta_n)-V(\mu_n)} \to \frac{U(\mu)-U(\eta)}{V(\eta)-V(\mu)} > 1$, and so there exists N such that $\frac{U(\mu_n)-U(\eta_n)}{V(\eta_n)-V(\mu_n)} > 1$ for all $n \geq N$. Lemma D.5 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 $(\mu, \eta) \in \Omega$.

(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 D.5, $\tau(\mu_n, \eta_n) = \min\{k \in N : k > \frac{\ln \frac{U(\mu_n)-U(\eta_n)}{V(\eta_n)-V(\mu_n)}}{\ln \frac{\gamma}{\delta}}\}$ for each n. Since $\frac{U(\mu_n)-U(\eta_n)}{V(\eta_n)-V(\mu_n)} \to \frac{U(\mu)-U(\eta)}{V(\eta)-V(\mu)}$, we have $\lim_{n\to\infty} \tau(\mu_n, \eta_n) = \tau(\mu, \eta)$, which establishes the result.

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

It must be that $V(\mu) < V(\eta)$. Thus 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 D.5, $\tau(\mu_n, \eta_n) = 0 \leq \tau(\mu, \eta)$ 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 (μ_n, η_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 (μ, η) and for all m,

$\frac{U(\mu_n)-U(\eta_n)}{V(\eta_n)-V(\mu_n)} \in (0,1)$. Note that $\tau(\mu_n, \eta_n) = 0$ for all these discarded (μ_n, η_n), and that by the argument in (ii), $\limsup_{n\to\infty} \tau(\mu_{n(m)}, \eta_{n(m)}) \leq \tau(\mu, \eta)$. Thus $\limsup_{n\to\infty} \tau(\mu_n, \eta_n) \leq \tau(\mu, \eta)$, as desired.

(b) $V(\mu) = V(\eta)$.

Continuity of U and V implies $\lim_{n\to\infty}(U(\mu_n) - U(\eta_n)) > 0$ and $\lim_{n\to\infty}(V(\eta_n) - V(\mu_n)) = 0$. 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 $\frac{U(\mu_n)-U(\eta_n)}{V(\eta_n)-V(\mu_n)} \notin (0,1]$ and $V(\eta_n) \neq V(\mu_n)$, or $V(\eta_n) = V(\mu_n)$. It follows from Lemma D.5 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 $\mathcal{C}$ satisfies Commitment is Normative. Note that $\tau(\mu, \eta) = 0$ and $(\mu, \eta) \in \Omega$ imply that τ is continuous at (μ, η).

Lemma D.8. *(a)*$x >_t x \cup y \Longrightarrow \tau(x^{+t}, (x \cup y)^{+t}) = 0$ *and* $(x^{+t}, (x \cup y)^{+t}) \in \Omega$.

(b) $x >_t x \cup y >_t y \Longrightarrow \tau((x \cup y)^{+t}, y^{+t}) = 0$ *and* $((x \cup y)^{+t}, y^{+t}) \in \Omega$.

Proof. (a) Let $x >_t x \cup y$, that is,

$$W(x) + \frac{\gamma^t}{\delta^t}\overline{V}(x) > W(x \cup y) + \frac{\gamma^t}{\delta^t}\overline{V}(x \cup y).$$

From $x \subset x \cup y$ it follows that $\overline{V}(x \cup y) \geq \overline{V}(x)$. 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 D.6 and D.7, $\tau(x^{+t}, (x \cup y)^{+t}) = 0$ and $(x^{+t}, (x \cup y)^{+t}) \in \Omega$.

(b) Note that by Lemma E.3(a) below, $W(x) > W(x \cup y)$ implies

$$\overline{V}(x \cup y) = \overline{V}(y) > \overline{V}(x).$$

Hence, $x \cup y >_t y$ implies

$$W(x \cup y) + \frac{\gamma^t}{\delta^t}\overline{V}(x \cup y) > W(y) + \frac{\gamma^t}{\delta^t}\overline{V}(y).$$

It follows that $W(x \cup y) > W(y)$, which implies $U((x \cup y)^{+t}) > U(y^{+t})$. By Lemmas D.6 and D.7, $\tau((x \cup y)^{+t}, y^{+t}) = 0$ and $((x \cup y)^{+t}, y^{+t}) \in \Omega$. ■

E. Appendix: Proof of Theorem 3.1 (Sufficiency)

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

E.1. Properties of $\gtrapprox$

Define the choice correspondence $\mathcal{C}^*(\cdot, \gtrapprox)$ by

$$\mathcal{C}^*(x, \gtrapprox) \equiv \{\mu \in x : \mu \gtrapprox \eta \text{ for all } \eta \in x\}.$$

Say that $\gtrapprox$ rationalizes $\mathcal{C}(\cdot)$ if $\mathcal{C}(x) = \mathcal{C}^*(x, \gtrapprox)$ for all x.

Lemma E.1. *$\gtrapprox$ is the unique preference relation that rationalizes $\mathcal{C}(\cdot)$. Furthermore $\gtrapprox$ satisfies the vNM axioms.*

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

We want to show that $\{\eta : \eta \gtrapprox \mu\}$ is closed for all μ; that $\{\eta : \mu \gtrapprox \eta\}$ is closed for all μ follows by an analogous argument. Take $\{\eta_n\}$ such that $\eta_n \gtrapprox \mu$ for all n and $\eta_n \to \eta$. Consider the sequence of menus $\{\{\eta_n, \mu\}_n\}$. Since Z is endowed with the Hausdorff metric, $\eta_n \to \eta$ implies $\{\eta_n, \mu\} \to \{\eta, \mu\}$. By definition of $\gtrapprox$, $\eta_n \in \mathcal{C}(\{\eta_n, \mu\})$ for all n. By Continuity (upper hemicontinuity of $\mathcal{C}$) and by [4, Thm 16.20], $\eta \in \mathcal{C}(\{\eta, \mu\})$, that is, $\eta \gtrapprox \mu$, as desired. This establishes that $\{\eta : \eta \gtrapprox \mu\}$ is closed.

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 $\gtrapprox'$ (with domain Δ) by

$$\mu \gtrapprox' \eta \Longleftrightarrow \exists x \text{ such that } \mu, \eta \in x \text{ and } \mu \in \mathcal{C}(x).$$

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

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

Step 3: $\mathcal{C}^*(\cdot, \gtrapprox)$ is nonempty.

The domain Z consists of compact menus, $\gtrapprox$ is continuous (Step 1) and thus, $\mathcal{C}^*(\cdot, \gtrapprox) \neq \phi$ by [4, Thm 2.41].

Step 4: $\gtrapprox$ is the unique weak order that rationalizes $\mathcal{C}(\cdot)$.

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

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

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

For $t > 0$, define $\gtrapprox_t$ over Z by

$$x \gtrapprox_t y \Longleftrightarrow x^{+t} \gtrapprox y^{+t}.$$

We establish that each $\gtrapprox_t$ satisfies the following axioms.[22]

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

B2 (Continuity**) The sets $\{y : x \gtrapprox_t y\}$ and $\{y : y \gtrapprox_t x\}$ are closed.

B3 (Independence**) $x >_t y \Longrightarrow \alpha x + (1-\alpha)z >_t \alpha y + (1-\alpha)z$.

B4 (Set-Betweenness**) $x \gtrapprox_t y \Longrightarrow x \gtrapprox_t x \cup y \gtrapprox_t y$.

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 δ_z denote the lottery degenerate at menu z. Define $\varphi : \Delta_s(Z) \longrightarrow Z$ 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 $\gtrapprox_t$ satisfies Order*, Continuity*, Independence* and Set-Betweenness*, that is, $\gtrapprox_t$ is a Self-Control preference [10].

---

[22] These axioms are discussed in [18].

Lemma E.2. *$\gtrapprox_t$ satisfies Order**, Continuity**, Independence** and Set-Betweenness**.*

Proof. It is clear that by Lemma E.1, $\gtrapprox_t$ satisfies Order** and Continuity**. By Set-Betweenness, $\gtrapprox_t$ satisfies Set-Betweenness**. To see that $\gtrapprox_t$ satisfies Independence**, observe that by Independence, for any x, y, z,

$$x^{+t} > y^{+t} \Longrightarrow \alpha x^{+t} + (1-\alpha)z^{+t} > \alpha y^{+t} + (1-\alpha)z^{+t},$$

and that by Indifference to Timing,

$$\begin{aligned} \alpha x^{+t} + (1-\alpha)z^{+t} &\approx (\alpha x + (1-\alpha)z)^{+t} \\ \alpha y^{+t} + (1-\alpha)z^{+t} &\approx (\alpha y + (1-\alpha)z)^{+t}. \end{aligned}$$

Independence** then follows by Order** and definition of $\gtrapprox_t$. ■

By Lemma E.2 and [10, Theorem 1], each $\gtrapprox_t$ is represented by $W_t : Z \to \mathbb{R}$ such that

$$W_t(x) = \max_{\mu \in x}\{U_t(\mu) + V_t(\mu)\} - \max_{\eta \in x} V_t(\eta), \tag{E.1}$$

where $U_t, V_t : \Delta \to \mathbb{R}$ are linear and continuous.

Lemma E.3. *For all x, y,*

(a) $x >_t x \cup y \Longleftrightarrow \max_y V_t > \max_x V_t$ and $W_t(x) > W_t(y)$.

(b) $x \cup y >_t y \Longleftrightarrow \max_x U_t + V_t > \max_y U_t + V_t$ and $W_t(x) > W_t(y)$.

(c) $x >_t x \cup y >_t y \Longleftrightarrow \max_x U_t + V_t > \max_y U_t + V_t$ and $\max_y V_t > \max_x V_t$.

Proof. This is Lemma C.1 adapted to the current setting. ■

Lemma E.4. *$\mu \gtrapprox \eta \Longleftrightarrow U_t(\mu) + V_t(\mu) \geq U_t(\eta) + V_t(\eta)$.*

Proof. Fix any $t > 0$. By Commitment is Normative and the second part of Set-Betweenness, $\gtrapprox_t$ is nondegenerate and thus by Lemma C.3, there exists ρ, ν such that

$$\{\rho\} >_t \{\rho, \nu\} >_t \{\nu\}.$$

By Order**, $\{\rho\} >_t \{\nu\}$ and by Lemma E.3(b) and Sophistication, $\rho > \nu$ and $U(\rho) + V(\rho) > U(\nu) + V(\nu)$. These observations will be used to prove the result.

$\Longrightarrow$: Suppose that $\mu \gtrapprox \eta$. If $\{\eta\} >_t \{\mu\}$, then by Sophistication, $\{\eta, \mu\} \not>_t \{\mu\}$ and it follows by Lemma E.3(b) that $U(\mu) + V(\mu) \geq U(\eta) + V(\eta)$. If, on the other hand, $\{\mu\} \gtrapprox_t \{\eta\}$, then by Independence** and Independence,

$$\{\mu\alpha\rho\} >_t \{\eta\alpha\nu\} \text{ and } \mu\alpha\rho > \eta\alpha\nu,$$

for all $\alpha \in (0,1)$. By Sophistication and Lemma E.3(b), $U(\mu\alpha\rho) + V(\mu\alpha\rho) > U(\eta\alpha\nu) + V(\eta\alpha\nu)$ for all $\alpha \in (0,1)$. By continuity of $U + V$, it follows that $U(\mu) + V(\mu) \geq U(\eta) + V(\eta)$, as desired.

$\Longleftarrow$: Suppose that $U(\mu)+V(\mu) \geq U(\eta)+V(\eta)$. If $\{\eta\} >_t \{\mu\}$, then by Lemma E.3(b), $\{\eta, \mu\} \not>_t \{\mu\}$ and by Sophistication, $\mu \gtrapprox \eta$. If, on the other hand, $\{\mu\} \gtrapprox_t \{\eta\}$, then for all $\alpha \in (0,1)$,

$$\{\mu\alpha\rho\} >_t \{\eta\alpha\nu\} \text{ and } U(\mu\alpha\rho) + V(\mu\alpha\rho) > U(\eta\alpha\nu) + V(\eta\alpha\nu).$$

By Lemma E.3(b) and Sophistication, $\mu\alpha\rho > \eta\alpha\nu$ for all $\alpha \in (0,1)$. Thus, continuity of $\gtrapprox$ implies $\mu \gtrapprox \eta$. ■

Lemma E.5. *$\gtrapprox_t$ satisfies Separability**.*

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

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

$$\{\frac{1}{2}(c,x) + \frac{1}{2}(\overline{c},\overline{x})\} \approx_t \{\frac{1}{2}(c,\overline{x}) + \frac{1}{2}(\overline{c},x)\}.$$

It follows that $\gtrapprox_t$ satisfies GP's version of Separability, and so, by [11, Lemma 9(1)], U is additively separable, thus establishing Step 1.

Step 2: Show that

$$\begin{aligned} U_t(\mu) + V_t(\mu) &= \int_{C\times Z} \left(u_t(c) + \widehat{W}_t(y) + v_t(c) + \widehat{V}_t(y)\right) d\mu(c,y) \\ \text{and } V_t(\mu) &= \int_{C\times Z} \left(v_t(c) + \widehat{V}_t(y)\right) d\mu(c,y). \end{aligned}$$

Since V is linear and continuous, there exists a continuous function $\overline{v}_t : C \times Z \to \mathbb{R}$ such that for all $\mu \in \Delta$,

$$V_t(\mu) = \int \overline{v}_t(c,x) d\mu.$$

By Separability, $\frac{1}{2}(c,x) + \frac{1}{2}(\overline{c},\overline{x}) \approx \frac{1}{2}(c,\overline{x}) + \frac{1}{2}(\overline{c},x)$. Hence, by Lemma E.4,

$$\begin{aligned} & U_t(\frac{1}{2}(c,x) + \frac{1}{2}(\overline{c},\overline{x})) + V_t(\frac{1}{2}(c,x) + \frac{1}{2}(\overline{c},\overline{x})) \\ = \; & U_t(\frac{1}{2}(c,\overline{x}) + \frac{1}{2}(\overline{c},x)) + V_t(\frac{1}{2}(c,\overline{x}) + \frac{1}{2}(\overline{c},x)). \end{aligned}$$

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

$$U_t(\tfrac{1}{2}(c,x)+\tfrac{1}{2}(\overline{c},\overline{x})) + V_t(\tfrac{1}{2}(c,x)+\tfrac{1}{2}(\overline{c},\overline{x})) = U_t(\tfrac{1}{2}(c,\overline{x})+\tfrac{1}{2}(\overline{c},x)) + V_t(\tfrac{1}{2}(c,\overline{x})+\tfrac{1}{2}(\overline{c},x))$$

$$\Longrightarrow V_t(\tfrac{1}{2}(c,x)+\tfrac{1}{2}(\overline{c},\overline{x})) = V_t(\tfrac{1}{2}(c,\overline{x})+\tfrac{1}{2}(\overline{c},x))$$

$$\Longrightarrow V_t(c,x)+V_t(\overline{c},\overline{x})) = V_t(c,\overline{x})+V_t(\overline{c},x))$$

$$\Longrightarrow \overline{v}_t(c,x)+\overline{v}_t(\overline{c},\overline{x}) = \overline{v}_t(c,\overline{x})+\overline{v}_t(\overline{c},x)$$

$$\Longrightarrow \overline{v}_t(c,x) = \overline{v}_t(c,\overline{x})-\overline{v}_t(\overline{c},\overline{x})+\overline{v}_t(\overline{c},x).$$

Define $v_t(c) \equiv \overline{v}_t(c,\overline{x})-\overline{v}_t(\overline{c},\overline{x})$ and $\widehat{V}_t(x) \equiv \overline{v}_t(\overline{c},x)$. We can then write,

$$V_t(\mu) = \int_{C\times Z}\Big(v_t(c)+\widehat{V}_t(y)\Big)\,d\mu(c,y).$$

It also follows that, $U_t(\mu)+V_t(\mu) = \int_{C\times Z}\Big(u_t(c)+\widehat{W}_t(y)+v_t(c)+\widehat{V}_t(y)\Big)\,d\mu(c,y)$.

Step 3: Prove the result.

Take μ,η,π and ν such that $\mu^1=\pi^1, \mu^2=\pi^2, \eta^1=\nu^1$ and $\eta^2=\nu^2$. By Step 2,

$$\begin{aligned} U_t(\mu)+V_t(\mu) &= U_t(\pi)+V_t(\pi)\ , \qquad V_t(\mu)=V_t(\pi),\\ U_t(\eta)+V_t(\eta) &= U_t(\nu)+V_t(\nu), \qquad V_t(\eta)=V_t(\nu). \end{aligned}$$

Observe that

$$\max_{\{\mu,\eta\}}\{U_t+V_t\} = \max_{\{\pi,\nu\}}\{U_t+V_t\} \text{ and } \max_{\{\mu,\eta\}} V_t = \max_{\{\pi,\nu\}} V_t.$$

Hence by (E.1), $\{\mu,\eta\} \approx_t \{\pi,\nu\}$. ■

Lemma E.6. $\gtrapprox_t$ *satisfies Indifference to Timing***.

Proof. The proof follows steps that are similar to those in Lemma E.5: for the $\widehat{W}_t$ and $\widehat{V}_t$ in the proof of Lemma E.5, use Indifference to Timing to show that $\widehat{W}_t$ is linear, then show that $\widehat{V}_t$ is linear, and then use (E.1) to establish the result. ■

E.2. Normative Preference $\succsim^*$

For each $t>0$, denote the restriction of $\gtrapprox_t$ to $\Delta \subset Z$ by $\gtrapprox_{t|\Delta}$, and define $\gtrapprox_{0|\Delta}=\gtrapprox$. Then $\{\gtrapprox_{t|\Delta}\}_{t=0}^{\infty}$ is a set of preference relations defined over Δ. Since $\mathcal{C}(\cdot)$ satisfies WARP, Continuity and Reversal, $\{\gtrapprox_{t|\Delta}\}$ satisfies the conditions in Lemma B.2. Thus, there is a well-defined normative preference $\succsim^*$ defined over Δ and a well-defined function $\tau : \Delta\times\Delta \to \mathbb{R}$ as in Lemma B.1.

Lemma E.7. $\succsim^*$ *satisfies order, continuity and independence*

Proof. Lemma B.2 establishes that $\succsim^*$ is complete, transitive and continuous. To prove independence, it suffices by Lemma B.3 to show that each $\gtrapprox_{t|_\Delta}$ satisfies independence. Recall that $\gtrapprox_{t|_\Delta}$ is the restriction of $\gtrapprox_t$ to Δ, and apply Lemma E.2. ■

In the remainder of the subsection we establish a stationarity property of $\succsim^*$.

Lemma E.8. $(c, \mu) \succsim^* (c, \eta) \Longleftrightarrow (c', \mu) \succsim^* (c', \eta)$

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

By Separability, for all t,

$$(\frac{1}{2}(c, x) + \frac{1}{2}(c', x'))^{+t} \approx (\frac{1}{2}(c, x') + \frac{1}{2}(c', x))^{+t}.$$

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 B.4(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 $(c, \mu) \succsim^* (c, \eta)$ and $(c', \mu) \prec^* (c', \eta)$. Since $\succsim^*$ satisfies the vNM axioms,

$$(c, \mu)\frac{1}{2}(c', \eta) \succ^* (c, \eta)\frac{1}{2}(c', \mu). \tag{E.2}$$

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 $(c, \eta)\frac{1}{2}(c', \mu) \succ^* (c, \mu)\frac{1}{2}(c', \eta)$, contradicting (E.2). ■

Lemma E.9. $\mu^{+1}\alpha\eta^{+1} \sim^* (\mu\alpha\eta)^{+1}$

Proof. By Indifference to Timing, $\mu^{+1}\alpha\eta^{+1} \approx (\mu\alpha\eta)^{+1}$, and by Lemma E.6, 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 B.4(c) to obtain $\mu^{+1}\alpha\eta^{+1} \sim^* (\mu\alpha\eta)^{+1}$. ■

Lemma E.10. $\mu \succsim^* \eta \Longrightarrow (c, \mu) \succsim^* (c, \eta)$.

Proof. Step 1: $\mu \gtrapprox_{\tau(\mu,\eta)} \eta \Longleftrightarrow \mu^{+1} \gtrapprox_{\tau(\mu^{+1},\eta^{+1})} \eta^{+1}$.

First show that if $\tau(\mu, \eta) > 0$, then

$$\tau(\mu^{+1}, \eta^{+1}) = \tau(\mu, \eta) - 1. \tag{E.3}$$

For this purpose, suppose that $\tau(\mu,\eta) > 0$ and that, without loss of generality, $\mu^{+t} > \eta^{+t}$ for all $t \geq \tau(\mu,\eta)$ and $\mu^{+t} \lessapprox \eta^{+t}$ for all $t < \tau(\mu,\eta)$. It follows that

$$\begin{aligned} \mu^{+(t+1)} &> \eta^{+(t+1)}, \text{ for all } t \geq \tau(\mu,\eta)-1. \\ \text{and } \mu^{+(t+1)} &\lessapprox \eta^{+(t+1)}, \text{ for all } t < \tau(\mu,\eta)-1. \end{aligned}$$

The assertion follows. Now prove the result. It follows by definition of τ if $\tau(\mu,\eta) = 0$ (in which case $\tau(\mu^{+1},\eta^{+1}) = 0$ as well). When $\tau(\mu,\eta) > 0$, then note that $\mu^{+\tau(\mu,\eta)} = (\bar{c},\mu)^{+(\tau(\mu,\eta)-1)}$ and $\eta^{+\tau(\mu,\eta)} = (\bar{c},\eta)^{+(\tau(\mu,\eta)-1)}$. The result follows from (E.3).

Step 2: The result.

If $\mu \succsim^* \eta$, then by Lemma B.4(b), there exists a sequence $\{(\mu_n,\eta_n)\}$ such that $(\mu_n,\eta_n) \to (\mu,\eta)$ and $\mu_n \gtrapprox_{\tau(\mu_n,\eta_n)} \eta_n$ for all n. It follows by Step 1 that $\{(\mu_n^{+1},\eta_n^{+1})\}$ is a sequence such that $(\mu_n^{+1},\eta_n^{+1}) \to (\mu^{+1},\eta^{+1})$ and $\mu_n^{+1} \gtrapprox_{\tau(\mu_n^{+1},\eta_n^{+1})} \eta_n^{+1}$ for all n. But then by Lemma B.4(b), $\mu^{+1} \succsim^* \eta^{+1}$. The result follows by Lemma E.8. ■

Lemma E.11. $\mu \succsim^* \eta \Longleftrightarrow (c,\mu) \succsim^* (c,\eta)$.

Proof. By Lemma E.8, it suffices to show $\mu \succsim^* \eta \Longleftrightarrow (\bar{c},\mu) \succsim^* (\bar{c},\eta)$. Define a binary relation $\succsim^{**}$ over Δ by $\mu \succsim^{**} \eta \Longleftrightarrow (\bar{c},\mu) \succsim^* (\bar{c},\eta)$. We need to show $\mu \succsim^* \mu' \Longleftrightarrow \mu \succsim^{**} \mu'$. This follows from the three observations below:

(a) By Lemma E.10, $\mu \succsim^* \eta \Longrightarrow \mu \succsim^{**} \eta$.

(b) The preference $\succsim^{**}$ is non-trivial, that is, there exist $\rho,\nu \in \Delta$ such that $\rho \succ^{**} \nu$: By the nondegeneracy condition in Set-Betweenness, there exists x,y such that $x^{+t} > (x \cup y)^{+t}$. By Commitment is Normative, we can assume that $t > 1$. Commitment is Normative also implies $\tau(x^{+t},(x \cup y)^{+t}) = 0$ and τ is continuous at $(x^{+t},(x \cup y)^{+t})$, and so $(x^{+t},(x \cup y)^{+t}) \in \Omega$. By Lemma B.4(a), $x^{+t} \succ^* (x \cup y)^{+t}$. It follows that $x^{+(t-1)} \succ^{**} (x \cup y)^{+(t-1)}$, that is, $\succsim^{**}$ is non-trivial.

(c) The preference $\succsim^{**}$ satisfies the vNM axioms: Follows from Lemmas E.7 and E.9. ■

E.3. FT Preference $\succsim$

The candidate for the FT preference $\succsim$ over Z that generates $\mathcal{C}(\cdot)$ is defined by

$$x \succsim y \Longleftrightarrow (c,x) \succsim^* (c,y),$$

for some $c \in C$. By Lemma E.8, the preference $\succsim$ is invariant to the choice of c. We verify that $\succsim$ is an FT preference by checking that it satisfies the conditions in [18, Theorems 3.1 and 5.3].

Lemma E.12. *For any x, y,*

(a) $x >_t x \cup y$ for some $t \Longleftrightarrow x \succ x \cup y$.

(b) $x >_t x \cup y >_t y$ for some $t \Longleftrightarrow x \succ x \cup y \succ y$.

(c) $x >_t x \cup y \approx_t y$ for some $t \Longrightarrow x \succ x \cup y \sim y$.

Proof. Suppose $x >_t x \cup y$, or equivalently, $x^{+t} > (x \cup y)^{+t}$. By Commitment is Normative, $\tau(x^{+t}, (x \cup y)^{+t}) = 0$ and τ is continuous at $(x^{+t}, (x \cup y)^{+t})$, and so $(x^{+t}, (x \cup y)^{+t}) \in \Omega$. Hence, by Lemma B.4(a), $x^{+t} \succ^* (x \cup y)^{+t}$, and repeated application of Lemma E.11 establishes $x^{+1} \succ^* (x \cup y)^{+1}$, which in turn establishes the sufficiency part of (a). For the converse, note that by Lemma B.4(a), $x^{+1} \succ^* (x \cup y)^{+1}$ implies $x >_t x \cup y$ for $t = \tau(x^{+1}, (x \cup y)^{+1})$. A similar argument establishes (b).

Turn to (c). In what follows we prove that $x \cup y \approx_{t'} y$ for all $t' \geq t$, since then $(x \cup y)^{+1} \approx_{\tau((x \cup y)^{+1}, y^{+1})} y^{+1}$, and so by Lemma B.4(c), $(x \cup y)^{+1} \sim^* y^{+1}$, that is, $x \cup y \sim y$.

As before, $x >_t x \cup y$ implies $x >_{t'} x \cup y$ for all $t' \geq t$. It follows by Set-Betweenness that $x >_{t'} y$ for all $t' \geq t$. Then by Lemma E.3(b), for all $t' \geq t$,

$$x \cup y \approx_{t'} y \Longleftrightarrow \max_{\mu \in y}\{U_{t'} + V_{t'}\} \geq \max_{\eta \in x}\{U_{t'} + V_{t'}\}. \tag{E.4}$$

where $U_{t'}$ and $V_{t'}$ are as in (E.1). By Lemma E.4, for all $t' \geq t$,

$$\max_{\mu \in y}\{U_{t'} + V_{t'}\} \geq \max_{\eta \in x}\{U_{t'} + V_{t'}\} \Longleftrightarrow \max_{\mu \in y}\{U_t + V_t\} \geq \max_{\eta \in x}\{U_t + V_t\}. \tag{E.5}$$

By hypothesis, $x \cup y \approx_t y$. Thus, by (E.4) and (E.5), $x \cup y \approx_{t'} y$ for all $t' \geq t$, as desired. ■

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

Lemma E.13. *$\succsim$ satisfies Order**, Continuity**. Moreover, $\succsim$ satisfies Independence**: for all $\alpha \in (0, 1)$, $\{\mu\} \succ \{\eta\} \Longrightarrow \{\alpha\mu + (1 - \alpha)\nu\} \succ \{\alpha\eta + (1 - \alpha)\nu\}$.*

Proof. Follows from Lemmas E.7 and E.11, and definition of $\succsim$. ■

Lemma E.14. *$\succsim$ satisfies Stationarity**: $z \succsim z' \Longleftrightarrow \{(c, z)\} \succsim \{(c, z')\}$.*

Proof. This follows from Lemmas E.11 and E.8, and by definition of $\succsim$. ■

Recall the set Ω of points in $\Delta \times \Delta$ on which τ is upper semicontinuous defined by

$$\Omega = \{(\mu, \eta) \in \Delta \times \Delta : (\mu_n, \eta_n) \to (\mu, \eta) \Longrightarrow \limsup_{n \to \infty} \tau(\mu_n, \eta_n) \le \tau(\mu, \eta)\}.$$

Lemma E.15. *$\succsim$ satisfies Set-Betweenness**: $x \succsim y \Longrightarrow x \succsim x \cup y \succsim y$.*

Proof. Begin by establishing that if $\mu >_{\tau(\mu,\eta)} \eta$ then

$$(\mu^{+1}, \eta^{+1}) \in \Omega \Longrightarrow (\mu, \eta) \in \Omega. \tag{E.6}$$

Suppose $(\mu, \eta) \notin \Omega$, so that there exists a sequence $\{(\mu_n, \eta_n)\}$ that converges to (μ, η) and $\limsup_{n\to\infty} \tau(\mu_n, \eta_n) > \tau(\mu, \eta)$. Without loss of generality, $\tau(\mu_n, \eta_n) > \tau(\mu, \eta)$ for all n. Suppose by way of contradiction that $\limsup_{n\to\infty} \tau(\mu_n, \eta_n) = T < \infty$. Thus, there exists N such that for all $n \ge N$, $T + 1 > \tau(\mu_n, \eta_n)$. Also, for large enough n, $\mu_n >_{\tau(\mu,\eta)} \eta_n$, and since

$$T + 1 > \tau(\mu_n, \eta_n) > \tau(\mu, \eta),$$

it follows that for all large enough n, $\eta_n \gtrapprox_{T+1} \mu_n$. By continuity of $\gtrapprox_{T+1}$, $\eta \gtrapprox_{T+1} \mu$. But since $T+1 > \tau(\mu, \eta)$, this contradicts the hypothesis that $\mu >_{\tau(\mu,\eta)} \eta$. Therefore, $\limsup_{n\to\infty} \tau(\mu_n, \eta_n) = \infty$.

To complete the argument, observe that $\{(\mu_n^{+1}, \eta_n^{+1})\}$ is a sequence that converges to (μ^{+1}, η^{+1}), and by (E.3), $\limsup_{n\to\infty} \tau(\mu_n^{+1}, \eta_n^{+1}) = \infty$.[23] It follows that $(\mu^{+1}, \eta^{+1}) \notin \Omega$, thus proving (E.6).

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

$$x^{+1} \succsim^* y^{+1} \Longrightarrow x^{+1} \succsim^* (x \cup y)^{+1} \succsim^* y^{+1}.$$

Denote $\tau(x^{+1}, y^{+1})$ by τ and consider two cases.

---

[23] Note that we had assumed $\tau(\mu_n, \eta_n) > \tau(\mu, \eta)$ for all n, and so $\tau(\mu_n, \eta_n) > 0$ for all n, as required by (E.3).

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

Then by Lemma B.4(a), $x^{+1} >_\tau y^{+1}$. By Set-Betweenness, $x^{+1} \gtrapprox_t (x \cup y)^{+1} \gtrapprox_t y^{+1}$ for all $t \geq \tau$. Hence

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

and by Lemma B.4(c), $x^{+1} \succsim^* (x \cup y)^{+1} \succsim^* y^{+1}$.

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

By Lemma E.11, $x^{+\tau} \backsim^* y^{+\tau}$. Note that $\tau(x^{+\tau}, y^{+\tau}) = 0$. By Lemma B.4(a), either $x^{+\tau} \approx y^{+\tau}$ or $[x^{+\tau} \not\approx y^{+\tau}$ 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 B.4(c), $x^{+\tau} \backsim^* (x \cup y)^{+\tau} \backsim^* y^{+\tau}$. By repeated application of Lemma E.11, $x^{+1} \backsim^* (x \cup y)^{+1} \backsim^* y^{+1}$, as desired.

(ii) Now suppose $[x^{+\tau} \not\approx y^{+\tau}$ and $(x^{+\tau}, y^{+\tau}) \notin \Omega]$ and without loss of generality, $x^{+\tau} > y^{+\tau}$. Since $\tau(x^{+\tau}, y^{+\tau}) = 0$, we have $x^{+\tau} >_t y^{+\tau}$ for all t. By Set-Betweenness, $x^{+\tau} \gtrapprox_t (x \cup y)^{+\tau}$ for all t. If $x^{+\tau} >_t (x \cup y)^{+\tau}$ for some t, then Commitment is Normative implies that τ is continuous at $(x^{+\tau+t}, y^{+\tau+t})$, which implies $(x^{+\tau+t}, y^{+\tau+t}) \in \Omega$. By repeated application of (E.6), $(x^{+\tau}, y^{+\tau}) \in \Omega$, 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 B.4(c) and repeated application of Lemma E.11 implies $x^{+1} \backsim^* (x \cup y)^{+1}$. Furthermore, since by hypothesis, $x^{+1} \backsim^* y^{+1}$, transitivity of $\succsim^*$ implies $(x \cup y)^{+1} \backsim^* y^{+1}$. Put together, $x^{+1} \backsim^* (x \cup y)^{+1} \backsim^* y^{+1}$, as desired. ■

Lemma E.16. *$\succsim$ satisfies Indifference to Timing**: for any $\mu, \eta, \pi, \nu \in \Delta_s$, if $\mu^1 = \pi^1, \eta^1 = \nu^1, \varphi(\mu^2) = \varphi(\pi^2)$ and $\varphi(\eta^2) = \varphi(\nu^2)$, then $\{\mu, \eta\} \backsim \{\pi, \nu\}$.*

Proof. Observe that by Lemma E.6, $\{\mu, \eta\} \approx_t \{\pi, \nu\}$ for all $t \geq 1$. Thus, $\tau(\{\mu, \eta\}^{+1}, \{\pi, \nu\}^{+1}) = 0$ and $\{\mu, \eta\}^{+1} \approx \{\pi, \nu\}^{+1}$. From Lemma B.4(c), we see that $\{\mu, \eta\}^{+1} \backsim^* \{\pi, \nu\}^{+1}$, and so $\{\mu, \eta\} \backsim \{\pi, \nu\}$. ■

Lemma E.17. *$\succsim$ satisfies Temptation Stationarity**:*

$$x \succ x \cup y \Longleftrightarrow \{(c, x)\} \succ \{(c, x), (c, y)\}.$$

Proof. By Lemma E.12(a) and Menus Can Tempt, $x \succ x \cup y \iff x >_t x \cup y$ for some $t \iff \{(c,x)\} >_{t'} \{(c,x),(c,y)\}$ for some $t' \iff \{(c,x)\} \succ \{(c,x),(c,y)\}$, as desired. ■

Lemma E.18. *$\succsim$ is nondegenerate.*

Proof. Use the nondegeneracy condition in Set-Betweenness and Lemma E.12(b). ■

The above Lemmas establish that $\succsim$ satisfy the conditions of [18, Thm 3.1], and so there exist $\delta, \gamma \in (0,1)$, continuous functions $u, v : C \to \mathbb{R}$ and continuous linear functions $U, V : \Delta(C \times Z) \to \mathbb{R}$ and $W, \overline{V} : Z \to \mathbb{R}$ such that for all $x \in Z$,

$$\begin{aligned}
W(x) &= \max_{\mu \in x}\{U(\mu) + V(\mu) - \max_{\eta \in x} V(\eta)\}, \\
U(\mu) &= \int_{C \times Z} (u(c) + \delta W(y))\, d\mu(c,y), \\
V(\mu) &= \int_{C \times Z} \big(v(c) + \gamma \overline{V}(y)\big)\, d\mu(c,y), \text{ where } \overline{V}(x) = \max_{\eta \in x} V(\eta)
\end{aligned}$$

and furthermore, W represents $\succsim$. It remains to show that $\gamma < \delta$ and that $\succsim$ generates $\mathcal{C}(\cdot)$.

Lemma E.19. *$\succsim$ generates $\mathcal{C}(\cdot)$.*

Proof. By nondegeneracy of $\succsim$, U is not an affine transformation of V (Lemma C.2). Lemmas E.12(a) and E.12(b) establish that for all t, $\succsim$ has more preference for commitment than $\gtrapprox_t$ and $\succsim$ has more self-control than $\gtrapprox_t$; see GP for definitions of these terms. By [10, Theorem 8], for each t,

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

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

$$U_t + V_t = (\alpha + \alpha')U + (2 - \alpha - \alpha')V. \tag{E.7}$$

Furthermore, by [10, Theorem 9],

$$U_t + V_t = \beta(U + V) + (1-\beta)V = \beta U + V, \tag{E.8}$$

for $\beta \in [0,1]$. Together, (E.7) and (E.8) 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 E.4,

$$\mu \gtrapprox \eta \iff U(\mu) + V(\mu) \geq U(\eta) + V(\eta).$$

By Lemma E.1, $\gtrapprox$ rationalizes $\mathcal{C}(\cdot)$. Therefore the above displayed equivalence implies that for all $x \in Z$,

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

as desired. ■

Lemma E.20. $\gamma < \delta$.

Proof. By Set-Betweenness, there is x, y and t such that $x \subset y$ and $x^{+t} > y^{+t}$. Thus,

$$\begin{aligned}
& x^{+t} > y^{+t} \\
\iff & U(x^{+t}) + V(x^{+t}) > U(y^{+t}) + V(y^{+t}) \qquad \text{by Lemma E.4} \\
\iff & \sum_{i=0}^{t-1} \delta^i u(\bar{c}) + \gamma^i v(\bar{c}) + \delta^t W(x) + \gamma^t \overline{V}(x) > \sum_{i=0}^{t-1} \delta^i u(\bar{c}) + \gamma^i v(\bar{c}) + \delta^t W(y) + \gamma^t \overline{V}(y) \\
\iff & W(x) + \tfrac{\gamma^t}{\delta^t}\overline{V}(x) > W(y) + \tfrac{\gamma^t}{\delta^t}\overline{V}(y).
\end{aligned}$$

Since $\overline{V}(y) \geq \overline{V}(x)$ (recall that $x \subset y$) and $\frac{\gamma^t}{\delta^t} > 0$, conclude that $W(x) > W(y)$, and hence by Lemma E.3(a) that $\overline{V}(y) > \overline{V}(x)$. Suppose by way of contradiction that $\frac{\gamma^t}{\delta^t} > 1$. Then the preceding implies $y^{+T} \gtrapprox x^{+T}$ for a large enough $T > t$, that is, $\tau(x^{+t}, y^{+t}) > 0$. But, by Commitment is Normative, $\tau(x^{+t}, y^{+t}) = 0$, a contradiction. Thus, $\frac{\gamma^t}{\delta^t} \leq 1$.

Suppose by way of contradiction that $\frac{\gamma^t}{\delta^t} = 1$. Then, for all μ, η and $t > 0$,

$$\begin{aligned}
& \mu^{+t} \gtrapprox \eta^{+t} \\
\iff & U(\mu^{+t}) + V(\mu^{+t}) \geq U(\eta^{+t}) + V(\eta^{+t}) \\
\iff & W(\mu) + \tfrac{\gamma^t}{\delta^t}\overline{V}(\mu) \geq W(\eta) + \tfrac{\gamma^t}{\delta^t}\overline{V}(\eta) \\
\iff & W(\mu) + \overline{V}(\mu) \geq W(\eta) + \overline{V}(\eta) \\
\iff & U(\mu) + V(\mu) \geq U(\eta) + V(\eta) \\
\iff & \mu \geq \eta.
\end{aligned}$$

That is, $\tau(\mu, \eta) = 0$ for all μ, η. This contradicts the latter part of Reversal. Hence, $\frac{\gamma^t}{\delta^t} < 1$, and the assertion follows. ■

F. Appendix: Proof of Theorem 3.1 (Uniqueness)

Lemma F.1. *If $U(\mu) = U(\eta)$ and $\mu > \eta$, then $(\mu, \eta) \notin \Omega$.*

Proof. By the definition of $\gtrapprox$, $U(\mu) = U(\eta)$ and $\mu < \eta$ imply $V(\mu) < V(\eta)$. By nondegeneracy, there exists ν, ρ such that $U(\nu) > U(\rho)$ and $V(\nu) < V(\rho)$. Consider the sequence $\{(\mu\alpha_n\nu, \rho\alpha_n\eta)\}$ that converges to (μ, η). Since U, V are linear, for each n, $U(\mu\alpha_n\nu) > U(\rho\alpha_n\eta)$ and $V(\mu\alpha_n\nu) < V(\rho\alpha_n\eta)$ and by Lemma D.6,

$$\mu\alpha_n\nu >_t \rho\alpha_n\eta \text{ for all } t \geq \tau(\mu\alpha_n\nu, \rho\alpha_n\eta). \tag{F.1}$$

The hypothesis ($U(\mu) = U(\eta)$ and $\mu < \eta$) implies that for all t,

$$\mu <_t \eta. \tag{F.2}$$

Suppose by way of contradiction that $\limsup_{n\to\infty} \tau(\mu\alpha_n\nu, \rho\alpha_n\eta) = T \leq \tau(\mu, \eta)$. Then there exists N such that

$$\tau(\mu\alpha_n\nu, \rho\alpha_n\eta) < T + 1, \quad \text{for all } n \geq N. \tag{F.3}$$

However, $\mu <_{T+1} \eta$ by (F.1) and so there exists N' such that

$$\mu\alpha_n\nu <_{T+1} \rho\alpha_n\eta, \quad \text{for all } n \geq N'.$$

But by (F.1) this implies $\tau(\mu\alpha_n\nu, \rho\alpha_n\eta) > T + 1$ for all $n \geq \max\{N, N'\}$, contradicting (F.3). ■

Lemma F.2. *If $x \succ y$, then there exists T such that $(c, x) >_t (c, y)$ for all $t \geq T$.*

Proof. The hypothesis implies $U(c, x) > U(c, y)$, and the result follows from Lemma D.6. ■

Lemma F.3. *If $x \sim y$, then $\tau(x^{+1}, y^{+1}) = 0$.*

Proof. Since, for any t,

$$(c, x) \gtrapprox_t (c, y) \iff W(x) + \frac{\gamma^{t+1}}{\delta^{t+1}}\overline{V}(x) \geq W(y) + \frac{\gamma^{t+1}}{\delta^{t+1}}\overline{V}(y),$$

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

Lemma F.4. *$\mathcal{C}(\cdot)$ is generated by a unique FT preference $\succsim$.*

Proof. Suppose, by way of contradiction, that $\succsim$ and $\succsim'$ are two FT preferences that generate $\mathcal{C}(\cdot)$ and that there exist x and y such that $x \succ y$ and $y \succsim' x$. Let (U, V) and (U', V') be representations of $\succsim$ and $\succsim'$, respectively. If $x \succ y$ and $y \succ' x$, then by Lemma F.2, $\succsim$ and $\succsim'$ do not generate the same choice correspondence, a contradiction. So suppose that $x \succ y$ 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 D.7, $(x^{+1}, y^{+1}) \in \Omega$. We show that $(x^{+1}, y^{+1}) \notin \Omega$ also holds, thereby establishing the desired contradiction. By Lemma F.3, $y \sim' x$ implies $\tau(x^{+1}, y^{+1}) = 0$. Then Lemma F.2 and $x \succ y$ 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 F.1, $(x^{+1}, y^{+1}) \notin \Omega$. ■

G. Appendix: Proof of Theorem 4.2

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

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

For each $t > 0$, define $\gtrapprox_t$ on Δ by

$$\mu \gtrapprox_t \eta \Longleftrightarrow \mu^{+t} \gtrapprox \eta^{+t}.$$

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

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

Lemma G.1. *The sequence $\{\varphi_t\}$ uniformly converges to U.*

Proof. The sequence $\{\varphi_t\}$ is a sequence of continuous real functions defined on a compact space Δ. Since $V \geq 0$ and $\frac{\gamma}{\delta} < 1$, the sequence is monotone decreasing and φ_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 $x \succ y$. 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,

$$U(\mu) \geq U(\eta) \Longrightarrow U(\mu\alpha\rho) > U(\eta\alpha\nu), \text{ for all } \alpha \in (0,1)\,. \tag{G.1}$$

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 on Δ with its graph $\Gamma(B) \subset \Delta \times \Delta$.

Lemma G.2. $\Gamma(\succsim_U) = \lim_{t\to\infty}\Gamma(\gtrapprox_t)$.

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

$$\varphi_{t(n)}(\mu_n) \geq \varphi_{t(n)}(\eta_n).$$

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

Next establish $\Gamma(\succsim_U) \subset Li\Gamma(\gtrapprox_t)$. Let $(\mu,\eta) \in \Gamma(\succsim_U)$ and take any neighborhood V of (μ,η). By (G.1), there exists $\alpha \in (0,1]$ s.t $(\mu\alpha\rho, \eta\alpha\nu) \in V$ and $U(\mu\alpha\rho) > U(\eta\alpha\nu)$. By Lemma G.1, there exists $T < \infty$ such that $\varphi_t(\mu\alpha\rho) > \varphi_t(\eta\alpha\nu)$ for all $t \geq T$, that is, $(\mu\alpha\rho, \eta\alpha\nu) \in \Gamma(\gtrapprox_t)$ for all $t \geq T$. Hence,

$$V \cap \Gamma(\gtrapprox_t) \neq \phi \text{ for all but a finite number of } t,$$

that is, $(\mu,\eta) \in Li\Gamma(\gtrapprox_t)$. This completes the proof. ■

By Lemma G.2, $\succsim_U = \succsim^*$, 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 $\succsim$, [10, Theorem 4] implies that for any α 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 $\succsim^*$.

H. Appendix: Proof of Theorem 5.1

Necessity of each axiom is either trivial or as in Appendix D. We prove sufficiency. Define $\{\gtrapprox_t\}$ over Z as in Appendix E.1 and note that by Preferences Reverse Tomorrow, $\gtrapprox_t = \gtrapprox_{t'}$ for all $t, t' > 0$. Let $\succsim' = \gtrapprox_1$, and take $\succsim'$ is the candidate CT preference. It is readily verified that $\succsim'$ is nondegenerate and satisfies the axioms Order**, Independence**, Continuity**, Set-Betweenness**, Separability** and Indifference to Timing** stated in Appendix E.1. Preferences Reverse Tomorrow implies that $\succsim'$ satisfies Stationarity**, and by Lemmas E.1 and E.4, Sophistication implies that $\succsim'$ generates $\mathcal{C}(\cdot)$.

Given the restrictions on $\succsim'$, it can be shown that there is a representation W' such that

$$W'(x) = \max_{\mu\in x}\int_{C\times Z}\Big(u(c) + \delta W'(y) + v(c) + \widehat{V}(y)\Big)\,d\mu(c,y) - \max_{\eta\in x}\int_{C\times Z}\Big(v(c) + \widehat{V}(y)\Big)\,d\eta(c,y),$$

where $\widehat{V}$ is linear and continuous. See the proof of [18, Theorem 3.1]. We proceed to show that $\widehat{V}$ is cardinally equivalent to W.

Lemma H.1. $\{(c,x)\} \succsim' \{(c,y)\} \Longrightarrow \{(c,x)\} \sim' \{(c,x),(c,y)\}$

Proof. Follows from Menus Do Not Tempt. ■

Lemma H.2. *There exists $\gamma \geq 0$ and θ such that for all x,*

$$\widehat{V}(x) = \gamma W'(x) + \theta.$$

Proof. If $\widehat{V}$ is constant (equal to some θ), then the result follows with $\gamma = 0$. Therefore suppose that $\widehat{V}$ is nonconstant. Suppose by way of contradiction that $\widehat{V}$ is not ordinally equivalent to W'. That is, there exists x and y such that

$$\begin{aligned} \widehat{V}(x) &\geq \widehat{V}(y) \text{ and } W'(x) < W'(y), \\ \text{or } \widehat{V}(x) &> \widehat{V}(y) \text{ and } W'(x) \leq W'(y). \end{aligned}$$

We show that in either case, nonconstancy of $\widehat{V}$ and W' implies the existence of menus for which both inequalities are strict. We prove this for the case that $\widehat{V}(x) = \widehat{V}(y)$ and $W'(x) < W'(y)$. The same argument can be applied to the other case, that is, $\widehat{V}(x) > \widehat{V}(y)$ and $W'(x) = W'(y)$. So suppose $\widehat{V}(x) = \widehat{V}(y)$ and $W'(x) < W'(y)$. There are two possibilities to consider. First, there is z such that $\widehat{V}(x) > \widehat{V}(z)$.[24] If $W'(x) < W'(z)$, there is nothing to prove. If $W'(z) \leq W'(x)$, then $W'(z) < W'(y)$, and since $\widehat{V}(y) > \widehat{V}(z)$, the assertion is proved for this case as well. Second, there is z such that $\widehat{V}(z) > \widehat{V}(x)$. The argument is similar. Since $\widehat{V}$ is nonconstant, one of the two possibilities must be true, and hence we are done.

Therefore we can assume without loss of generality that

$$W'(x) > W'(y) \text{ and } \widehat{V}(x) < \widehat{V}(y).$$

Observe that

$$W'\{(c,x)\} = u(c) + \delta W'(x) > u(c) + \delta W'(y) = W'\{(c,y)\}.$$

If $\delta W'(x) + \widehat{V}(x) > \delta W'(y) + \widehat{V}(y)$, then

$$W'\{(c,x)\} = u(c) + \delta W'(x) > u(c) + \delta W'(x) + \widehat{V}(x) - \widehat{V}(y) = W'(\{(c,x),(c,y)\}).$$

---

[24] A $\widehat{V}-$best and worst menu exists since $\widehat{V}$ is continuous and Z is compact.

This contradicts Lemma H.1. On the other hand, if $\delta W'(x) + \widehat{V}(x) \leq \delta W'(y) + \widehat{V}(y)$, then

$$W'\{(c,x)\} = u(c) + \delta W'(x) > u(c) + \delta W'(y) = W'\{((c,x),(c,y)\},$$

again contradicting Lemma H.1. Thus we establish that $\widehat{V}$ is ordinally equivalent to W'. Since $\widehat{V}$ and W' are also linear, it follows that they must be cardinally equivalent. Thus, there exists $\gamma \geq 0$ and θ such that for all x,

$$\widehat{V}(x) = \gamma W'(x) + \theta,$$

as was to be shown. ■

Without loss of generality, we can set $\theta = 0$. Therefore, $\succsim'$ is represented by the function defined by:

$$\begin{aligned} W'(z) &= \max_{\mu \in z} \int_{C \times Z} (u(c) + \delta W'(x) + v(c) + \gamma W'(x))\, d\mu(c,x) \\ &\quad - \max_{\eta \in z} \int_{C \times Z} (v(c) + \gamma W'(y))\, d\eta(c,y). \end{aligned}$$

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

$(c,x) \in \mathcal{C}(\{(c,x),(c,y)\})$

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

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

$\iff x \succsim' y.$

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

I. Appendix: Proof of Theorem 5.2

The argument used in the proof of Theorem 4.2 goes through. The only modification is that each $\gtrapprox_t$ 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

[1] Ainslie, G. (1975): 'Specious Reward: A Behavioral Theory of Impulsiveness and Impulse Control', *Psychological Bulletin* 82, pp 463-496.

[2] Ainslie, G. (1992): *Picoeconomics*, Cambridge University Press.

[3] Ainslie, G. and V Haendel (1983): ‘The Motives of the Will’, in Gotteheil et al (Eds), *Etiologic Aspects of Alcohol and Drug Abuse* , Springfield, Il: Thomas.

[4] Aliprantis, C. and K. Border (1994): *Infinite Dimensional Analysis: a Hitchhiker's Guide*, 2nd Edition, Springer Verlag.

[5] Benartzi, S. and R. Thaler (2004): ‘Save More Tomorrow$^{\mathrm{TM}}$: Using Behavioral Economics to Increase Employee Savings’, *Journal of Political Economy*, Vol. 112, No. 1, pp. S164-S187.

[6] Brown, A. and C. Pearcy (1995): *An Introduction to Analysis*, Springer, New York.

[7] Fishburn, P. (1970), *Utility Theory for Decision Making*, John Wilet and Son, New York. Reprinted by Krieger Press, Huntington, New York, 1979.

[8] Fuller, R., and H. Roth (1979): ‘Disulfiram for the Treatment of Alcoholism’, *Annals of Internal Medicine* 90, pp 901-904.

[9] Goldstein, A. (2001): *Addiction: From Biology to Drug policy*, Second Edition, New York: Oxford University Press.

[10] Gul, F. and W. Pesendorfer (2001): ‘Temptation and Self-Control’, *Econometrica* 69, pp 1403-1435.

[11] Gul, F. and W. Pesendorfer (2003): ‘Self-Control and the Theory of Consumption’, *Econometrica*, 72, pp 119-158.

[12] Hildenbrand, W. (1974): *Core and Equilibria of a Large Economy*, Princeton University Press.

[13] Kirby, K., and R. Hernnstein (1995): ‘Preference Reversals due to Myopic Discounting of Delayed Reward’, *Psychological Science*, pp 83-89.

[14] Krusell, P., B. Kurusçu and A. Smith, Jr. (2002): ‘Time Orientation and Asset Prices’, *Journal of Monetary Economics* 49(1), pp 107-135.

[15] Laibson, D. (1997): ‘Golden Eggs and Hyperbolic Discounting’, *Quarterly Journal of Economics* 112, pp 443-77.

[16] Mas-Collel, A., M. D. Whinston and J. R. Green (1995): *Microeconomic Theory*, New York: Oxford University Press.

[17] Mazur, J. E. (1987): ‘An Adjustment Procedure for Studying Delayed Reinforcement’ in M. L. Commons, J. E. Mazur, J. A. Nevin and H. Rachlin (Eds), *Quantitative Analyses of Behavior: Vol. 5. The Effect of Delay and Intervening Events on Reinforcement Value*, pp 55-73, Hillsdale, NJ: Erlbaum.

[18] Noor, J. (2005): ‘Commitment and Self-Control’, University of Rochester Working Paper.

[19] Noor, J. (2005): ‘Normative Preference and Choice’, University of Rochester Working Paper.

[20] O’Donoghue, T., and M. Rabin (1999): ‘Doing it Now or Later’, *American Economic Review*, 89, pp 103-24.

[21] Rachlin, H. (1970): *Introduction to Modern Behaviorism*, San Francisco: W. H. Freeman.

[22] Rachlin, H. and L. Green (1972): ‘Commitment, Choice and Self-Control’, *Journal of the Experimental Analysis of Behavior* 17, pp 15-22.

[23] Strotz R. (1955): ‘Myopia and Inconsistency in Dynamic Utility Maximization’, *Review of Economic Studies* XXIII, pp 165-180.