Revealed Willpower*

Yusufcan Masatlioglu[†] Daisuke Nakajima[‡] Emre Ozdenoren[§]

Abstract

The willpower as limited cognitive resource model has been proposed by experimental psychologists and used by economists to capture lack of self control and explain various behavioral paradoxes. This paper provides the first behavioral foundation for the limited willpower model which bridges the standard utility maximization and the Strotz models. In our model, we observe the agent's ex-ante preferences and ex-post choices and derive a representation that captures key behavioral traits of willpower constrained decision making. When the willpower stock is not too high or too low, choices reflect a compromise between the ex-ante preference and ex-post temptation and violate WARP. In a riskless domain, we provide simple axioms that give a clean comparison with other self control models. In the lottery domain, our characterization has stronger uniqueness properties. We illustrate by means of an IO application that the model is tractable and provides distinct insights.

1 Introduction

Standard theories of decision making assume that people choose what they prefer and prefer what they choose. However, introspection suggests that implementation of choice may not be automatic and there is often a wedge between preferences and actual choices. Recently psychologists and economists have emphasized the lack of self control in decision making as an important reason for this wedge.[1] When people face temptation, they make choices that are in conflict with their commitment preferences. Procrastination, impulse purchases, and succumbing to the temptation of unhealthy foods are some common examples of such behavior.

*We would like to thank Kfir Eliaz, Collin Raymond, Larry Samuelson, Joel Sobel, Rani Spiegler, Tomasz Strzalecki, and Neslihan Uler. Parts of this research were conducted while Masatlioglu was visiting the Institute for Advanced Studies at Hebrew University. Nakajima appreciates the financial support provided by Japan Society for the Promotion of Science (JSPS KAKENHI Grant Number 26780113).

†Department of Economics, University of Michigan. Email: yusufcan@umich.edu.

‡Otaru University of Commerce. Email: nakajima@res.otaru-uc.ac.jp.

§London Business School. Email: eozdenoren@london.edu.

[1] Models of self-control problems include quasi-hyperbolic time discounting (e.g., Laibson (1997); O'Donoghue and Rabin (1999)), temptation costs (e.g., Gul and Pesendorfer (2001, 2004)), and conflicts between selves or systems (e.g., Shefrin and Thaler (1988); Bernheim and Rangel (2004); Fudenberg and Levine (2006)).

People do not always succumb to temptation and are sometimes able to overcome temptations by using cognitive resources. This ability is often called willpower.[2] There is a growing experimental psychology literature demonstrating that willpower is a limited resource, and it is more than a mere metaphor (e.g., Baumeister and Vohs (2003); Faber and Vohs (2004); Muraven et al. (2006).)[3]

In this paper we propose a simple and tractable model to capture limited willpower and provide behavioral foundations for it that allow us to go beyond simple intuition, understand what type of observed behavior would characterize the limited willpower model, and pave the way for designing new experiments to test the willpower theory. Moreover, the representation theorem gives a precise meaning to the term willpower stock. We use the model to study an example of monopolistic contracting where consumers have limited willpower but are unaware of their willpower problems. The example illustrates that the limited willpower model has policy related implications that are distinct from those of other models of self-control.

In Section 2, we present the limited willpower model which is based on three ingredients. The first, commitment utility u, represents the decision maker's (henceforth DM) commitment preferences. The other two ingredients are temptation values v and the willpower stock w which jointly determine how actual choices depart from what commitment utility would dictate. The key to determining the actual choice is the willpower constraint. Specifically, the DM is able to consider an alternative x from a set A only if $\max_{y\in A} v(y) - v(x) \leq w$. Otherwise, she does not have enough willpower to choose this alternative. The DM then picks the alternative that maximizes her commitment utility from the set of alternatives that satisfies the willpower constraint. Formally, the ex-post choice from a set A is the outcome of the following maximization problem:

$$\max_{x\in A} \quad u(x) \quad \textit{subject to} \quad \max_{y\in A} v(y) - v(x) \leq w$$

The limited willpower model bridges the standard utility maximization and the Strotz models. When the willpower stock is very large, the willpower-constrained DM behaves like a standard agent who chooses the most preferred alternative (according to u). When the willpower stock is lower, the constraint starts to bind and a wedge between preferences and choices appears – the DM can only choose alternatives that are close enough, in terms of temptation, to the most tempting one. In the other extreme, when the willpower stock is very low, the DM behaves like a Strotzian

[2]Loewenstein (2000) emphasizes the role of both positive and negative visceral urges in generating the wedge between choice and preference. Following self-control literature we focus on temptations or positive visceral urges. However similar issues are relevant when individuals face fear or guilt inducing alternatives that cause negative visceral reactions. In those cases willpower might be necessary to motivate oneself to choose an alternative that causes a more intense negative urge than another feasible alternative.

[3]Models motivated by these experiments have been used by economists to capture lack of self-control (Ozdenoren et al. (2012); Fudenberg and Levine (2012)) in a dynamic settings. In those models agents take into account willpower consequences of their current choices for their future choices.

agent who always succumbs to temptation. Notice that the DM's choice will satisfy WARP in the two extreme cases for different reasons. While in the former choices reflect the ex-ante preference alone, in the latter, temptation ranking solely determines the choices. In comparison, in the limited willpower case, when the willpower stock is not too high or too low, choices reflect a compromise between the ex-ante preference and the temptation ranking and violate WARP. To see this suppose there are three alternatives, one alternative (x) is ex-ante best but least tempting, a second (z) is ex-ante worst but most tempting and a third (y) is in the middle both in terms of ex-ante preference and temptation. A standard DM chooses x and a Strotzian agent chooses z from $\{x, z\}$ and from $\{x, y, z\}$, and in both cases WARP will be satisfied. A DM with limited willpower may choose z from $\{x, z\}$ and y from $\{x, y, z\}$ violating WARP.[4] Such behavior, resembling the compromise effect, occurs when the DM does not have enough willpower to choose the least tempting alternative, but has enough willpower to choose the moderately tempting alternative.

To derive the limited willpower representation, we use a novel data set given by the DM's ex-ante preferences ($\succsim$) and ex-post choices (c).[5] Our data set differs from the classical one given by the DM's ex-ante preferences over menus of alternatives. There are two advantages to working with our data set. First, for a DM, revealing ex-ante preferences over *all menus of alternatives* is a cognitively demanding task. In contrast, revealing her ex-ante preferences over *alternatives* (which can be identified with preferences over singleton menus) is a simpler and more natural task. Second, menu preferences implicitly assume sophistication and derive the DM's ex-post choices from her ex-ante ranking of menus.[6] Our data set allows us to remain agnostic about whether the DM is sophisticated or naive about anticipating her ex-post choices and take ex-post choices as given. This is useful in applications where the agents are often assumed to be naive about their self-control problems (as we do in our monopolistic contracting example in Section 5).

We first provide two representation theorems in an environment with finitely many, riskless alternatives (in contrast to lotteries over them.) Working in a riskless domain has two important advantages. First, psychology experiments on willpower typically involve choices without risk. Second, it enables us to identify the key behavioral postulates underpinning the limited willpower model without additional structure.

In this simple domain, our first axiom is that the DM's ex-ante preferences are complete and

[4]Similar examples of WARP violations feature in Fudenberg and Levine (2006); Dekel et al. (1998); Noor and Takeoka (2010). In a menu preference framework, Dillenberger and Sadowski (2012) point out that when agents anticipate experiencing guilt or shame when they deviate from a social norm, their choices can also violate WARP.

[5]Ahn and Sarver (2013) also utilized two kinds of behavioral data. As opposed to ours, to derive their result, they utilize both the entire menu preferences and the *random* ex-post choice from menus.

[6]There is evidence of "hot-cold empathy gaps" where individuals are not able to appreciate the intensity of temptation, or other visceral urges at an ex-ante state. Loewenstein and Schkade (1999) review several studies that find people tend to underestimate the influence on their behavior of being in a hot state (such as hunger, drug craving, curiosity, sexual arousal, etc.).

transitive. The second axiom, Independence of Preferred Alternative (IPA), says that dropping any alternative that is strictly preferred to the chosen one should not affect the actual choice.[7] The third axiom, Choice Betweenness, states that the choice from the union of two sets is "between" the choices made separately from each set with respect to preference.[8] Our first representation theorem shows that these three axioms are necessary and sufficient for $(\succsim, c)$ to be represented by a *generalized* limited willpower model in which the willpower stock depends on the chosen alternative. To characterize the limited willpower model we need an additional axiom. This fourth axiom, Consistency, formalizes the intuition that if y is more tempting than z and the DM prefers x but cannot choose it against z, then the DM should not be able to choose x against y either. Our second representation shows that these four axioms are necessary and sufficient for $(\succsim, c)$ to be represented by a limited willpower model.

These four axioms are simple and intuitive, but the simplicity comes at a cost. The finite representation is not unique even in an ordinal sense. This observation motivates us to consider an environment where alternatives are lotteries over outcomes. To prove the representation theorem in the lottery domain, we introduce several additional axioms to take advantage of the additional structure provided by the lottery domain. In addition to some technical axioms we use two independence axioms imposed on the choice correspondence. These axioms are suitable relaxations of the full independence axiom in our setting. Given that we are ultimately interested in conflict between preferences and choices, we introduce an axiom, Conflict, to rule out cases where a conflict between them never arises. In other words, Conflict requires that there are situations where the DM prefers one lottery over another but when she has to make a choice between the two she chooses the less preferred lottery. Finally, Limited Agreement axiom restricts the extent of the conflict that occurs between preferences and choices. It says that when the two lotteries are very close, the DM's choices must agree with her preferences.

Issues of temptation and self-control have been studied using the preference over menus framework pioneered by Kreps (1979).[9] Under sophistication, DM's preferences over menus have a two-period interpretation and reveal her anticipated second period choices from the menu. In Gul and Pesendorfer (2001)'s costly self-control model the implied second period choices satisfy WARP. However, others have noted that under temptation intuitive choice patterns like the compromise effect might be observed and these choices violate WARP. Working in the lottery domain, Noor and Takeoka (2010) retain all the axioms of the costly self-control model except for independence.

[7] This axiom can be viewed as a relaxation of WARP that says that any unchosen alternative can be dropped without affecting actual choices. In contrast, IPA allows only strictly preferred alternatives to be dropped without affecting actual choices.

[8] Although at first glance this axiom seems like a translation of Gul and Pesendorfer's Set Betweenness axiom to our domain, the two axioms are independent. We discuss this point in Section 3.

[9] For a survey see Lipman and Pesendorfer (2011).

They show that the resulting menu dependent self-control model generates implicit second period choices that violate WARP. In the limited willpower model, DM's choices also generate these choice patterns. However, the two approaches differ for two reasons. First, since our axiomatization relies directly on the DM's choices, and does not require the additional assumption of sophistication. Second, we can go beyond the observation that choices violate WARP, and understand the testable implications of different representations of ex-post choices. For example, the choices generated by the menu dependent self-control and the limited willpower models both satisfy IPA and Choice Betweenness. As a result, they are both special cases of the generalized limited willpower model. However, the choices generated by the convex self-control model typically violate the Consistency axiom, showing that the models are distinct. In Section 3 we expand on this point.

In Section 5, we present a simple example, based on a two-period model of monopolistic contracting between a monopolist and a naive consumer who has self-control problems. Several papers have studied contracts in this setting motivated by the fact that consumers sign up for phone plans, gym-membership, or credit card plans that seem exploitative. Our example illustrates that the willpower model delivers distinct predictions compared to these papers. Specifically, we show that when the consumer has positive willpower the optimal contract consists of three alternatives and the consumer's choices reflect a form of the "compromise effect" which is induced endogenously by the contract. The example has several interesting implications. Interestingly, the optimal contract includes an alternative that neither the consumer nor the firm believes would be chosen from an ex-ante perspective, and indeed is not chosen ex-post. Hence, the example suggests a new reason for why contracts in the real-world may seem excessively complex.

In Section 6.1, we discuss the implication of our results for the interpretation of the large literature on "willpower experiments" (for example Baumeister et al. (1994); Baumeister and Vohs (2003)). These experiments have demonstrated that individuals depleted by prior acts of self-restraint tend to behave later as if they have less self-control which is often viewed as prima facie evidence of limited willpower. We argue that there is an alternative explanation where prior act of self restraint alter subjects' temptation rankings. We then derive conditions that would allow an experimenter to eliminate this alternative explanation and conclude that the only impact of a prior act of self restraint is on the willpower stock.

Finally, in Section 6.2 we discuss identification of the utility function u, temptation ranking v and the willpower stock w. We show that, in the lottery domain, identification can be based on the DM's ex-post choices and ex-ante preferences can be constructed from ex-post choices (even though the choices violate WARP). To the best of our knowledge the limited willpower model is the first model that captures self-control problems entirely from ex-post choice data. We should highlight that the lottery domain is the key for this result.[10]

2 Model

Let X be a finite set of alternatives. The DM's *ex-ante* preferences $\succsim$ are over Z. In the riskless domain $Z = X$ and in the lottery domain $Z = \Delta$ where Δ is the set of all simple lotteries over X. These preferences can be interpreted as the DM's commitment preferences. The DM's *ex-post* choices are captured by a choice correspondence c that assigns a non-empty subset of A to each $A \in \mathcal{X}$ where $\mathcal{X}$ is the set of all non-empty subsets of Z.

We say that $(\succsim, c)$ has a *generalized limited willpower representation* if there exists (u, v, w) where $u : Z \to R$ represents preference $\succsim$ and c is given by

$$c(A) = \underset{x \in A}{\operatorname{argmax}} \quad u(x) \quad \textit{subject to} \quad \max_{y \in A} v(y) - v(x) \leq w(x)$$

where $v : Z \to R$ captures the temptation values and $w : X \to \mathbb{R}_+$ is the willpower function. If w is a constant function, we call it simply a *limited willpower representation.* Furthermore, in the lottery domain, both u and v are linear.

In the standard model where there is no willpower problem, a DM chooses the alternative that maximizes u from any menu. A DM who has limited willpower also maximizes u but faces a constraint. The willpower requirement of alternative x is given by the difference between the temptation value of the most tempting alternative on the menu, $\max_{y \in A} v(y)$, and the temptation value of x. The DM can choose x only if its willpower requirement is less than the willpower stock, w. Otherwise, the DM does not have enough willpower to choose this alternative. Notice that the willpower requirement is menu dependent. This is because willpower depletion not only depends on how tempting the chosen alternative is but also on the most tempting alternative on the menu.

As a simple example consider three alternatives: going to the gym (g), reading a book (g) or watching TV (t). Suppose ex-ante $g \succ b \succ t$. Suppose $v(g) = 0$, $v(b) = 2$, $v(t) = 4$. Table 1 shows the DM's choices from two sets, $\{g, b, t\}$ and $\{g, b\}$ for varying levels of willpower stock.[11] When willpower stock is high, $w = 5$, the DM chooses according to her ex-ante preferences. When willpower stock is low the DM also behaves like a standard preference maximizer, except that she chooses the most tempting alternative. When the willpower stock is intermediate, $w = 3$, then the

[10] In the context of rational attention, Ellis (2013) made a similar point where ex-post choices rationalize rational inattention. However, his choice data is richer that ours, namely the DMs choices from each feasible set of acts and conditional on each state of the world.

[11] We will abuse the notation and write $c(x, y, \dots)$ instead of $c(\{x, y, \dots\})$ and when the choice is unique, $x = c(x, y, \dots)$. Similarly, we omit braces and write $A \cup x$ instead of $A \cup \{x\}$.

model has interesting implications – decisions can be driven by a compromise between the ex-ante preference and temptation. To see this suppose all three alternatives are available. The DM is not able to choose g since $v(t) - v(g) = 4 > 3 = w$. In this case she chooses the compromise alternative b since $v(t) - v(b) = 2 < 3$. However when only g and b are available, there is no need to compromise (since $v(b) - v(g) = 2 < 3$) and the DM chooses g.

	$w = 1$	$w = 3$	$w = 5$
$c(g, b, t)$	t	b	g
$c(g, b)$	b	g	g

Table 1: Choices for different levels of the willpower stock

3 Riskless Domain

In this section we introduce the axioms and provide our first representation theorem in the riskless domain with finitely many alternatives. Working in this domain, we are able to identify the key axioms characterizing the limited willpower model with minimal structure. Our first axiom is standard.

Axiom 1. $\succsim$ *is complete and transitive.*

For simplicity throughout this section we assume that for all $x, y \in X$, if $x \neq y$, either $x \succ y$ or $y \succ x$ and use $\succ$ notation instead of $\succsim$. Since there are no indifferences, the choice must be unique, $|c(S)| = 1$ for all S. We will relax this assumption when we move to the lottery domain. The second axiom is Independence of (Unchoosable) Preferred Alternative (IPA); better options that are not chosen can be removed without affecting the actual choice.

Axiom 2. *(IPA) If* $x \succ y$ *and* $y \in c(A \cup x)$ *then* $c(A) = c(A \cup x)$.

This axiom can be viewed as a relaxation of WARP. Recall WARP allows any unchosen alternative to be dropped without affecting actual choices. In contrast, IPA allows only preferred unchosen alternatives to be dropped without affecting actual choices.

IPA is based on the intuitive notion that when a tempting alternative is also the most preferred available alternative, it should be chosen. Hence any unchosen alternative that is strictly preferred to the chosen one must have a relatively low temptation value. IPA says that dropping such alternatives should not affect the actual choice. Let's revisit the example in Section 2 with three alternatives, g, b and t with $g \succ b \succ t$. Suppose reading a book is uniquely chosen when all three options are available, i.e. $b = c(g, b, t)$. This means the most preferred alternative (g) is not chosen, and hence, is not the most tempting alternative and is irrelevant in the sense that dropping it from the menu should not affect the choice behavior of the DM. That is, we must have $c(g, b, t) = c(b, t)$. On the other hand, it is possible that removing t, the least preferred alternative, might influence the choice. If b is not as tempting as t, the DM can choose the best alternative g when t is removed, i.e. $g = c(g, b) \neq c(g, b, t)$. Hence, WARP is not satisfied in the presence of limited willpower.

The next axiom is Choice Betweenness: the choice from the union of two sets is "between" the choices made separately from each set with respect to preference.

Axiom 3. *(CB) If $c(A) \succsim c(B)$ then $c(A) \succsim c(A \cup B) \succsim c(B)$.*

To understand this axiom take the union of two choice sets $A \cup B$ and w.l.o.g. suppose A contains one of the chosen alternatives from $A \cup B$. Consider two (not necessarily mutually exclusive) cases. First, suppose A contains the most tempting item in $A \cup B$. In this case, the DM should not be able to choose a strictly better alternative from A (since she needs to overcome the same temptation from $A \cup B$ as from A) but should still be able to choose the alternative originally chosen from $A \cup B$, i.e., $c(A) \sim c(A \cup B)$.[12] Note that in this case the axiom is automatically satisfied since $c(A \cup B)$ must be in between $c(A)$ and $c(B)$ in terms of preference. As a second case suppose B contains contains the most tempting item in $A \cup B$. In this case the DM should be able to choose at least as preferred an alternative from A as she can from $A \cup B$ since she needs to overcome a weaker temptation from A. Moreover, the alternative chosen from B cannot be strictly preferred since the most tempting alternative is contained in B. Thus, the axiom should be satisfied in this case as well.[13]

A closely related axiom is Gul and Pesendorfer's Set Betweenness (SB). Although at first glance CB seems like a translation of SB to our domain, the two axioms are independent. To make this point precise, suppose $\succ_0$ is a preference relation over non-empty subsets of X. We say $\succ_0$ satisfies SB if $A \succ_0 B$ implies $A \succ_0 A \cup B \succ_0 B$. We let $x \succ y$ iff $\{x\} \succ_0 \{y\}$. We will now provide two examples that show that SB and CB are indeed independent axioms.

In the first example, $\succ_0$ satisfies SB, but $(\succ, c)$ violates CB. For this example, we use the *costly self-control representation* axiomatized by Noor and Takeoka (2010) in the menu preference framework.[14] We say that $\succ_0$ has a costly self-control representation if it can be represented by

[12]Implicit in these arguments is that only the most tempting alternatives matter in influencing the DM's choices. Clearly, this is also the case in the representation since only the alternative with the highest v value matters in determining which alternatives are choosable from a choice set.

[13]In fact, $c(A \cup B)$ can be strictly between $c(A)$ and $c(B)$. Continuing with our earlier example, let $A = \{g, b\}$ and $B = \{t\}$. Recall that both g and b are strictly better than t, so $c(A) \succ c(B)$. The choice from all three options, b, is strictly better than t, the worst alternative, so $c(A \cup B) \succ c(B)$. Moreover, from the set A, g is chosen, thus $c(A) \succ c(A \cup B) \succ c(B)$.

[14]Noor and Takeoka (2010)'s axiomatization is in the lottery domain. Here we adopt their representation to the

$V : \mathcal{X} \to \mathbb{R}$ given by

$$V(A) = \max_{x \in A} \quad u(x) - \varphi(max_{y \in A} v(y) - v(x))$$

where $u, v : X \to \mathbb{R}$ and $\varphi : \mathbb{R} \to \mathbb{R}$. The DM's choices, naturally implied by the model, are given by

$$c(A) = \operatorname*{argmax}_{x \in A} \quad u(x) - \varphi(max_{y \in A} v(y) - v(x)).$$

It is easy to see that if $\succ_0$ has a costly self-control representation then it satisfies SB. To see that $(\succ, c)$ can violate CB let $X = \{x, y, z\}$, $\varphi(a) = a^{.5}$, $u(x) = 2$, $u(y) = 1$, $u(z) = 0$, and $v(x) = 0$, $v(y) = 1.5$, $v(y) = 3$. In this case direct calculation shows that $x = c(x, z) = c(x, y, z) \succ y = c(x, y) \succ z = c(y, z)$. Hence, $(\succ, c)$ does not satisfy CB since $c(x, y, z) \succ c(x, y) \succ c(y, z)$.

In the second example, $(\succ, c)$ satisfies CB but $\succ_0$ violates SB. Suppose $\succ_0$ is represented by a function $W : \mathcal{X} \to \mathbb{R}$ defined as follows. If A has 2 or more elements:

$$W(A) = \max_{x \in A} u(x) - \left(\max_{y, z \in A, y \neq z} (v(y) + v(z)) - v(x) \right)$$

and for singleton sets $W(\{x\}) = u(x)$ where $u, v : X \to \mathbb{R}$. The above model is a variation of Gul and Pesendorfer (2001) where the self-control cost is linear but, differently from that model, here the DM is tempted by not just the most tempting but also the second most tempting alternative in the set. The DM's choices are given by

$$c(A) = \operatorname*{argmax}_{x \in A} u(x) + v(x).$$

It is easy to see that $(\succ, c)$ satisfies CB. To see that $\succ_0$ violates SB, let $X = \{x, y, z\}$, $u(x) = 7$, $u(y) = 3$, $u(z) = 2$, $v(x) = 0$, $v(y) = 1$ and $v(z) = 2$. Then, $\{x, y\} \succ_0 \{x, z\} \succ_0 \{x, y, z\}$.

Next, we present our first representation theorem in the riskless domain.

Theorem 1. $(\succ, c)$ *satisfies satisfies Axioms 1-3 if and only if it admits a generalized limited willpower representation.*

Obviously the generalized limited willpower representation contains limited willpower representation as a special case. Less obvious is that the generalized limited willpower representation is also closely related to the costly self-control representation that we discussed earlier. If the cost function φ is linear, this is the model of Gul and Pesendorfer, which satisfies WARP. More interestingly when the cost function is not linear, the model generates WARP violations. The previous literature focused especially on the cases where the cost function is either convex or concave. Theorem 1 sheds light on an important distinction between these cases. The convex cost function representa-

riskless environment.

tion satisfies our Axioms 1-3, hence it is a special case of the generalized limited willpower model. The concave cost function representation, on the other hand, is not a special case since, as shown earlier, it violates CB. This shows that although the generalized limited willpower representations is quite broad, it rules out some choice patterns. This discussion also highlights an advantage of our domain in that it allows transparent comparisons of various models directly through their choice implications.

Our next goal is to characterize the limited willpower model. To do this we need one more assumption. Consider four alternatives $x, y, z, t \in X$. Suppose, $y \succ c(y, z)$, that is the DM prefers y to z but is unable to choose it. Intuitively this means that z is more tempting than y. If, in addition, $c(t, z) = t$, then t must be more tempting than y as well, otherwise the DM would not be able to choose t. If $x \succ c(x, y)$, then the DM prefers x but cannot choose it against y because y is too tempting. Since t is even more tempting than y, the DM should be unable to choose x against t either. This intuitive conclusion would hold for the limited willpower model but it is not implied by IPA and CB. This is our next axiom, Consistency.

Axiom 4. *(Consistency) Let $y \succ c(y, z)$ and $c(t, z) = t$. If $x \succ c(x, y)$ then $c(x, t) = t$.*

Now, we are ready to prove the main representation theorem in the riskless domain.

Theorem 2. $(\succ, c)$ *satisfies Axioms 1-4 if and only if it admits a limited willpower representation.*

The costly self-control model with convex cost function violates the Consistency axiom.[15] Hence, Consistency provides a direct test to separate the limited willpower model from costly self-control models. Moreover, the test is based only on ex-ante preferences and ex-post choices and does not require information on menu preferences.

In the finite domain, limited willpower representation lacks uniqueness even in an ordinal sense. To see this suppose $X = \{x, y, z\}$ with $x \succ y \succ z$ and the DM maker chooses the most preferred alternative from any menu. This behavior is consistent with a limited willpower representation where there is no self-control problem (e.g. $v_1(x) > v_1(y) > v_1(z)$ and $w_1 = 0$). It is also consistent with a limited willpower representation where self-control problem exists but the DM has enough willpower to overcome it (e.g. $v_2(z) > v_2(y) > v_2(x)$ and $w > v_2(z) - v_2(x)$.) In the next section, we move to the lottery domain where the additional structure allows us to prove a representation theorem with stronger uniqueness properties.

[15]To see this consider the following example. Suppose $\varphi(a) = a^2$, $u(x) = 9$, $u(t) = 4.9$, $u(y) = .9$, $u(z) = 0$, and $v(x) = 0$, $v(t) = 2$, $v(y) = 3$, $v(z) = 4$. In this case direct calculation shows that $x \succ y = c(x, y) \succ z = c(y, z)$, $c(t, z) = t$ and $c(x, t) = x$, hence Consistency is violated.

4 Lottery Domain

4.1 Axioms

In the lottery domain, in addition to completeness and transitivity, we assume that $\succsim$ also satisfies the standard independence axiom. Hence, we replace Axiom 1 with the following one.

Axiom 5. *(EU) $\succsim$ admits an expected utility representation.*

In the lottery domain, we maintain the two main behavioral axioms, IPA and CB and allow indifferences. One of the implications of IPA is that the DM must be indifferent between all chosen lotteries with respect to preferences. In other words, $p, q \in c(A)$ implies $p \sim q$. Given this fact, we abuse the notation and write $p \succsim c(A)$ if $p \succsim q$ and $q \in c(A)$. Similarly, we use $c(A) \succsim c(B)$ if $p \succsim c(B)$ and $p \in c(A)$.

We impose the next axiom to allow for indifferences in preferences, which is necessary in this domain. This axiom relaxes the classical Independence of Irrelevant Alternatives (IIA) which requires that if a lottery is chosen from a larger set and it is in a smaller subset, then it must also be chosen from a smaller set. In the presence of willpower problems IIA may not necessarily hold because the larger set might have alternatives that are more tempting then the ones in the smaller one. Thus the DM might be able to choose better lotteries from the smaller set. However, if the lotteries chosen from the smaller set are not better than the ones from the larger set, any lottery that is chosen from the larger one must also be chosen from the smaller one. This is our next axiom:

Axiom 6. *If $c(A) \sim c(B)$ and $p \in A \subset B$ then $p \in c(B)$ implies $p \in c(A)$.*

To allow for multi-valued choice, we impose two additional axioms. The first one says that if a lottery is chosen from two choice sets, it will be also chosen from their union. If the DM has enough willpower to choose p from A and B, she can also choose p from $A \cup B$. Indeed, if the choice is single-valued, this axiom is implied by CB.

Axiom 7. *If $p \in c(A) \cap c(B)$ then $p \in c(A \cup B)$.*

The second one says that, if there are at least two chosen lotteries, removing one of them will lead to (weak) improvement. Assume both p and q are in $c(A)$. This means that the DM has enough willpower to choose p from A, hence from any subset of A including p. Removing q can only relax the willpower constraint, which might lead to the choice of a better lottery. Again, if the choice is single-valued, this axiom is trivially satisfied.

Axiom 8. *If $c(A) \setminus q \neq \emptyset$ and $q \in c(A)$ then $c(A \setminus q) \succsim c(A)$.*

The next two axioms impose independence conditions on the choice correspondence.[16] We use the notation $p\alpha q$ as a short hand for $\alpha p + (1-\alpha)q$ where α is a scalar. The standard (or full) independence axiom adapted to choice correspondences would say that if a lottery is chosen over another one, when both are mixed with a third alternative using the same mixture weight, the mixture of the better lottery is still chosen, i.e., $p \in c(p,q)$ implies $p\alpha r \in c(p\alpha r, q\alpha r)$ where $\alpha \in [0,1]$. The following example illustrates that full independence is too strong for the limited willpower model.

Example 1. *Assume $u(x) = 1$ and $u(y) = 0$, $v(x) = 0$ and $v(y) = 3$, and $w = 2$. Because $v(y) - v(x) = 3 > 2 = w$, the DM does not have enough willpower to choose x whenever y is available. Hence the DM ends up choosing y from $\{x,y\}$, i.e. $c(x,y) = y$. Now replace x with the half-half mixture $x\frac{1}{2}y$. Since v is linear, the temptation of the mixture is higher than the temptation value of x. Indeed, since $v(y) - v(x\frac{1}{2}y) = \frac{1}{2}v(y) - \frac{1}{2}v(x) = 1.5 < 2 = w$, the DM has enough willpower to overcome the temptation and chooses the mixture, $c(x\frac{1}{2}y, y) = x\frac{1}{2}y$. Hence full independence is violated.*

Full independence fails in the above example because when the tempting alternative y is present, choosing the half and half mixture of the preferred alternative x and the tempting alternative y requires less willpower than choosing x. Thus x is not choosable but $x\frac{1}{2}y$ is choosable when y is available.

Note that, in the example, x is better than y but it is not choosable over y. At the same time y is (trivially) better than and choosable over itself. Thus $x\frac{1}{2}y$ involves a mixture of two alternatives that are both better than y but one is choosable and the other not choosable over y. The example illustrates that in such cases independence might fail. Our next two axioms relax the full independence axiom to take into account choosability of alternatives.

Consider two binary choice problems: $c(p,q)$ and $c(p',q')$ where p and p' are better than q and q', respectively. Suppose the DM cannot choose the better options in either situation. We interpret this to mean that she does not have enough willpower to choose the better alternative in either case. Now consider a third choice problem $c(p\alpha p', q\alpha q')$. Part ii of the next axiom requires that the DM is unable to choose the mixture of the better alternatives in this new choice problem. In other words, if both p and p' are unchoosable over q and q', respectively, then $p\alpha p'$ is unchoosable over $q\alpha q'$. Part i of the axiom considers the opposite case when both p and p' are choosable over q and q' respectively. In this case the axiom requires that $p\alpha p'$ is choosable over $q\alpha q'$. While the full independence axiom might fail if we mix choosable and unchoosable alternatives, the model enjoys independence if we account for choosability as given by our next axiom, Temptation Independence.

[16] In our setup choices are not necessarily captured by a single preference relation, therefore we need to impose additional independence conditions directly on the choice correspondence.

Axiom 9. *(Temptation Independence) Suppose $p \succ q$. Then for all $\alpha \in (0,1]$,*

(i) If $p = c(p,q)$, $p' \in c(p',q')$ and $p' \succsim q'$ then $p\alpha p' = c(p\alpha p', q\alpha q')$,

(ii) If $q = c(p,q)$, $q' = c(p',q')$ and $p' \succ q'$ then $q\alpha q' = c(p\alpha p', q\alpha q')$.

The next axiom, Invariance, relaxes the full independence axiom in a different way. A version of full independence can be formulated as follows:

$$\text{If } c(p\alpha r, q\alpha r) = p\alpha r \text{ then } c(p\alpha' r', q\alpha' r') = p\alpha' r' \text{ for any } r' \text{ and } \alpha'.$$

When α' is not equal to α, full independence might fail in our model. To see that we revisit Example 1 and remember $c(x\frac{1}{2}y, y) = x\frac{1}{2}y$ and $c(x,y) = y$. We could write these choices as $c(x\frac{1}{2}y, y\frac{1}{2}y) = x\frac{1}{2}y$ and $c(x1y, y1y) = y1y$. Hence, different mixing ratios might affect choosability and the choice. The invariance axiom relaxes full independence by comparing only cases where $\alpha' = \alpha$. Specifically, consider two cases. The DM chooses either from $\{p\alpha r, q\alpha r\}$ or from $\{p\alpha r', q\alpha r'\}$. The axiom requires that as long as both lotteries p and q are are mixed with a third lottery using the same weight α, neither the ranking nor the choosability of the mixtures will be affected. Thus if the DM chooses $p\alpha r$ from $\{p\alpha r, q\alpha r\}$ then she must choose $p\alpha r'$ from $\{p\alpha r', q\alpha r'\}$.

Axiom 10. *(Invariance) If $c(p\alpha r, q\alpha r) = p\alpha r$ then $c(p\alpha r', q\alpha r') = p\alpha r'$ for any r'.*

We also make a standard continuity assumption for binary choice sets.

Axiom 11. *(Continuity) Suppose $p_n \to p$ and $q_n \to q$ with $p_n \succsim q_n$ for all n. If $p_n \in c(p_n, q_n)$ then $p \in c(p,q)$.*

As emphasized above, we are ultimately interested in conflict between preferences and choices. Axiom 10, Conflict, rules out cases where a conflict between preferences and choices never arises. In other words, Conflict requires that there are situations where the DM prefers one lottery over another but when she has to make a choice between the two she chooses the less preferred lottery.

Axiom 12. *(Conflict) There exist p and q such that $p \succ q$ and $c(p,q) = q$.*

Axiom 11, Limited Agreement, on the other hand, restricts how much conflict can occur between preferences and choices. It says that when the two lotteries are very close, the DM's choices must be in agreement with her preferences.

Axiom 13. *(Limited Agreement) For all $p \succ q$, there exists $\alpha > 0$ such that $p\alpha q = c(p\alpha q, q)$.*

We now state our main theorem in the lottery domain.

Theorem 3. $(\succsim, c)$ *satisfies IPA (Axiom 2), CB (Axiom 3) and Axioms 5-13 if and only if it admits a limited willpower representation.*[17]

It is routine to verify the if part of the theorem. For the only if part, we first identify u by the help of Axiom 5. We then define an auxiliary definition- just choosability. We say, given $p \succ q$, p is **just choosable** over q if (i) $p = c(p, p\alpha q)$ if $0 < \alpha < 1$ but $p \neq c(p, p\alpha q)$ if $\alpha < 0$ and (ii) $p\alpha q = c(p\alpha q, q)$ if $0 < \alpha < 1$ but $p\alpha q \neq c(p\alpha q, q)$ if $\alpha > 1$. We show that the just choosability relation is linear on the line passing through p and q. We then show that there exist two neighborhoods of p and q such that the linearity is still satisfied. We then utilize this finding to define v for all lotteries and show that the representation holds for any binary set. The last step in the proof extends the representation from binary sets to any finite set.

Since u represents $\succsim$, which admits an expected utility representation, u is unique up to any positive affine transformation. The next proposition states that the temptation ranking and the willpower stock are unique up to a *common* positive linear transformation and an additive shift to the former.

Theorem 4. *If* (u, v, w) *and* (u', v', w') *are limited willpower representations of* $(\succsim, c)$ *then there exist scalars* $\alpha > 0, \alpha' > 0, \beta, \beta'$ *such that* $u' = \alpha u + \beta$, $v' = \alpha' v + \beta'$ *and* $w' = \alpha' w$.

Conflict and Limited Agreement are crucial to provide the uniqueness result. When there is no observable conflict between the DM's preferences and choices, the temptation ranking v and the willpower stock w are not pinned down. To see this, suppose the DM's choice correspondence c can be represented by her preferences $\succsim$. We can then find a limited willpower representation of c by setting v equal to u and choosing any non-negative amount of willpower. Conflict and Limited Agreement eliminate these trivial cases.

5 Contracting with Consumers with Limited Willpower: An Example

In this section we illustrate through a simple example that the willpower model delivers distinct predictions in economic applications compared with other widely studied models of limited self-control.[18]

We consider a two-period model of monopolistic contracting between a monopolist and a consumer. Let's denote the set of alternatives available to the monopolist by A. In the first period,

[17] Remember both u and v are linear in the lottery domain.

[18] In this example we compare the predictions of the Strotz model with the limited willpower model. However, very similar comparisons would hold with respect to the costly self-control or hyperbolic discounting models.

the monopolist offers the consumer a contract that consists of a menu of alternatives $C \subseteq A$ with corresponding prices. The consumer can accept or reject the contract. If the consumer accepts the contract, in the second period she chooses an alternative from the menu and pays its price to the monopolist. If the consumer rejects the contract then she receives her outside option normalized to zero. We assume that both parties are committed to the contract once accepted.[19]

We denote the price of alternative (or service s) by p_s, the cost of providing it by $c(s)$, its utility to the consumer by $u(s)$, and its temptation value by $v(s)$. We call $e(s) = v(s) - u(s)$ as the excess temptation associated with alternative s. We assume that the consumer has limited willpower and utility and temptation values are both quasilinear in prices.[20] We denote $U(s, p_s) = u(s) - p_s$ and $V(s, p_s) = v(s) - p_s$.

The monopolist's profit from selling alternative s is $p_s - c(s)$.[21] Following Eliaz and Spiegler (2006) and Spiegler (2011) we assume that the consumer is naive in the sense that she believes she has no self-control problem, i.e., she believes that she will choose from the menu alternative s that maximizes $U(s, p_s)$.[22] In reality, the consumer's second period choices are governed by the limited willpower model, that is, she might be tempted by the other alternatives available in the contract C. This means that from the menu the consumer chooses alternative s that maximizes $U(s, p_s)$ subject to $V(s', p_{s'}) - V(s, p_s) \leq w$ where w is the willpower stock and s' is the alternative on the menu that maximizes V. We assume that the monopolist knows that the consumer has limited willpower and can predict perfectly the consumer's second period choices.[23]

In our example we will assume that the monopolist has only four potentially available alternatives: s_1, s_2, s_3, s_4. Here think of alternative s_1 as the basic service level, and the others as upgrades. These upgrades provide higher utility and temptation values but are also costlier to produce. For each alternative the utility and temptation values and the production cost are given in the following table:

[19] This framework fits into many real world situations. For example, when signing up for a phone plan, gym-membership, or a credit card, purchasing a holiday package, or making a hotel reservation consumers often sign a contract that specifies a basic level of consumption but can be "upgraded" at the time of consumption.

[20] Broadly speaking, the idea that temptation would decrease in price seems reasonable in many situations. When the price of a good increases, the consumer must forego other potentially tempting goods. Moreover, when the price is sufficiently high the good might become unaffordable. Quasilinearity of temptation values in prices is clearly a partial equilibrium way of capturing the impact of prices on temptation and a restrictive assumption. Yet it provides tractability and is implicitly invoked in the literature on changing tastes where it is usually assumed that both the present and future utilities are quasilinear in prices.

[21] We assume that the production cost is incurred only for the service that the consumer chooses from the menu.

[22] Other important contributions to this literature include DellaVigna and Malmendier (2004); Heidhues and Koszegi (2010).

[23] More precisely, we solve for the optimal contract for the monopolist given its beliefs about the consumer's behavior. To do this we do not need to know whether the monopolist (or the consumer) holds correct beliefs about the consumer's second period behavior.

	u	v	c
s_1	4	6	1
s_2	8	12	4
s_3	12	18	9
s_4	16	24	16

To begin suppose the consumer can commit to an alternative at the time of purchasing the contract.[24] Or equivalently, the monopolist can only offer a contract with one option or the consumer is a standard utility maximizer with no willpower limitation. In any of these cases, the monopolist would offer the consumer s_2 (which is the option that maximizes $u - c$), and charge the consumer $p(s_2) = u(s_2) = 8$, yielding a profit of 4 for the monopolist.

Next, consider the case where the consumer has zero willpower (i.e. has Strotz preferences). In this case, the optimal contract involves two alternatives. One of the alternatives is used as bait to attract the naive consumer, but the other one (which we call the indulging alternative) is what the consumer actually chooses in the second period. To maximize its profit the monopolist chooses s_3 – the alternative that maximizes $v - c$ – as the indulging alternative, and s_1 – the alternative with lowest overall excess temptation $v - u$ – as bait.[25] Since the price of the bait can be at most its u-value, the monopolist sets $p(s_1) = 4$. The consumer's willpower constraint implies that the price of the indulging alternative must be less than $v(s_3) - (v(s_1) - u(s_1)) = 16$. Hence with Strotz preferences, the optimal contract is $\{s_1, s_3\}$ with prices $p(s_1) = 4$ and $p(s_3) = 16 - \varepsilon$ where ε is a slight discount that makes the willpower constraint strict. The monopolist's profit is $p(s_3) - c(s_3) = 7 - \varepsilon$.

The case of $w = 0$ replicates Eliaz and Spiegler (2006)'s finding that the monopolist's optimal contract is an indulging contract for naive consumers with Strotz preferences. The novelty of our example arises when $w > 0$. It turns out that, as long as the consumer's willpower is not too high, the monopolist can improve its profit by offering a *compromising contract* that consists of three alternatives. We consider two cases $w = 2$ and $w = 4$. We refer to these as low and medium willpower cases.

To begin consider the low willpower case with $w = 2$. Suppose the monopolist continues to offer the alternatives $\{s_1, s_3\}$. Since the consumer can resist some temptation, it is easy to see that now the monopolist needs to lower the price of s_3 from 16 to $16 - w = 14$ reducing its profit to 5. Now, consider the contract offering three service levels $\{s_1, s_3, s_4\}$ with prices $p(s_1) = 4$, $p(s_3) = 16 - \varepsilon$ and $p(s_4) = 20 - \varepsilon$ where ε is arbitrarily small. Again, in period 1, the consumer believes that

[24] Of course, a naive consumer would not find such commitment necessary.

[25] The consumer is willing to pay at most $p_b = u(s_b)$ for the "bait" s_b. The consumer's willpower constraint implies that $v(s_3) - p_3 \geq v(s_b) - p_b = v(s_b) - u(s_b)$. Hence $p_3 \leq v(s_3) - (v(s_b) - u(s_b)) = e(s_b)$. This implies that the monopolist would choose s_b as the alternative with the lowest excess temptation e.

she will choose the bait s_1. In reality, since $v(s_1) - p(s_1) = 2 < v(s_4) - p(s_4) - w = 2 + \varepsilon$, the consumer does not have enough willpower to choose s_1 in period 2 when the tempting alternative s_4 is available. Moreover, she prefers to choose s_3 over s_4 since $u(s_3) - p(s_3) = -4 + \varepsilon > u(s_4) - p(s_4) = -5 + \varepsilon$ and, since $v(s_3) - p(s_3) = 2 + \varepsilon = v(s_4) - p(s_4) - w = 2 + \varepsilon$, she is able to do so. Note that the monopolist does not intend to sell the tempting alternative since it is too costly to produce, but by making it available the monopolist can make sure that the consumer is unable to choose s_1 and ends up choosing s_3 instead.

We highlight two key points about the low willpower case. First, note that the monopolist's profit is once again $7 - \varepsilon$ – the profit level that it makes when the traveler has no willpower – but the menu now must include a third tempting alternative. Second, in terms of utility value, temptation value and the price, the chosen alternative is in the middle. Thus, the consumer's choice exhibits the compromise effect, and importantly the effect is induced *endogenously* by the monopolist's choice of contract.

Next, consider the medium willpower case, $w = 4$. Clearly, faced with the previous contract the consumer would now be able to choose s_1 – an undesirable outcome for the monopolist. One solution would be to lower the prices of both s_3 and s_4 by 2. Since the reduction in the price of s_4 compensates for the change in the willpower stock, the consumer cannot choose s_1, and since prices of s_3 and s_4 are reduced by the same amount, then she would choose her preferred alternative s_3 over s_4. However this new contract would lower the monopolist's profit from $7 - \varepsilon$ to $5 - \varepsilon$.

In fact, the monopolist can do better by replacing s_3 with s_2 in the contract. To see how this contract works, let $p(s_1) = 4$, $p(s_2) = 10 - \varepsilon$ and $p(s_4) = 18 - \varepsilon$. Clearly, given the prices of s_1 and s_4, the consumer cannot choose s_1 in period 2. In addition, since $u(s_2) - p(s_2) = -2 + \varepsilon = u(s_4) - p(s_4)$, she is indifferent between s_2 and s_4 and since $v(s_2) - p(s_2) = 2 + \varepsilon = v(s_4) - p(s_4) - w$, she is able to choose s_2 over s_4. Thus under this contract the consumer would choose s_2 and the monopolist's profit is $6 - \varepsilon$.

The medium willpower case allows us to make further observations. First, as the willpower level increases, the monopolist switches from selling the indulging alternative s_3 to selling the commitment alternative s_2. Rather surprisingly, even though it sells the same product, its profits exceeds the profit level under commitment. This is because the consumer accepts the contract under the naive belief that it will consume the frugal option s_1 and does not believe that she would pay the high price of s_2. However, once she accepts the contract, she ends up consuming the rather more indulgent alternative s_2. Of course, as the consumer's willpower increases the monopolist would have to reduce the price of s_2 and eventually offer only s_2 at the commitment price.

It is also interesting that when consumers have limited willpower the optimal menu contains products that neither the consumer nor the monopolist believe would be consumed. The consumer naively views some of the alternatives as irrelevant, whereas the monopolist views them as tempting options that makes the frugal choices unpalatable to the consumer.

In summary we make the following observations from this example. When the consumer is a standard utility maximizer the monopolist needs to offer only one alternative to maximize its profit. When the consumer has Strotz preferences (i.e. zero willpower), the optimal contract consists of two alternatives. However, as soon as the consumer has positive willpower, the optimal contract consists of three alternatives and the consumer's choices reflect a form of the "compromise effect" which is induced endogenously through the monopolist's choice of contract. Moreover, both the design of the menu itself and the product that is actually consumed depend subtly on the consumer's willpower stock.

These results are in fact quite robust but a complete discussion of contracting with naive consumers with limited willpower is beyond the scope of this paper. We refer interested readers to our companion paper (Masatlioglu et al. (2014)).

6 Discussion

6.1 Defining "More Willpower" and Psychology Experiments

Psychologists (Baumeister and Vohs (2003); Faber and Vohs (2004); Muraven et al. (2006)) have run experiments aiming to demonstrate that individuals who perform prior acts of self-restraint tend to behave later as if they have less self-control. The typical experiment has two phases. Every subject participates in the second phase but only a randomly chosen subset participates in the first, with the remainder serving as a control group. In the first phase, subjects are asked to perform a task that requires self restraint; in the second phase, a single choice from a feasible set of alternatives is observed. The choice in the second phase requires self control. For example, the subject decides for how long to squeeze a hand grip where choosing to squeeze longer requires more self-control. Subjects who participate in the first phase seem to give into temptation in the second phase. Experimental psychologists view these experiments as an apparent demonstration of limited willpower.

We interpret these experiments in the context of our model as follows. We assume that the subject's preference relation does not depend on whether she is in the control or treatment group. However, being in the control or treatment group might affect the subject's choices. We denote subject's choices in the control and treatments groups as c_{cont} and c_{treat} respectively. In the experiment, while the subject in the treatment group gives in to temptation, the subject in the control group does not, in the sense that, $c_{cont}(x, y) = x \succ c_{treat}(x, y) = y$. It is apparent that these choices can be represented with a common (u, v) and different willpower stocks, $w_{cont} > w_{treat}$.

We would like to extend this intuition to cases where we observe multiple choices. Assume the subject in the treatment group gives in to temptation more often than the subject in the control group, in the sense that, $c_{cont}(A) \succsim c_{treat}(A)$ for all A and the relation is strict for some A. We first illustrate that, in our model, giving in to temptation when in the treatment group is not sufficient to conclude that the subject's willpower stock is depleted when we observe multiple choices. To see this consider the following example. The subject's preferences are $x \succ y \succ z$. The choices in the control group is $c_{cont}(x, y) = c_{cont}(x, z) = c_{cont}(x, y, z) = x$ and $c_{cont}(y, z) = z$. Notice that the subject gives in to temptation only when she faces $\{y, z\}$. The choices in the treatment group is $c_{treat}(x, z) = c_{treat}(y, z) = c_{treat}(x, y, z) = z$ and $c_{treat}(x, y) = y$. In the treatment group, she always gives in to temptation. Hence the subject in the treatment group gives into temptation more often than the subject in the control group. However, these choices cannot be represented with a common (u, v) and different willpower stocks. To see this suppose there was a common (u, v) and the willpower levels are such that $w_{cont} > w_{treat}$. Since $c_{cont}(x, z) = x$ and $c_{cont}(y, z) = z$, we should have $v(z) - v(y) > w_{cont}$ and $v(z) - v(x) < w_{cont}$ implying $v(y) < v(x)$. Since x is more tempting than y, independent of the willpower stock, x will be always chosen when the feasible set is $\{x, y\}$. This contradicts the fact that $c_{treat}(x, y) = y$.

The previous example illustrates that to identify when a subject in the treatment group has less willpower we need to make sure that the subject's temptation ranking v is the same whether she is in the control group or the treatment group. To see how we do this, suppose $p \succ q, q'$ and the control subject chooses p from $\{p, q\}$ and q' from $\{p, q'\}$. Suppose the treatment subject who has less willpower is unable to choose p in either case. Now consider the mixtures $p\beta q$ and $p\beta q'$. As β increases from zero, both of these mixtures become less tempting. If the control and the treatment subjects have the same v, then the former mixture is less tempting than the latter. Therefore, for small β the treatment subject chooses both mixtures over p, for large enough β she chooses p over both mixtures and for some intermediate range chooses p over $p\beta q$ and $p\beta q'$ over p, but never the reverse. The second part of the following definition formalizes this intuition.

Definition 1. *Let $\succsim$ be the subject's preference over Δ. Then, the subject in the control group has more willpower than the subject in the treatment group if and only if*

(i) $c_{cont}(A) \succsim c_{treat}(A)$ *for all* A*,*

(ii) Suppose $p \succ q, q'$*,* $c_{cont}(p, q) = p$ *and* $c_{cont}(p, q') = q'$*. If* $c_{treat}(p, p\beta q) = p\beta q$ *then* $c_{treat}(p, p\beta q') = p\beta q'$ *where* $\beta \in (0, 1]$.

If the subject's temptation ranking is the same whether she is in the control group or the treatment group and gives into temptation more often in the treatment group, she must have less willpower when she is in the treatment group compared to when she is in the control group. The next theorem shows that this is indeed the case.

Theorem 5. *Let the behaviors of a subject both in the control and treatment group be represented by a limited willpower model. Then the subject in the control group has more willpower than the subject in the treatment group (according to Definition 1) if and only if their behavior can be represented with a common u and v where $w_{cont} \geq w_{treat}$.*

Clearly, experiments that are designed to identify willpower as a cognitive resource need to control for variations in utilities of various alternatives. Our results says that controlling for preferences over alternatives is not sufficient and – even for within subject designs – the experimenter needs to control for variations in relative temptation values. Definition 1 tells us that at least in principle this can be done, but such experiments would require observation of more complex choices.

6.2 Relationship between Ex-ante Preferences and Ex-post Choices

Suppose an outside observer knows that the DM's ex-ante preferences and ex-post choices conform to the limited willpower model and wants to identify the utility function u, temptation ranking v and the willpower stock w. How much information does the observer need for this identification? In this section, we show that, in the lottery domain, the observer only needs the DM's ex-post choices and can construct her ex-ante preferences from ex-post choices (even though the choices violate WARP).

More formally, suppose $(\succsim, c)$ admits a limited willpower representation and suppose that we observe choices c, but do not know anything about $\succsim$ directly. That is, $(\succsim, c)$ is such that for some (u, v, w) with $w > 0$

$$c(A) = \underset{x \in A}{\operatorname{argmax}} \quad u(x) \quad \textit{subject to} \quad v(x) \geq \max_{y \in A} v(y) - w$$

where $u, v : \Delta \to R$ are linear and u represents $\succsim$. We wonder whether one can identify $\succsim$ (i.e., u) from c.

In the standard approach, preferences are revealed by choices in a straightforward way. We say x is revealed to be at least as good as y iff $x \in c(x, y)$. For choices that satisfy the limited willpower model this is no longer true. The difficulty in the identification is that even when $u(x) > u(y)$, y will be uniquely chosen from $\{x, y\}$ if $v(y) - v(x) > w$.

Despite this difficulty it is possible to uniquely identify $\succsim$ from choices. Suppose the DM prefers x over y. Either the DM has enough willpower to choose x or not. If it is the former, any mixture

$x\alpha y$ is also chosen over y. If it is the latter, the DM will have enough willpower to choose a mixture $x\alpha y$ for small enough α, and hence there must exist $x\alpha y$ that is uniquely chosen over y. Motivated by this, given choice correspondence c, we define a revealed preference relation $\succ^c$ as follows.

Definition 2. *We say $x \succ^c y$ if one of the following is true*

(i) $x = c(x, y)$ and there exists no $\alpha \in (0, 1)$ such that $y \in c(x\alpha y, y)$[26]*,*

(ii) $y = c(x, y)$ and there exists some $\alpha \in (0, 1)$ such that $x\alpha y = c(x\alpha y, y)$.

As usual we say $x \succsim^c y$ iff $y \not\succ^c x$. The next theorem shows that $\succsim^c$ is indeed the unique preference relation consistent with choices c if $(\succsim, c)$ has a limited willpower representation.

Theorem 6. *Suppose $(\succsim, c)$ satisfies Axioms 1-11. Then $\succsim$ and $\succsim^c$ are the same.*

In fact, using Theorem 6 we can conduct our analysis based only on c, even when we do not observe $\succsim$ directly.[27] To do this, we first construct the preference relation $\succsim^c$ from c. We can then check whether $(\succsim^c, c)$ satisfies the axioms in Theorem 3. If the answer is yes, we can derive the limited willpower representation of c. If the answer is no, then c does not have a willpower representation since $\succsim^c$ is the only possible candidate by Theorem 6.

In principle, checking whether $x \succ^c y$ for arbitrary $x, y \in \Delta$ can be difficult. If $x = c(x, y)$ we need to be sure there exists no $\alpha \in (0, 1)$ such that $y \in c(x\alpha y, y)$. If $y = c(x, y)$ we need to see if there exists some $\alpha \in (0, 1)$ such that $x\alpha y = c(x\alpha y, y)$. In general this requires checking whether in the former case $y \in c(x\alpha y, y)$ or in the latter case $x\alpha y \in c(x\alpha y, y)$ for arbitrarily small α. However, it turns out that it is sufficient to check these for a fixed $\widehat{\alpha}$ independent of x and y. To see this, let $y^v \in \arg\max_{y \in X} v(y)$ and $y^u \in \arg\min_{y \in X} v(x)$. For given $w > 0$, let $\bar{\alpha}(v(y^v) - v(y^u)) = w$. Clearly, if $u(x) \geq u(y)$ then $x\alpha y \in c(x\alpha y, y)$ for any $\alpha < \bar{\alpha}$.

7 Conclusion

Starting from Kreps (1979), researchers have been studying a two-period choice model, in which a DM picks a menu among several menus in the planning period (menu preferences) under the assumption that she is going to make a choice from each menu in the consumption period. This new and rich data set allows researchers to study phenomena like temptation and self-control. Menu

[26] This is equivalent to saying that for all $\alpha \in (0, 1)$ such that $x\alpha y = c(x\alpha y, y)$.

[27] A similar identification does not apply to the convex self-control model of Noor and Takeoka (2010). Consider the following (non-degenerate) example. Suppose $u(x) = 1$, $u(y) = 0$, $v(x) = 0$, $v(y) = 1$ and $\varphi(a) = a^2 + a$. In this case, the DM always chooses the lottery with more weight on y and it is not possible to distinguish between the preference and the temptation.

preferences are not only useful but also necessary to study self-control within Gul and Pesendorfer's model because consumption choices alone cannot reveal whether the DM has a self-control problem. However, the reliability of menu preferences depends on the ability of decision maker to predict her own future behavior (i.e. sophistication).

In this paper, to derive the limited willpower representation, we use a novel data set: ex-ante preferences and ex-post choices. Revealing the ex-ante preferences over *alternatives* is a simpler and more natural task than revealing ex-ante preferences over *all menus of alternatives*. More importantly, our data set allows us to remain agnostic about whether the DM is sophisticated or naive about anticipating her ex-post choices. To derive the representation, we introduce a new axiom called Choice Betweenness. We show that this axiom is independent of the Set Betweenness axiom that is commonly invoked in the menu preferences domain. In other words, the two axioms do not imply each other. We also show that in the lottery domain we can work with an even simpler data set since ex-ante preferences can be derived from ex-post choices. This is a surprising finding since self-control can be revealed from ex-post choice data alone (as opposed to menu preferences).

Finally we would like to highlight an important avenue for exploration in future work which is the implications of limited willpower in a dynamic setting with multiple tasks. In the current manuscript, we consider a model where willpower is needed in a single choice task. In fact, people often use willpower in multiple tasks, and using more willpower in one task might mean less willpower is left for another. Moreover, the model is static. In reality, there are dynamic effects in the sense that the amount of willpower used in one period can affect the willpower stock in the next period. For example, people reward themselves with a drink after a difficult day or week at work. and the willpower stock in one period does not carry over to the next period. Incorporating these considerations in an axiomatic framework can lead to new insights about behavior and a rich set of testable implications.

References

Ahn, David S and Todd Sarver, "Preference for flexibility and random choice," *Econometrica*, 2013, *81* (1), 341–361.

Baumeister, R. and K. Vohs, "Willpower, Choice, and Self-control," in "Time and decision: Economic and psychological perspectives on intertemporal choice," Russell Sage Foundation Publications, 2003, chapter 6, pp. 201–216.

Baumeister, R.F., T.F. Heatherton, and D.M. Tice, *Losing control: How and why people fail at self-regulation.*, Academic Press, 1994.

Bernheim, B.D. and A. Rangel, "Addiction and cue-triggered decision processes," *American Economic Review*, 2004, pp. 1558–1590.

Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini, "Standard State-Space Models Preclude Unawareness," *Econometrica*, January 1998, *66* (1), 159–174.

DellaVigna, S. and U. Malmendier, "Contract design and self-control: Theory and evidence," *The Quarterly Journal of Economics*, 2004, *119* (2), 353–402.

Dillenberger, David and Philipp Sadowski, "Ashamed to be selfish," *Theoretical Economics*, 2012, *7* (1), 99–124.

Eliaz, K. and R. Spiegler, "Contracting with diversely naive agents," *The Review of Economic Studies*, 2006, *73* (3), 689–714.

Ellis, Andrew, "Foundations for optimal inattention," 2013. Mimeo.

Faber, R. and K. Vohs, "To buy or not to buy?: Self-control and self-regulatory failure in purchase behavior," in "Handbook of self-regulation: Research, theory, and applications.," Guilford Press, 2004, pp. 509–54.

Fudenberg, D. and D.K. Levine, "A dual-self model of impulse control," *The American Economic Review*, 2006, pp. 1449–1476.

_ **and D.K Levine**, "Timing and Self-Control," *Econometrica*, 2012, *80*, 1–42.

Gul, F. and W. Pesendorfer, "Temptation and Self-Control," *Econometrica*, 2001, *69* (6), 1403–1435.

_ **and** _ , "Self-Control and the Theory of Consumption," *Econometrica*, 2004, *72* (1), 119–158.

Heidhues, P. and B. Koszegi, "Exploiting naivete about self-control in the credit market," *The American Economic Review*, 2010, *100* (5), 2279–2303.

Kreps, D. M, "A Representation Theorem for "Preference for Flexibility"," *Econometrica*, May 1979, *47* (3), 565–77.

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

Lipman, Barton and Wolfgang Pesendorfer, "Temptation," 2011. Mimeo.

Loewenstein, G., "Willpower: A Decision-theorist's Perspective," *Law and Philosophy*, 2000, *19*, 51–76.

Loewenstein, George and David Schkade, "Wouldnt it be nice? Predicting future feelings," *Well-being: The foundations of hedonic psychology*, 1999, pp. 85–105.

Masatlioglu, Y., D. Nakajima, and E. Ozdenoren, "Exploiting Naive Consumers with Limited Willpower," 2014. Working paper, University of Michigan.

Muraven, M., D. Shmueli, and E. Burkley, "Conserving self-control strength," *Journal of personality and social psychology*, 2006, *91* (3), 524.

Noor, J. and N. Takeoka, "Uphill self-control," *Theoretical Economics*, 2010, *5* (2), 127–158.

O'Donoghue, T. and M. Rabin, "Doing it now or later," *American Economic Review*, 1999, pp. 103–124.

Ozdenoren, E., S. W. Salant, and D. Silverman, "Willpower and the Optimal Control of Visceral Urges," *Journal of the European Economic Association*, 2012, *10* (2), 342–368.

Shefrin, H. M. and R. H. Thaler, "The Behavioral Life-cycle Hypothesis," *Economic Inquiry*, 1988, *26* (4), 609–643.

Spiegler, R., *Bounded rationality and industrial organization*, Oxford University Press, USA, 2011.

A Proofs

Proofs of Theorem 1 and 2

Before we provide the proofs of Theorem 1 and 2, we provide a brief sketch. To prove Theorem 1, we first define a binary relation $\rhd'$. We say that $x \rhd' y$ if $y \succ x = c(xy)$. In words, x blocks y if x is worse than y but DM cannot choose y when x is available. Next, we define a second binary relation $\rhd''$. We say that $x \rhd'' y$ if $x \succsim y$ and there exist a and b such that $a \rhd' y$, $x \rhd' b$, and $a \not\rhd' b$. We say that $x \rhd y$ if $x \rhd' y$ or $x \rhd'' y$. Next we show that $\rhd$ is an interval order, i.e. it is irreflexive and $x \rhd b$ or $a \rhd y$ holds whenever $x \rhd y$ and $a \rhd b$. The binary relation $\rhd$ is an interval order if and only if there exist functions v and w such that

$$\Gamma_{\rhd}(S) = \{x \in S : \max v(y) - v(x) \leq w(x)\}.$$

Finally, to complete the proof of the first step, we show that S is indifferent to the $\succsim$-best element in $\Gamma_{\rhd}(S)$.

In the proof of Theorem 2 we use consistency to show that we can construct a semi order $\overline{\rhd}$ (i.e., $\overline{\rhd}$ is an interval order and if $x \overline{\rhd} y \overline{\rhd} z$ then $x \overline{\rhd} t$ or $t \overline{\rhd} z$ for any t) by properly modifying $\rhd$ such that S is indifferent to the $\succsim$-best element in $\Gamma_{\overline{\rhd}}(S)$. To complete the proof we note that the binary relation $\overline{\rhd}$ is a semi order if and only if there exist a function v and a scalar w such that

$$\Gamma_{\rhd}(S) = \{x \in S : \max v(y) - v(x) \leq w\}.$$

Proof of Theorem 1

We first show that Axiom 1-3 imply an important implication of our model.

Claim 1. *Suppose $(\succ, c)$ satisfies Axiom 1-3. Then, If $x \succ c(A \cup x)$ then $c(B) = c(B \cup x)$ for all $B \supset A$.*

Proof. Let L(n) stand for the statement of Claim 1 that is restricted to when $|B - A| \leq n$. Notice that Axiom 2 is L(0). First, we shall show L(1). That is, $x \succ c(A \cup x)$ (so $c(A) = c(A \cup x)$ by Axiom 2) implies $c(A \cup y) = c(A \cup x \cup y)$ for any y.

Case 1: $y \succ c(A \cup x \cup y)$: By Axiom 2, $c(A \cup x) = c(A \cup x \cup y)$. By the assumption, we have $x \succ c(A \cup x) = c(A \cup x \cup y)$. By applying Axiom 2, we get $c(A \cup y) = c(A \cup x \cup y)$.

Case 2: $y \prec c(A \cup x \cup y)$: By Axiom 3, $y \prec c(A \cup x \cup y) \precsim c(A \cup x) \prec x$. By Axiom 1 we get $c(A \cup x \cup y) \prec x$. Then by Axiom 2, we get the desired result, $c(A \cup y) = c(A \cup x \cup y)$.

Case 3: $y \sim c(A \cup x \cup y)$: We have three sub-cases:

- If $y = c(A \cup y)$, then $c(A \cup y) = c(A \cup x \cup y)$.
- If $y \succ c(A \cup y)$, then Axiom 2 implies $c(A \cup y) = c(A)(= c(A \cup x))$. Applying Axiom 3, we get $c(A \cup x \cup y) = c(A \cup y)$, which is a contradiction because $c(A \cup x \cup y) = y \succ c(A \cup y)$.
- If $y \prec c(A \cup y)$, then Axiom 3 implies $(c(A \cup x) =)c(A) \succsim c(A \cup y)$. Applying Axiom 3 again, it must be $c(A \cup x \cup y) \succsim c(A \cup y) \succ y$, which is a contradiction because $c(A \cup x \cup y) = y$.

Now suppose that L(k) is true up when $1 \leq k \leq n-1$. We shall prove L(n). Assume $x \succ c(A\cup x)$ and let $B = A \cup \{y_1, y_2, \ldots, y_n\}$ where all of y_i's are distinct and excluded from A. Our goal is to show $c(B) = c(B \cup x)$. Without loss of generality, assume $y_1 \succ y_2 \succ \cdots \succ y_n$.

Case 1: $y \succ c(A\cup x\cup y)$ for some $y \in \{y_1, y_2, \ldots, y_n\}$: Since $(B\setminus y)\cup x \supset A\cup x$ and the difference of their cardinality is $n-1$, we can utilize L($n-1$). Then we get $c((B\setminus y)\cup x) = c((B\setminus y)\cup x\cup y)(= c(B\cup x))$. Applying L(1) to $x \succ A\cup x$, we have $(y \succ)c(A\cup x\cup y) = c(A\cup y)$. Applying L($n-1$) to this yields $c(B\setminus y) = c((B\setminus y)\cup y) = c(B)$. Notice that $c(B\setminus y) = c((B\setminus y)\cup x)$ because $x \succ c(A\cup x)$ and L($n-1$). These three equalities imply $c(B) = c(B\cup x)$.

Case 2: $y \prec c(A\cup x\cup y)$ for some $y \in \{y_1, y_2, \ldots, y_n\}$: By Axiom 3 we have $c(A\cup x) \succsim c(A\cup x\cup y) \succ y$. Since $x \succ c(A\cup x)$ and Axiom 1, we have $x \succ c(A\cup x\cup y)$. Because $|B\setminus (A\cup y)| = n-1$, by applying L($n-1$) we have $c(B) = c(B\cup x)$.

Case 3: $y_i = c(A\cup y_i\cup x)$ for all $i = 1, \ldots, n$: In this case, we have

$$y_1 = c(A\cup y_1\cup x) \succ y_2 = c(A\cup y_2\cup x) \succ \cdots \succ y_n = c(A\cup y_n\cup x)$$

Since $c(A\cup y_i\cup x) = c(A\cup y_i)$ by L(1), the above relations still hold when x is removed:

$$y_1 = c(A\cup y_1) \succ y_2 = c(A\cup y_2) \succ \cdots \succ y_n = c(A\cup y_n)$$

Recursively applying Axiom 3 implies

$$(c(A\cup y_1\cup x) =)y_1 \succsim c(A\cup \{y_1, y_2, \ldots, y_n\})(= c(B)) \succsim y_n(= c(A\cup y_n\cup x))$$

In other words,

$$c(A\cup y_1\cup x) \succsim c(B) \succsim c(A\cup y_n\cup x)$$

Since $(A\cup y_1\cup x)\cup B = B\cup x$, Axiom 3 implies $c(B\cup x) \succsim c(B)$. Similarly, since $(A\cup y_n\cup x)\cup B = B\cup x$, Axiom 3 implies $c(B) \succsim c(B\cup x)$. Therefore, by Axiom 1, $c(B) = c(B\cup x)$. □

For any binary relation R, let $\Gamma_R(S)$ be the set of R-undominated elements in S, that is,

$$\Gamma_R(S) = \{x \in S: \text{ there exists no } y \in S \text{ such that } yRx\}$$

Instead of constructing v and w, we shall construct a binary relation over X, denoted by $\rhd$ such that $c(S)$ is the $\succ$-best element in $\Gamma_\rhd(S)$.[28] It is known (Fishburn (1979)) that, if (and only if) $\rhd$ is an interval order[29], there exist functions v and ε such that

$$\Gamma_\rhd(S) = \{x \in S: v(y) - v(x) \leq w(x) \forall y \in S\} = \{x \in S: \max_{y\in S} v(y) - v(x) \leq w(x)\}$$

so that we can get the desired representation.

Now, for any $x \neq y$, we define $x \rhd y$ when either $x \rhd' y$ or $x \rhd'' y$ where $\rhd'$ and $\rhd''$ are defined as follow:

1. $x \rhd' y$ if $y \succ x = c(xy)$

[28] In our framework, the $\succ$-best element is equal to the $\succsim$-best element.

[29] $\rhd$ is called an interval order if it is irreflexive and $x \rhd b$ or $a \rhd y$ holds whenever $x \rhd y$ and $a \rhd b$.

2. $x \rhd'' y$ if $x \succ y$ and there exist a and b such that $a \rhd' y$, $x \rhd' b$, and $a \not\rhd' b$.

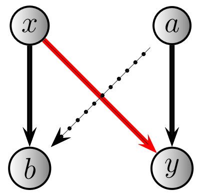


Figure 1: Black and Red arrows represent $\rhd'$ and $\rhd''$, respectively. Solid and dashed arrows indicate the existence and non-existence of relations, respectively.

Note that $x \rhd' y$ and $x \rhd'' y$ cannot happen at the same time. In addition, $\rhd'$ and $\rhd''$ are both irreflexive.

We need to show that (i) $\rhd$ is an interval order and (ii) the $\succ$-best element in $\Gamma_{\rhd}(S)$ is equal to $c(S)$.

Claim 2. *$\rhd'$ is asymmetric and transitive.*

Proof. By construction, $x \rhd' y$ and $y \rhd' x$ cannot happen at the same time. Suppose $x \rhd' y$ and $y \rhd' z$, i.e., $z \succ c(yz) = y \succ c(xy) = x$. Then by Claim 1, $c(xyz) = c(xz)$ because $y \succ c(xy)$. By Axiom 3, $(z \succ) c(yz) \succsim c(xyz) \succsim c(xy)$. Hence, we have $z \succ c(xyz) = c(xz)$. Hence we have $z \succ x = c(xz)$, so $x \rhd' z$. □

Claim 3. *If $x \rhd' y$ and $a \rhd' b$ but neither $x \rhd' b$ or $a \rhd' y$, then it must be $x \rhd'' b$ or $a \rhd'' y$ but not both.*

Proof. First we shall show that $x \rhd'' b$ and $a \rhd'' y$ cannot happen at the same time. Suppose it does. Then by definition of $\rhd'$ and $\rhd''$, we have $y \succ x \succ b \succ a \succ y$. Axiom 1 is violated.

Now, we shall show that either $x \rhd'' b$ or $a \rhd'' y$ must be defined. Suppose not. Then, along with the definition of $\rhd'$, we have $b \succ x = c(xy)$, and $y \succ a = c(ab)$. Therefore, $c(xyab)$ must be weakly worse than x or a because it must be weakly worse than $c(xy)$ or $c(ab)$ by Axiom 3.

Since neither (x, b) nor (a, y) belongs to $\rhd'$ or $\rhd''$, we have $c(xb) = b \succ x$, and $c(ay) = y \succ a$. By Axiom 3, $c(xyab)$ must be weakly better than $c(xb)$ or $c(ay)$ so it must be weakly better than y or b.

Hence, either x or a must be weakly better than either y or b. Since we have already seen $b \succ x$ and $y \succ a$, the only possibilities are $a \succsim b$ or $x \succsim y$, neither of which is possible because $a \rhd' b$ and $x \rhd' y$. □

Claim 4. *$\rhd$ is an interval order.*

Proof. We need to show that $\rhd$ is irreflexive. By definition, we cannot have (i) $x \rhd' y$ and $y \rhd' x$, (ii) $x \rhd' y$ and $y \rhd'' x$, or (iii) $x \rhd'' y$ and $y \rhd'' x$. Hence $\rhd$ is irreflexive.

Next we show that $x \rhd b$ or $a \rhd y$ holds whenever $x \rhd y$ and $a \rhd b$. We shall prove this case by case:

Case 1: $x \rhd' y$ and $a \rhd' b$: If we have $x \rhd' b$ or $a \rhd' y$, then we are done. Assume not, then Claim 3 implies we must have $x \rhd'' b$ or $a \rhd'' y$ (not both). Then $x \rhd b$ or $a \rhd y$.

Case 2: $x \rhd' y$ and $a \rhd'' b$: In this case, by definition of $\rhd''$ and Claim 3, there exist s and t such that $a \rhd' t$ and $s \rhd' b$ but not $s \rhd t$. Focus on $x \rhd' y$ and $a \rhd' t$, we must have either $a \rhd y$ (it is done in this case) or $x \rhd t$ (so eihter $x \rhd' t$ or $x \rhd'' t$). If $x \rhd' t$, then by looking at $x \rhd' t$ and $s \rhd' b$ Claim requires $x \rhd b$ because it is not $s \rhd t$. Thus, we consider the final sub-case: $x \rhd'' t$. If so, we have $x \rhd' y$ and $s \rhd' b$ so it must be either $x \rhd b$ (then done) or $s \rhd y$. If $s \rhd y$, then it must be $s \rhd' y$ (i.e. not $s \rhd'' y$) because $y \succ x \succsim t \succ a \succsim b \succ s$. Therefore, we have $s \rhd' y$ and $a \rhd' t$ with not $s \rhd t$. Hence it must be $a \rhd y$.

(a) Case 2 (b) Case 3

Figure 2: The Proof of Claim 4

Case 3: $x \rhd'' y$ and $a \rhd'' b$: By definition of $\rhd''$, there exist s and t such that $x \rhd' t$ and $s \rhd' y$ with not $s \rhd t$. Then by focusing on $x \rhd' t$ and $a \rhd'' b$, we must have either $x \rhd b$ (done) or $a \rhd t$. Suppose the latter. Then we have $s \rhd' y$ and "$a \rhd' t$ or $a \rhd'' t$," so the previous two cases are applicable so we conclude $a \rhd y$ because it is not $s \rhd t$.

□

Claim 5. *$c(S)$ is equal to the $\succ$-best element in $\Gamma_{\rhd'}(S)$.*

Proof. First, we prove that $\Gamma_{\rhd'}(S)$ does not include any element that is strictly better than $c(S)$. Suppose $x \in \Gamma_{\rhd'}(S)$. Let S' and S'' be the subsets of $S \setminus x$ consisting of elements that are better than x and strictly worse than x, respectively. That is,

$$S' := \{y \in S : \ y \succ x\} \text{ and } S'' := \{y \in S : \ x \succ y\}.$$

Then, we have $c(S' \cup x) \succ x$ by definition of c and $x = c(xy) \succ y$ for all $y \in S''$ by the definition of $\rhd'$. Then by applying Axiom 3 we get $c(S'' \cup x) = x$. Thus, $c(S' \cup x) \succ c(S'' \cup x)$ implies $c((S' \cup x) \cup (S'' \cup x)) = c(S) \succsim x$ again by Axiom 3.

Next, we shall show that $c(S) \in \Gamma_{\rhd'}(S)$. Suppose not. Then, there exists $y \in \Gamma_{\rhd'}(S)$ such that $y \rhd' c(S)$ by Claim 2 (especially $\rhd'$ is transitive). That is $c(S) \succ c(\{c(S), y\}) = y$. Thus, by Claim 1, we have $c(S \setminus c(S)) = c((S \setminus c(S)) \cup c(S)) = c(S)$, a contradiction.

Combining the first and second results, the $\succ$-best element in $\Gamma_{\rhd'}(S)$ is equal to $c(S)$. □

Claim 6. *$c(S)$ is equal to the $\succ$-best element in $\Gamma_{\rhd}(S)$.*

Proof. Since $\rhd \supseteq \rhd'$ by construction, we have $\Gamma_{\rhd}(S) \subseteq \Gamma_{\rhd'}(S)$. Therefore, by Claim 5, it is enough to show is that the $\succ$-best elements in $\Gamma_{\rhd'}(S)$ (which is $c(S)$) is included in $\Gamma_{\rhd}(S)$. Suppose $c(S) \notin \Gamma_{\rhd}(S)$. Since $\rhd$ is an interval order, it is automatically transitive. Therefore, there exists $y \in \Gamma_{\rhd}(S)$ such that $y \rhd c(S)$ but not $y \rhd' c(S)$. Therefore, it must be $y \rhd'' x$ so $y \succ c(S)$. Since $y \in \Gamma_{\rhd'}(S)$, y cannot be strictly better than $c(S)$ (see the proof of Claim 5). □

(The Representation $\Rightarrow$ The Axioms)

Showing that the first axiom is necessary is straightforward. For the second axiom, if $x \succ c(A \cup x)$ then A must have an element y with $v(y) > v(x) + w(x)$, so its superset B also includes y so $\Gamma(A) = \Gamma(B)$, so $c(A) = c(B)$.

The third axiom: Let x^* be the u-best element in $\Gamma(A \cup B)$. Then it must be in $\Gamma(A)$ or $\Gamma(B)$ as well so it is not possible that $A \cup B$ is strictly preferred to both A and B. Now we show that the union cannot be strictly worse than both. Let x_A and x_B be the u-best elements in A and B, respectively, and take v_A and v_B be the maximum values of v in A and in B, respectively. Then we have

$$v_A \leq u(x_A) + \varepsilon(x_A) \text{ and } v_B \leq u(x_B) + \varepsilon(x_B)$$

Therefore the maximum value of v in $A \cup B$ is the higher one between v_A and v_B, either x_A or x_B must be in $\Gamma(A \cup B)$ so $c(A \cup B)$ must be weakly better than either $c(A)$ or $c(B)$.

Proof of Theorem 2

We are now done proving the sufficiency of the axioms for the representation in Theorem 1. Next, we show the sufficiency of Axioms 1-4 for the representation in Theorem 2.

Claim 7. *If $x \succ c(xy) \succ c(yz)$ then, for all t, $c(xyzt)$ is either $c(xt)$ or $c(yz)$.*

Proof. Assume $x \succ c(xy) \succ c(yz)$, then it must be $x \succ y \succ z$. Consider $c(zt)$. If $c(zt) = t$ then by Axiom 4 we get $c(xt) = t$. Since $y \succ c(yz)$, by Claim 1, we have $c(zt) = c(yzt)$. By Axiom 3 we have $c(xt) = c(xyzt) = c(yzt)$. Hence $c(xyzt) = c(xt)$.

Now assume $c(zt) = z$. Since $x \succ c(xy)$, by Claim 1, we have $(z =)c(yz) = c(xyz)$. By Axiom 3, we have $c(zt) = c(xyzt) = c(xyz)(= c(yz))$. Hence $c(xyzt) = c(yz)$. □

Again as in the proof of Theorem 1, instead of defining $v(.)$ and $w > 0$, we shall construct a binary relation over X, denoted by $\overline{\triangleright}$ such that $c(S)$ is equal to the $\succ$-best element in $\Gamma_{\overline{\triangleright}}(S)$ (i.e. the set of $\overline{\triangleright}$-undominated elements in S). It is known (Fishburn (1979)) that if (and only if) $\overline{\triangleright}$ is a semi order[30], which is a special type of an interval order, there exist function v and positive number w such that

$$\Gamma_{\overline{\triangleright}}(S) = \{x \in S : \max_{y \in S} v(y) - v(x) \leq w\}$$

so we get the desired representation.

Next we define the (i, j)-representation for an arbitrary interval order P.

Claim 8. *Any interval order, P, has an (i, j)-representation (see Figure 3) if there exist two functions $i : X \to N$ and $j : X \to N$ such that*

i) For all $x \in X$, $i(x) \geq j(x)$,

ii) The ranges of i and j have no gap: That is if there exist x and y such that $i(x) > i(y)$ then for any integer n between $i(x)$ and $i(y)$ there is z with $i(z) = n$. Similarly for $j(\cdot)$,

[30] $\overline{\triangleright}$ is a semi order if it is an interval order and if $x\overline{\triangleright}y\overline{\triangleright}z$ then $x\overline{\triangleright}t$ or $t\overline{\triangleright}z$ for any t.

iii) xPy if and only if $i(x) < j(y)$.

Proof. The following proof is based on Mirkin (1979). Given an interval order, P, (xPy and zPw imply xPw or zPy) we can show that, for all x and y in X, $L(x) \subseteq L(y)$ or $L(y) \subseteq L(x)$, and, $U(x) \subseteq U(y)$ or $U(y) \subseteq U(x)$, where $L(x)$ and $U(x)$ are lower and upper contour sets of x with respect to P, respectively. That is, $L(x) = \{y \in X | \ xPy\}$ and $U(x) = \{y \in X | \ yPx\}$. Irreflexivity indicates that there is a chain with respect to lower contour sets (this is also true for upper contour sets), i.e., relabel elements of X, $|X| = n$ such that $L(x_j) \subseteq L(x_i)$ for all $1 \leq i \leq j \leq n$. Moreover, we can strict inclusions such as there exists $s \leq n$ such that $\emptyset = L(x_s) \subset L(x_{s-1}) \dots L(x_2) \subset L(x_1)$ where $\{x_1, x_2, \dots, x_s\} \subseteq X$. For all $k \leq s$, Define

$$I_k = \{x \in X \mid L(x_k) = L(x)\}$$

I_k is not empty for any k since $x_k \in I_k$ by construction. Clearly, the system $\{I_k\}_1^s$ is a partition of the set X, i.e. $\overset{s}{\underset{k=1}{\cup}} I_k = X$, $I_k \cap I_l = \emptyset$ when $k \neq l$. Define

$$i(x) := k \text{ if } L(x) = L(x_k) \text{ for some } x_k \text{ in } X.$$

Now construct another family of non-empty sets $\{J_m\}_1^s$, as follows

$$J_s = L(x_{s-1}) \setminus L(x_s), \ \cdots, \ \ J_2 = L(x_1) \setminus L(x_2), \ J_1 = X \setminus L(x_1)$$

Clearly, the system $\{J_m\}_1^s$ is another partition of the set X. Most importantly, we have $\emptyset = U(y_1) \subset U(y_2) \dots U(y_{s-1}) \subset U(y_s)$ where $y_i \in J_i$ for all $i \leq s$. Define

$$j(x) := k \text{ if } x \in J_k.$$

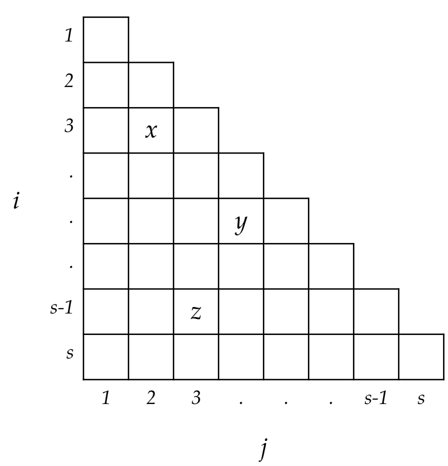


Figure 3: The graph of the (i, j)-representation. Condition ii) implies every row and column (not every cell) includes at least one alternative. Condition iii) implies $(x, y) \in P$ but $(x, z) \notin P$.

To see Condition i) holds, let $i(x) = i$. That means $x \in I_i$. If there exists no element z such that zPx, i.e. $U(x) = \emptyset$, then $j(x) = 1 \leq i(x)$. Otherwise find the largest integer j such that $x \in L(x_j)$. Note that j must be strictly less than i. Then by definition, $j(x) = j+1$, which is less than $i = i(x)$.

Since both $\{I_k\}_1^s$ and $\{J_k\}_1^s$ are partitions of X, there is no gap (Condition ii)). Finally, we have Condition iii) since $xPy \Leftrightarrow y \in L(x) \Leftrightarrow j(y) \geq i(x)+1 > i(x)$. □

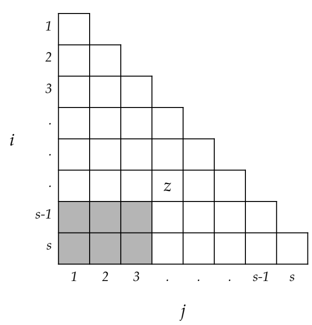
Let $\rhd$ be the interval order that is defined in the proof of Theorem 1. By Claim 8, it has an (i,j)-representation. We now modify the (i,j)-representation of $\rhd$ so that the resulting binary relation is a semiorder, say $\overline{\rhd}$, such that $c(S)$ is equal to the $\succ$-best element in $\Gamma_{\rhd}(S)$. In other words, we construct a semiorder based on the interval order we created without affecting the representation. To do this, we prove several claims relating the (i,j)-representation with the preference $\succ$.

Claim 9. *If $i(x) = j(y) - 1$, it must be $y \succ x$.*

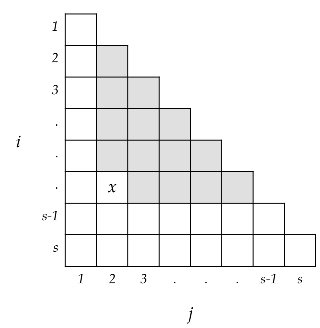
Proof. Since $i(x) < j(y)$ we know that $x \rhd y$. If $x \rhd' y$ then we are done since in that case $y \succ x = c(xy)$. So suppose that $x \rhd'' y$. Then by definition of $\rhd''$, there exist α and β such that $\alpha \rhd' y$ and $x \rhd' \beta$ and $\alpha \not\rhd' \beta$. Moreover, by Claim 3 $\alpha \not\rhd'' \beta$. So $\alpha \not\rhd \beta$. Since $\alpha \rhd y$ and $x \rhd \beta$, $i(\alpha) < j(y)$ and $i(x) < j(\beta)$. Since $\alpha \not\rhd \beta$, $i(\alpha) \geq j(\beta)$. Therefore it must be $i(x) \leq j(y) - 2$, a contradiction. □

Definition 3. *(i,j) is called a prohibited cell if there exists z such that $i(z) < i$ and $j(z) > j$. Otherwise, it is called a safe cell (see Figure 4a).*

To obtain a semi-order representation, we need to move each alternative that is in a prohibited cell to a safe cell and still the representation holds. The next definition describes a way in which alternatives can be moved.



(a) Prohibited cells because of z



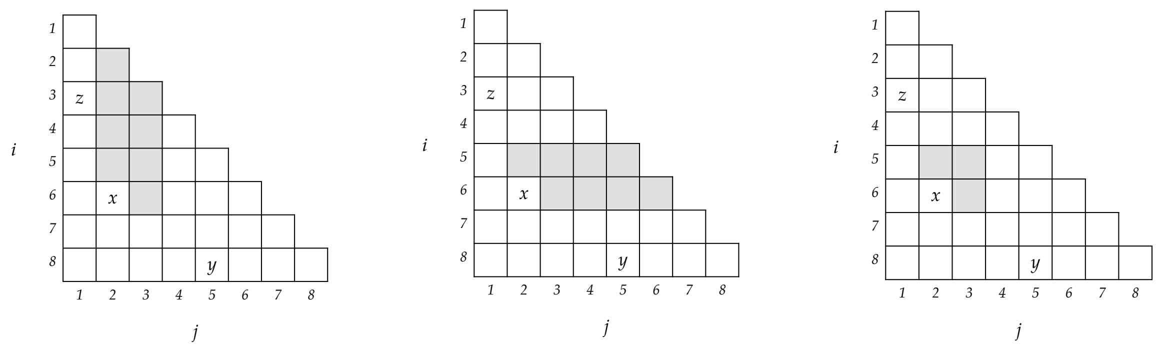
(b) Movable cells for x

Figure 4: Prohibited and movable cells

Definition 4. *An alternative x can be moved to the cell (i,j) where $i \geq j$ if (a) $i \leq i(x)$ and $j \geq j(x)$, (b) $x \succ y$ for all y with $i < j(y) \leq i(x)$, (c) $z \succ x$ for all z with $j(x) \leq i(z) < j$.*

Definition requires that the alternatives in prohibited cells must move up and right (Condition (a)). As an outcome x is moved a new cell, (i,j), it is possible that there exists y such that $i(x) \geq j(y)$ but $i < j(y)$. Condition (b) requires that in this case $x \succsim y$. Suppose to the contrary that $y \succ x$. Since $i < j(y)$, in the new representation $x \rhd' y$. But in the original representation we have $x \not\rhd y$. So the two representations must represent different preferences. Condition (c) can be understood similarly.

To understand this definition, we provide three examples (Figure 5). In Figure 5a, we have $x \succ y, z$. Since we have $z \not\rhd x$ and $x \succ z$, x cannot be moved a cell where z will eliminate x



(a) Movable cells if $x \succ z, y$ (b) Movable cells if $y, z \succ x$ (c) Movable cells if $y \succ x \succ z$

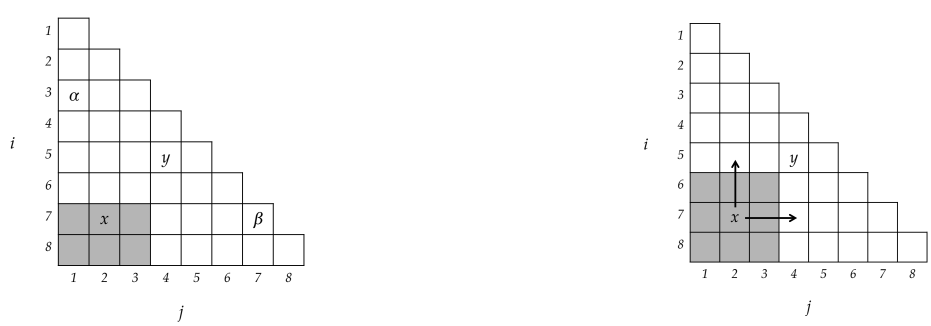
Figure 5: Examples of movable cells with different preferences

(Condition (c)). That is, $j \leq i(z) = 3$. On the other hand, since $x \succ y$, there is no restriction on movement on i. In Figure 5b, we have completely opposite situation $y, z \succ x$. Since we have $x \not\triangleright y$ and $y \succ x$, x cannot be moved a cell where x will eliminate y (Condition (b)). That is, $i \geq j(y) = 5$. On the other hand, since $z \succ x$, there is no restriction on movement on j. Finally, we provide an example where both Condition (b) and (c) induce restrictions because we have $y \succ x \succ z$.

Claim 10. *Suppose $\beta \succ y \succ \alpha$ and $\alpha \triangleright y \triangleright \beta$. If there exists x such that $x \not\triangleright \beta$ and $\alpha \not\triangleright x$, then $x \succ \beta$ or $\alpha \succ x$.*

Proof. Suppose $\beta \succ x \succ \alpha$ and we shall get a contradiction. Then we have $\beta = c(\beta x)$ and $c(x\alpha) = x$ because $x \not\triangleright \beta$ and $\alpha \not\triangleright x$. By the assumption, we have $\beta \succ c(\beta y) \succ c(\alpha y)$. By Claim 7, $\alpha\beta xy$ must be equal to either $c(\beta x) = \beta$ or $c(\alpha y) = \alpha$. Consider $c(\beta y)$ and $c(x\alpha)$, both of which are strictly worse than β and strictly better than α. Axiom 3 dictates that $\beta \succ \beta y \succsim c(\alpha\beta xy) \succsim x\alpha \succ \alpha$. Hence, $c(\alpha\beta xy)$ cannot be equal to β or α, which is a contradiction. $\square$

Given the assumptions of Claim 10, we have $j(x) \leq i(\alpha) < j(y)$ and $i(y) < j(\beta) \leq i(x)$. This means that $(i(x), j(x))$ is a prohibited cell because of y. This means that x needs to be moved. Claim 10 illustrate that x can be moved because $x \succ \beta$ or $\alpha \succ x$. The next claim shows that there is a unique way to move x. That is, x can be moved to either $(i(y), j(x))$ or $(i(x), j(y))$ but not to both.



(a) Claim 10 (b) Claim 11

Figure 6

Claim 11. *Let exist two alternatives x and y such that $i(x) > i(y)$ and $j(x) < j(y)$. Then x can be moved to either $(i(y), j(x))$ or $(i(x), j(y))$ but not to both.*

Proof. There exist two alternatives α and β such that $i(\alpha) = j(y) - 1$ and $j(\beta) = i(y) + 1$.[31] By Claim 8 and 9, we have $\beta \succ y \succ \alpha$ and $\alpha \rhd y \rhd \beta$. Since $j(\beta) \leq i(x)$ and $j(x) \leq i(\alpha)$, $x \not\rhd \beta$ and $\alpha \not\rhd x$ by Claim 8. Thus, by Claim 10, we have $x \succ \beta \succ \alpha$ (so x cannot be moved to $(i(x), j(y))$ because of α) or $\beta \succ \alpha \succ x$ (so x cannot be moved to $(i(y), j(x))$ because of β). Therefore, all we need to show is that x can be moved to either of them.

Case I: $x \succ \beta$. We show that x can be moved to $(i(y), j(x))$. First, Condition (a) holds trivially: $i(y) \leq i(x)$ and $j(x) \geq j(x)$. For Condition (b), take an element z such that $i(y) < j(z) \leq i(x)$ (so $y \rhd z$ but $x \not\rhd z$). Then, it must be either $y \succ z$ (which implies $x \succ z$) or $z \succ y = c(yz)$ in which case we have $z \succ y \succ \alpha$ and $\alpha \rhd y \rhd z$ (with $x \not\rhd z$ and $\alpha \not\rhd x$). By Claim 10, we should have $x \succ z$ or $\alpha \succ x$. Since we are considering the case $x \succ \beta (\succ \alpha)$, it must be $x \succ z$. Condition (c) is trivially satisfied because $j = j(x)$.

Case II: $\alpha \succ x$: Condition (a) and (b) will be now trivial while Condition (c) can be proven in the same way how we prove Condition (b) in case I. $\square$

Claim 12. *Let*

$$U_x = \{y : i(x) > i(y),\ j(x) < j(y) \text{ and } x \text{ can be moved to } (i(y), j(x))\} \cup \{x\}$$
$$R_x = \{y : i(x) > i(y),\ j(x) < j(y) \text{ and } x \text{ can be moved to } (i(x), j(y))\} \cup \{x\}$$

and let

$$i_x = \min_{y \in U_x} i(y) \text{ and } j_x = \max_{y \in R_x} j(y)$$

Then (i) x can be moved to (i_x, j_x), and (ii) (i_x, j_x) is a safe cell. That is, there is no z with $i(z) < i_x$ and $j(z) > j_x$.

Proof. Notice that by the definitions of movability, $i_x \leq i(x)$ and $j_x \geq j(x)$.

(i) Clearly, $i_x \leq i(x)$ and $j_x \geq j(x)$ as $x \in U_x, R_x$. First, we show that $i_x \geq j_x$. Take an alternative $y \in U_x$ such that $i(y) = i_x$. Since $y \in U_x$, x cannot be moved to $(i(x), j(y))$ by Claim 11. By the definition of movability, x cannot be moved to $(i(x), j)$ if $j \geq j(y)$. Hence for all $z \in R_x \setminus \{x\}$, $j(z) < j(y)$, which means $j_x = \max_{z \in R_x} j(z) \leq j(y)$. Since $i(y) \geq j(y)$, we have $j_x \leq j(y) \leq i(y) = i_x$.

Since x can be moved to $(i_x, j(x))$, then the second condition of the movability is satisfied. Similarly, we can prove the third requirement as well. Therefore, x can be moved to (i_x, j_x)

(ii) If $z \notin U_x, R_x$, then by Claim 11, it must be $(i_x \leq) i(x) \leq i(z)$ or $j(z) \leq j(x) (\leq j_x)$. If $z \in U_x$, then $i(z) \geq i_x$. If $z \in R_z$ then $j(z) \leq j_x$. $\square$

Now, define $x \bar{\rhd} y$ if and only if $j_y > i_x$.

Claim 13. *$\bar{\rhd}$ is a semi-order.*

[31] This is because since neither $i(y)$ is the smallest nor $i(y)$ is the largest integer within the range of i.

Proof. Since $i_x \geq j_x$ by Claim 12 for all x, $\bar{\rhd}$ is an interval order.

Next, we shall show that if (i, j) is a safe cell, there is no element x such that $i_x < i$ and $j_x > j$. Suppose there is such x. Notice that it must be $i \leq i(x)$ or $j \geq j(x)$ because (i, j) is a safe cell so it must be $i_x < i(x)$ or $j > j(x)$. Suppose $i_x < i(x)$. Then there exists y such that $i(y) = i_x$ and $j(x) < j(y)$ such that x can be moved to $(i(y), j(x))$. By Claim 11, x cannot be moved to $(i(x), j(y))$, so it cannot be moved to $(i(x), j')$ for any $j' \geq j(y)$. Since (i, j) is a safe cell and $i(y) = i_x < i$, it must be $j(y) \leq j(< j_x)$. Hence, x cannot be moved to $(i(x), j_x)$, which contradicts the definition of j_x unless $j_x = j(x)$. But if so, $j(y) > j_x > j$ but this contradicts that (i, j) is a safe cell. Analogously, we can show a contradiction if $j > j(x)$.

By Claim 12, all elements have been moved to safe cells, so there is no pair of elements x and y such that $i_x < i_y$ and $j_x > j_y$. Therefore, if $i_x < j_y \leq i_y < j_z$ (i.e. $x \bar{\rhd} y \bar{\rhd} z$) then for any w, it must be either $j_w > j_y$ or $i_w \leq i_y$, which implies $j_w > i_x$ or $i_w < j_z$ (i.e. $x \bar{\rhd} w$ or $w \bar{\rhd} z$). Therefore, $\bar{\rhd}$ is a semiorder. □

Claim 14. *If $x \rhd y$ then $x \bar{\rhd} y$.*

Proof. By definitions of i' and j', $i_x \leq i(x)$ and $j_x \geq j(x)$ for all x. Therefore, if $x \rhd y$, then $i_x \leq i(x) < j(y) \leq j_y$ so we have $x \bar{\rhd} y$. □

Claim 15. *If $x \bar{\rhd} y$ but not $x \rhd y$, then $x \succ y$.*

Proof. First, we shall note that both x and y must be in prohibited cells. If neither of them is in, $i_x = i(x)$ and $j_y = j(y)$ so $x \bar{\rhd} y$ and not $x \rhd y$ cannot happen at the same time. If only x is in a prohibited cell, then $i_x < j(y) \leq i(x)$ so x cannot be moved to (i_x, j_x). Similarly we can prove that it is not possible that only y is in a prohibited cell.

Next we shall show that $i_x < i(x)$ and $j_x > j(x)$. Since x can be moved to (i_x, j_x) while $y \succ x$, it must be $i_x \geq j(y)$ because $j(y) \leq i(x)$ (i.e. not $x \rhd y$). Combined with $x \bar{\rhd} y$, we get $j_y > j(y)$. Flipping x and y, one can prove $i_x < i(x)$.

Therefore, there must exist z and z' with $i(z) \in [j(y), j_y - 1]$ and $j(z') \in [i_x - 1, i(x)]$ (notice that these intervals are non-empty). Furthermore, we can take such z and z' so that $i(z) = j(z') - 1$ because $i_x - 1 < j_y - 1$ and $i(x) > j(y)$. Thus, $z' \succ z$ by Claim 9. Since x is movable to (i_x, j_x), we have $x \succ z'$. Similarly, we have $z \succ y$. Therefore, we conclude $x \succ y$. □

Claim 16. *$c(S)$ is equal to the $\succ$-best element in $\Gamma_{\bar{\rhd}}(S)$.*

Proof. We know $\bar{\rhd}$ is transitive, $\bar{\rhd} \supseteq \rhd$ and $x \succ y$ whenever $x \bar{\rhd} y$ but not $x \rhd y$. It is easy to see that this claim can be proven in the exactly same way as Claim 6. □

(Representation $\Rightarrow$ Axioms 1-4) Showing that the first axiom is necessary is straightforward. Let

$$\Gamma(A) = \{x \in A : \max_{y \in A} v(y) - v(x) \leq \overline{w}\}$$

For Axiom 2, if $x \succ c(A \cup x)$ then A must have an element y with $v(y) > v(x) + \overline{w}$, so it is clear that $\Gamma(A) = \Gamma(A \cup x)$ so $c(A) = c(A \cup x)$.

Axiom 3: Let x^* be the u-best element in $\Gamma(A \cup B)$. Then it must be in $\Gamma(A)$ or $\Gamma(B)$ so it is not possible that $c(A \cup B)$ is strictly preferred to both $c(A)$ and $c(B)$. Now we show that the union cannot be strictly worse than both. Let x_A and x_B be the u-best elements in $\Gamma(A)$ and $\Gamma(B)$, respectively, and take v_A and v_B be the maximum values of v in A and in B, respectively. Then we have

$$v_A \leq v(x_A) + \overline{w} \text{ and } v_B \leq v(x_B) + \overline{w}$$

Therefore the maximum value of v in $A \cup B$ is the higher one between v_A and v_B, either x_A or x_B must be in $\Gamma(A \cup B)$ so $c(A \cup B)$ must be weakly better than either $c(A)$ or $c(B)$.

Finally we show that the representation implies Axiom 4. Suppose $x \succ c(xy) \succ c(yz)$, then it must be $x \succ y \succ z$, $v(y) - v(x) > \overline{w}$ and $v(z) - v(y) > \overline{w}$. Therefore, $v(z) - v(x) > 2\overline{w}$.

Since $c(tz) = t$, we must either "$z \succ t$ and $v(t) - v(z) > \overline{w}$" or "$t \succ z$ and $v(z) - v(t) \leq \overline{w}$." In both cases, we have $v(t) - v(x) > \overline{w}$, hence we have $c(xt) = t$.

Proof of Theorem 3

By Axiom 5, $\succsim$ is represented by a linear function u. Hence the usual uniqueness result applies to u. Now, we need to prove that c is represented by our model. In other words, our task is to show that there exist v and w that represent c along with u.

Let $p\alpha q$ represent the lottery r where $r(x)$ is equal to $\alpha p(x) + (1 - \alpha)q(x)$ for all $x \in X$. We allow the possibility of $\alpha < 0$ or $\alpha > 1$ as long as $r(x) \in [0, 1]$ for all $x \in X$. Thus, whenever we refer to α greater than 1 or less than 0, we implicitly assume that the mixture is well-defined. Also, we understand that $c(A) = p$ is an abbreviation of $c(A) = \{p\}$.

Through the proof, we are often interested in whether $\succsim$ and the choice are consistent or not. To do so we define the following terminologies:

Definition 5. *Suppose $p \succ q$. p is* ***choosable*** *over q if $c(p, q) = p$. Similarly, p is* ***unchoosable*** *over q (q* ***blocks*** *p) if $c(p, q) = q$. Finally, p is* ***just choosable*** *over q if (i) p is choosable over $p\alpha q$ if $0 < \alpha < 1$ but unchoosable over $p\alpha q$ if $\alpha < 0$ and (ii) $p\alpha q$ is choosable over q if $0 < \alpha < 1$ but unchoosable over q if $\alpha > 1$.*

Notice that if p is just choosable over q, it is indeed choosable over q by Axioms 2 and 11. To see this take a sequence α_n in $(0, 1)$ approaching zero. Since $c(p, p\alpha_n q) = p$ for all α_n, Axiom 11 implies that $p \in c(p, q)$. Since $p \succ q$, Axiom 2 implies that $c(p, q) = p$, thus p is choosable over q.

Claim 17. *If p is choosable over q, then $p\alpha q$ is choosable over $p\beta q$ for any α and β with $0 \leq \beta \leq \alpha \leq 1$. Moreover, if p is unchoosable over q, then $p\alpha q$ is unchoosable over $p\beta q$ for any α and β with $\beta \leq 0$ and $\alpha \geq 1$.*

Proof of Claim 17: Assume p is choosable over q. Take α and β with $0 \leq \beta \leq \alpha \leq 1$. By Axiom 9, $p\alpha q$ is choosable over q because $q = q\alpha q$ and $c(p, q) = p$. Applying Axiom 9 again, we have $p\alpha q$ is choosable over $(p\alpha q)\frac{\beta}{\alpha}q = p\beta q$ given that $\beta \leq \alpha$.

Now take $\alpha \geq 1$ and $\beta \leq 0$ and assume p is unchoosable over q. If $p\alpha q$ is choosable over q, by the first part of the claim, we must have p is choosable over q since $(p\alpha q)\frac{1}{\alpha}q = p$, a contradiction. Hence, $p\alpha q$ is unchoosable over q. Similarly, if $p\alpha q$ is choosable over $p\beta q$, by the first part of the claim, we must have $p\alpha q$ is choosable over q since $(p\alpha q)\frac{-\beta}{\alpha-\beta}(p\beta q) = q$ $(0 \leq \frac{-\beta}{\alpha-\beta} \leq 1)$, a contradiction. Hence, $p\alpha q$ is unchoosable over $p\beta q$. $\square$

Claim 18. *Let $p \succ q$. If p is unchoosable over q, then there exists unique $\alpha \in (0,1)$ such that $p\alpha q$ is just choosable over q and p is just choosable over $p(1-\alpha)q$.*

Proof of Claim 18: The uniqueness of such α is directly implied by Axiom 9. Thus, we shall prove the existence. Let $C = \{\alpha' \in (0,1] : p\alpha' q \text{ is choosable over } q\}$. By Axiom 13, $\epsilon \in C$ for sufficiently small positive ϵ but $1 \notin C$. Let $\alpha = \sup C$, which is strictly positive. Furthermore, C is convex by Axiom 9. Let $p' = p\alpha q$. Since $p'\beta q = p(\alpha\beta)q$, $p'\beta q$ is choosable over q if $0 < \beta < 1$ and unchoosable if $\beta > 1$. We now show that p' is choosable over $p'\beta q$ if $0 < \beta < 1$ but unchoosable over $p'\beta q$ if $\beta < 0$. To prove the second part by a contradiction, assume p' is choosable over $p'\beta q$ for some when $\beta < 0$. Then define $\lambda = \frac{-\alpha\beta}{1-\alpha\beta} > 0$. Note that $p\lambda(p'\beta q)$ is equal to q. Since $\lambda > 0$, $p\lambda p'$ is equal to $p\alpha' q$ where $\alpha' > \alpha = \sup C$. Moreover, by Axiom 9, we have $p\lambda p'$ is choosable over $p\lambda(p'\beta q)$. This is a contradiction to the fact that $\alpha = \sup C$. When $0 < \beta < 1$, since $\alpha > \alpha\beta$, $\alpha\beta \in C$. By Claim 17, we conclude p' is choosable over $p'\beta q$. Therefore, $p' = p\alpha q$ is just choosable over q.

Similarly, define $C' = \{\alpha' \in (0,1] : p \text{ is choosable over } p(1-\alpha')q\}$. The exactly same argument shows that p is just choosable over $p(1-\sup C')q$. We shall show that $\sup C' = \sup C (= \alpha)$ simply by showing $C' = C$. Suppose $\beta \in C$, which means that $p\beta q$ is choosable over $q = q\beta q$. Thus, $p = p\beta p$ is choosable over $q\beta p = p(1-\beta)q$ by Axiom 10 so $\beta \in C'$. Similarly, $\beta \notin C$ implies $\beta \notin C'$. Therefore, $C = C'$. $\square$

Claim 19. *Let $p\gamma q$ be just choosable over q with $0 < \gamma \le 1$ and both $p\alpha q$ and $p\beta q$ exist. Then, for any $\alpha > \gamma$, $p\alpha q$ is just choosable over $p\beta q$ if and only if $\alpha - \beta = \gamma$.*

Proof of Claim 19: We shall show that $p\alpha q$ is just choosable over $p(\alpha-\gamma)q$. Let $p' = p\alpha q$. Then, $p'(\gamma/\alpha)q = p\gamma q$, hence $p'(\gamma/\alpha)q$ is just choosable over q. By Claim 18 and $\gamma/\alpha \in (0,1)$, $p'(= p\alpha q)$ is just choosable over $p'(1-(\gamma/\alpha))q(= p(\alpha-\gamma)q)$. $\square$

Claim 20. *Let $p\gamma q$ be just choosable over q with $0 < \gamma \le 1$ and both $p\alpha q$ and $p\beta q$ exist. Then $p\alpha q$ is just choosable over $p\beta q$ if and only if $\alpha - \beta = \gamma$.*

Proof of Claim 20: If $\alpha > \gamma$, then we are done by Claim 19. If $\alpha = \gamma$, the statement is trivially true. Hence assume $\alpha < \gamma$. By Claim 17, $p\alpha q$ is choosable over q, which implies β is less than 0. More importantly, we have $\beta < 1-\gamma < 1$. Let $p'' = p\beta q$. Then, $p(\frac{1-\gamma-\beta}{1-\beta})p''$ is equal to $p(1-\gamma)q$. Notice that $\frac{1-\gamma-\beta}{1-\beta} \in (0,1)$. Since p is just choosable over $p(\frac{1-\gamma-\beta}{1-\beta})p''$, by Claim 18, we have $p''(\frac{1-\gamma-\beta}{1-\beta})p$ is just choosable over p''. Given that $p''(\frac{1-\gamma-\beta}{1-\beta})p$ is equal to $p(\beta+\gamma)q$, we reach the desired conclusion: $p(\beta+\gamma)q$ is just choosable over $p\beta q$. $\square$

Claim 21. *Let p, q be interior of Δ^{N-1} and p is choosable over q. Then there exists $\varepsilon > 0$ such that for any p' in ε-neighborhood of p, say $N_\varepsilon(p)$, and any $q' \in \Delta^{N-1}$ with $p'_i - q'_i = p_i - q_i$ for all i, p' is choosable over q'. The statement is also true for "unchoosable" and "just choosable" relationships.*

Proof of Claim 21: We first show that the claim is true for "choosable" relationship. Let r be $p\frac{1}{2}q$. Since p and q are interior points find $p_0 \neq p$, $q_0 \neq q$ and $\alpha \in (0,1)$ such that $p = p_0\alpha r$ and $q = q_0\alpha r$. It is routine to check that $p'_i - p_i = (1-\alpha)(r'_i - r_i)$ for all i (see Figure 7). Take $\varepsilon > 0$ and let $N_\varepsilon(p)$ be the ε ball around p. Choose ε small enough so that $N_\varepsilon(p)$ belongs to the simplex. Then for each p' in $N_\varepsilon(p)$, choose r' so that $p' = p_0\alpha r'$. As ε gets smaller, all p''s get close to p.

Consequently, all r''s get close to r. Since r is in the interior of the simplex, we choose ε small enough so that all r''s are in the simplex as well. We then fix the ε, and set the neighborhood as $N_\varepsilon(p)$.

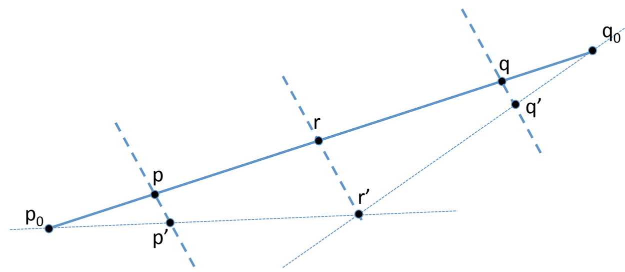


Figure 7: Paralel Shift

Next, take any p' from the neighborhood of p constructed above. We know show that for q' such that $p'_i - q'_i = p_i - q_i$ for all i, it must be $c(p', q') = p'$. Note that we have $c(p_0\alpha r, q_0\alpha r) = p_0\alpha r$. By Axiom 10, $c(p_0\alpha r', q_0\alpha r') = p_0\alpha r'$. Since $p'_i - p_i = q'_i - q_i = (1-\alpha)(r'_i - r_i)$, we have $q' = q_0\alpha r'$ (see Figure 7). Hence, $c(p', q') = p'$.

Since the proof of "unchoosable" relationship is similar to above, we omit it. If both the choosable and unchoosable relationships are preserved for small parallel shifts, "just choosable" relationship will also be preserved by Claim 17. $\square$

By Axiom 12 there exist g' and b such that $g' \succ b$ but $c(g', b) = b$. By Axiom 11, we can find such g' and b from the set of completely mixed lotteries. By Claim 18, there exists a lottery g between g' and b which is just choosable over b. Since b and g are interior points, there exist ε and ε' such that $N_\varepsilon(g) \cup N_{\varepsilon'}(b) \subset \Delta^{N-1}$ with $p \succ q$ for any $p \in N_\varepsilon(g)$ and $q \in N_{\varepsilon'}(b)$ (Axiom 11). Consider two subsets of $N_\varepsilon(g)$:

$$A(b) = \{p \in N_\varepsilon(g) : p \succ b = c(p, b)\} \quad \text{and} \quad B(b) = \{p \in N_\varepsilon(g) : p = c(p, b) \succ b\}$$

Note that $A(b) \cup B(b) = N_\varepsilon(g)$ and $A(b) \cap B(b) = \emptyset$.[32] By Axiom 9, both sets are convex. Therefore, there exists a hyperplane V_0 such that $g \in V_0$ and V_0 separates $A(b)$ and $B(b)$. By Axiom 11, $V_0 \cap N_\varepsilon(g)$ is a subset of B. We next show that all points (within $N_\varepsilon(g)$) on the hyperplane are just choosable over b. Take a point on V_0, say p. Consider the line passing through p and b. Since p is choosable over b, for all $0 < \alpha \leq 1$, $p\alpha b$ is also choosable over b. Since V_0 is a hyperplane, there exists α' such that $p\alpha' b$ is in $A(b)$. That is, $p\alpha' b$ is not choosable over b. Moreover, for all $1 < \alpha < \alpha'$ we have $p\alpha b$ is in $A(b)$. Then by Claim 18, there exists unique α'' such that $\alpha' > \alpha'' \geq 1$ and $p\alpha'' b$ is just choosable over b, hence $p\alpha'' b \in B(b)$. If $\alpha'' > 1$ then we have a contradiction since $A(b) \cap B(b) = \emptyset$. Hence, p is just choosable over b.

Similarly, there exists a hyperplane V_1 such that $b \in V_1$ and V_1 divides $N_{\varepsilon'}(b)$ into the two regions: $A(g) = \{q \in N_{\varepsilon'}(b) : g = c(q, g) \succ q\}$ and $B(g) = \{q \in N_{\varepsilon'}(b) : g \succ q = c(q, g)\}$, and g is just choosable over all points (within $N_{\varepsilon'}(b)$) on the hyperplane.

[32]Since $p \succ b$ for any $p \in N_\varepsilon(g)$, by Axiom 2, $c(p, b)$ must be either p or b.

Claim 22. *V_0 and V_1 are parallel.*

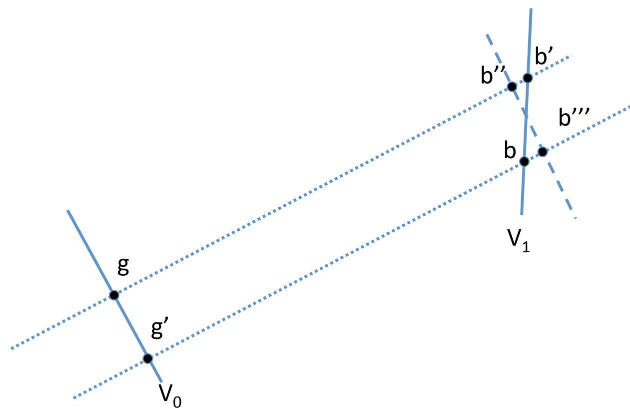


Figure 8: Hyperplanes are parallel

Proof of Claim 22: Figure 9 is helpful to follow the proof. Suppose two hyperplanes V_0 and V_1 are not parallel. Hence there exist two alternatives $g' \in V_0 \setminus \{g\}$ and $b' \in V_1 \setminus \{b\}$ such that $g'_i - b_i = \lambda(g_i - b'_i)$ for all i with $0 < \lambda < 1$. Geometrically, this means that (i) $g'b$ line is parallel to gb' line and (ii) the distance between b and g' is smaller than the distance between g and b'. Now consider b'' such that $b''_i - b'''_i = g_i - g'_i$ for all i. This is a shift parallel to V_0. Since V_0 and V_1 are not parallel, b'' lies strictly between g and b'. Since g is just choosable over b', g is choosable over b''. Since we can choose g' arbitrary close to g, we can invoke Claim 21 to show that g' is choosable over $b^{'''}$. This yields a contradiction since g' is just choosable over b and b lies strictly between g' and b''' (Claim 18). □

Claim 23. *There exist two open neighborhoods of g and b, say N_g and N_b, such that p is just choosable over q for any $p \in V_0 \cap N_g$ and $q \in V_1 \cap N_b$.*

Proof of Claim 23: Since g is just choosable over b, by Claim 21, there exists a neighborhood of g, $N_\varepsilon(g)$, such that for any $\hat{g}$ in $N_\varepsilon(g)$ and any $\hat{b} \in N_\varepsilon(b)$ with $\hat{g}_i - \hat{b}_i = g_i - b_i$ for all i, $\hat{g}$ is just choosable over $\hat{b}$. Now consider $N_\varepsilon(g) \cap V_0$. By Claim 22, if $\hat{g}$ is on V_0, then $\hat{b}$ must be on V_1. Hence we have $\hat{g}$ is just choosable over b and g is just choosable over $\hat{b}$.

Now take any $p \in N_{\varepsilon/2}(g) \cap V_0$ and $q \in N_{\varepsilon/2}(b) \cap V_1$. There exists $g' \in N_\varepsilon(g) \cap V_0$ such that p is the mid-point of g and g', i.e., $p = g\frac{1}{2}g'$. Similarly, there exists $b'' \in N_\varepsilon(b) \cap V_1$ such that q is the mid-point of b and b'', i.e., $q = b\frac{1}{2}b''$. Now define g'' in $N_\varepsilon(g) \cap V_0$ and $b' \in N_\varepsilon(b) \cap V_1$ such that $g''_i - b''_i = g'_i - b'_i = g_i - b_i$ for all i. We have show above that g' is just choosable over b and g is just choosable over b''. By applying Axiom 9, we get $g'\frac{1}{2}g$ is just choosable over $b\frac{1}{2}b''$. Since $p = g\frac{1}{2}g' = g'\frac{1}{2}g$ and $q = b\frac{1}{2}b''$, p is just choosable over q. □

Claim 24. *Let $p \in V_0 \cap N_g$ and $q \in V_1 \cap N_b$ and α, β (not necessarily between 0 and 1). Then $p\alpha q$ is just choosable over $p\beta q$ if and only if $\alpha - \beta = 1$.*

Proof of Claim 24: By Claim 23, p is just choosable over q. Consider the line crossing p and q. By Claim 20, on this line, $p\alpha q$ is just choosable over $p\beta q$ if and only if $\alpha - \beta = 1$. □

Now we shall construct v. Let $L_{bg} = \{p : p = b\alpha g \text{ for some } \alpha\}$, which is the set of all linear combinations of g and b in Δ^{N-1}. We first define v only over L_{bg} as follows:

$$v(p) = \alpha \text{ where } p = b\alpha g$$

Note that $v(b) = 1$ and $v(g) = 0$. Directly applying Claim 20, we have the following claim stating that the representation holds for any binary set contained in L_{bg}.

Claim 25. *For any $p, q \in L_{bg}$ with $p \succ q$, p is choosable over q if and only if $v(q) - v(p) \leq 1$.*

Our next step is to define v for all lotteries and show that the representation holds for any binary set. For any $p \in V_0$ $q \in V_1$, define $v(p) = 1$ and $v(q) = 0$. For all other points, define v so that v is linear.

Claim 26. *Suppose $p \succ q$. Then $p = c(p, q)$ if and only if $v(q) - v(p) \leq 1$.*

Proof of Claim 26: Take two points p and q such that $p \succ q$ and $1 \geq v(q) - v(p) := \lambda$. Suppose $q = c(p, q)$, i.e. p is not choosable over q. Consider $g_\alpha = b\alpha g$ with $\alpha < 0$. Since g is just choosable over b, g_α is not choosable over b. By Axiom 9, $g_\alpha \beta p$ is not choosable over $b\beta q$.

Given $\alpha < 0$ there exists $\beta < 1$ such that $v(b\beta q) - v(g_\alpha \beta p) = 1$. We can calculate the relation between α and β. To see this,

$$\begin{aligned} 1 &= (1-\alpha)\beta + (1-\beta)[v(q) - v(p)] \\ \beta(\alpha) &= \frac{1-\lambda}{1-\lambda-\alpha} \end{aligned}$$

Note that $\beta(\alpha)$ is less than 1 and goes monotonically to 1 as α approaches to 0. By choosing α close enough to zero, we have $g_\alpha \beta p \in N_g$ and $b\beta q \in N_b$ are true.

Now consider the line goes through these two points. Since $v(b\beta q) \neq v(g_\alpha \beta p)$, this line intersects with V_0 and V_1. Denote g' and b' as the intersections of this line and V_0 and V_1, respectively. There exists α' such that $g'\alpha' b' = g_\alpha \beta p$. We have

$$\begin{aligned} v(g'\alpha' b') &= v(g_\alpha \beta p) \\ \alpha' v(g') + (1-\alpha')v(b') &= \beta v(g_\alpha) + (1-\beta)v(p) \\ 1-\alpha' &= \beta\alpha + (1-\beta)v(p) \\ 1-\alpha'+1 &= \beta\alpha + (1-\beta)v(p) + (1-\alpha)\beta + (1-\beta)[v(q) - v(p)] \\ 2-\alpha' &= \beta + (1-\beta)v(q) \\ v(g'(\alpha'-1)b') &= v(b\beta q) \end{aligned}$$

Hence, $g'(\alpha'-1)b' = b\beta q$. By Claim 24, $g'\alpha' b'$ is just choosable over $g'(\alpha'-1)b'$. That is, $g_\alpha \beta p$ is just choosable over $b\beta q$, a contradiction.

Similarly, one can show that if $p \succ q$ and $1 < v(q) - v(p)$ then $c(p, q)$ is equal to q. $\square$

Claim 27. *Suppose $p \sim q$. Then $p \in c(p, q)$ if and only if $v(q) - v(p) \leq 1$.*

Proof of Claim 27: Take two points p and q such that $p \sim q$ and $1 \geq v(q) - v(p) := \lambda$. Consider a decreasing sequence of real numbers $\alpha_k > 0$ with $\lim_{k \to \infty} \alpha_k = 0$. Define $g\alpha_k p$ and $b\alpha_k q$. Since

$g \succ b$ and $p \sim q$, we have $g\alpha_k p \succ b\alpha_k q$ for all α_k. In addition, we have $v(b\alpha_k q) - v(g\alpha_k p) = \alpha_k + (1-\alpha_k)\lambda \leq 1$ for all α_k since $\lambda \leq 1$. By Claim 26, we have $g\alpha_k p = c(g\alpha_k p, b\alpha_k q)$. By Axiom 11, $p \in c(p,q)$.

To see the converse, suppose, towards a contradiction, $p \sim q$ and $1 < v(q) - v(p)(:= \lambda')$ but $p \in c(p,q)$. Take an $\alpha \in (0,1)$ and consider $g\alpha p$ and $b\alpha q$. Since $g \succ b$ and $p \sim q$, we have $g\alpha p \succ b\alpha q$. In addition, we have $v(b\alpha q) - v(g\alpha p) = \alpha + (1-\alpha)\lambda' > 1$ since $\lambda' > 1$. By Claim 26 and Axiom 2, we have $g\alpha p \notin c(g\alpha p, b\alpha q)$. This contradicts Axiom 9 (part i) since $g = c(g,b)$, $p \in c(p,q)$ and $p \sim q$. $\square$

Claim 26 and Claim 27 establish the representation for any binary set. The next goal is to extend the representation from binary sets to any finite set.

Claim 28. *$c(A) = \arg\max_{p\in A} u(p)$ subject to $\max_{q\in A} v(q) - v(p) \leq w$ for all finite set of lotteries A.*

Proof of Claim 28 Let c' be the choice correspondence represented by $(u, v, 1)$ as we constructed from c. We know that $c(A) = c'(A)$ whenever $|A| = 2$. Now suppose that $c(A) = c'(A)$ whenever $|A| \leq n$. Let $|B| = n+1$.

Suppose $x \in c(B)$ but $x \notin c'(B)$. By construction of c' there are two cases to consider. First case is that $x \notin c'(B)$ because there is another choosable element in B that is preferred to x. The second case is that x itself is not choosable. We will now consider the two case and in turn and show that each leads to a contradiction.

- There exists $y \in B$ such that $y \succ x$ but there is no $z \in B$ such that $v(z) - v(y) > 1$. This implies that $c'(B) \succsim y$. In addition, for any $T \subset B$ including y, we have $c'(T) \succsim y$. By the inductive hypothesis, $c(B \setminus z) \succsim y$ for any $z \in B \setminus y$. Since $|B| \geq 3$, take two alternatives z and z' in $B \setminus y$. Since $c(B \setminus z), c(B \setminus z') \succsim y$, by Axiom 3 we have $c(B) \succsim y \succ x$, which is a contradiction.
- Suppose there exists $y \in B$ with $v(y) - v(x) > 1$. Without loss of generality, assume that $v(y) = \max_{y' \in B}\ v(y')$.
  - We know that $y \succ x$ is not possible from the previous case.
  - If $y \sim x$ then we have $c'(x,y) = c(x,y) = y$. Hence $x \notin c(x,y) \sim x \sim c(B)$, contradicting Axiom 6.
  - If $y \prec x$, then x must be one of the best elements in B. If not, there exists $z \succ x$ and Axiom 2 implies $c(B \setminus z) = c(B)$ so $x \in c(B \setminus z)$, which contradicts the inductive hypothesis because $y \in B \setminus z$ and $c(B \setminus z) = c'(B \setminus z)$. Now, consider the sets $A = \{z : z \in B \text{ and } z \sim x\}$ and $B' = B \setminus (A \setminus x)$. Note that x is the unique best element in B'. By Axiom 8, $c(B') \succsim c(B)$ so $c(B') = x$, which also contradicts the inductive hypothesis ($y \in B'$ and $c(B') = c'(B')$).

Therefore, we conclude that $c(B) \subset c'(B)$.

For the other direction, suppose $x \in c'(B)$. Since $c(B) \subset c'(B)$, for any $y \in c(B)$, we have $y \in c'(B)$. By construction of c', we have $x \sim y$. If there is no alternative in B such that $z \succ x$, then $x \in c'(x,z) = c(x,z)$ for all $z \in B$. Thus, by Axiom 7, $x \in c(B)$ so $c'(B) \subset c(B)$.

Assume there are some alternatives which rank above x (or y), say z. Then by Axiom 2 $c(B) = c(B \setminus z)$. By induction hypothesis, $c(B) = c(B \setminus z) = c'(B \setminus z)$. Since $x \in c'(B \setminus z)$ by definition of c', we have $x \in c(B)$ so $c'(B) \subset c(B)$. $\square$

Proof of Theorem 4

Since v is linear, it is routine to show that (u, v', w') also represents c if there exist $\alpha' > 0$ and β' such that $v'(x) = \alpha' v(x) + \beta'$and $w' = \alpha'$. This proves our uniqueness result. $\square$

Proof of Theorem 5

It is straightforward to verify the if-part so we only show the only-if part. Suppose controls have more willpower than treatments under the same temptation.

Hence, each subject $i \in \{cont, treat\}$ is represented by some (u, v_i, w_i) and such (v_i, w_i) is unique if both v_i's are normalized so that $v_{cont}(x) = v_{treat}(x) = 0$ for some x, and $\Sigma_{x \in X}(v_i(x))^2 = 1$. Suppose that v_{cont} and v_{treat} have been normalized in such a way.

If $v_{cont} = v_{treat}$, clearly it must be $w_{cont} \geq w_{treat}$ so we have the desired result. Thus, we shall show that it must be $v_{cont} = v_{treat}$. Suppose $v_{cont} \neq v_{treat}$, then there exists $\alpha \in R^{|X|}$ with $\Sigma_{x \in X} \alpha_x = 0$ such that $\nabla v_{cont} \alpha > 0 > \nabla v_{treat} \alpha$ where $\nabla v_i = (\partial v_i / \partial x_i)_{i \in X}$. Take $g, b \in int(\Delta)$ with $g \succ b$ and $v_{cont}(b) - v_{cont}(g) = w_{cont}$ (see Figure 9). Since treatments choose a worse alternative whenever controls do, it must be $v_{treat}(b) - v_{treat}(g) \geq w_{treat}$.

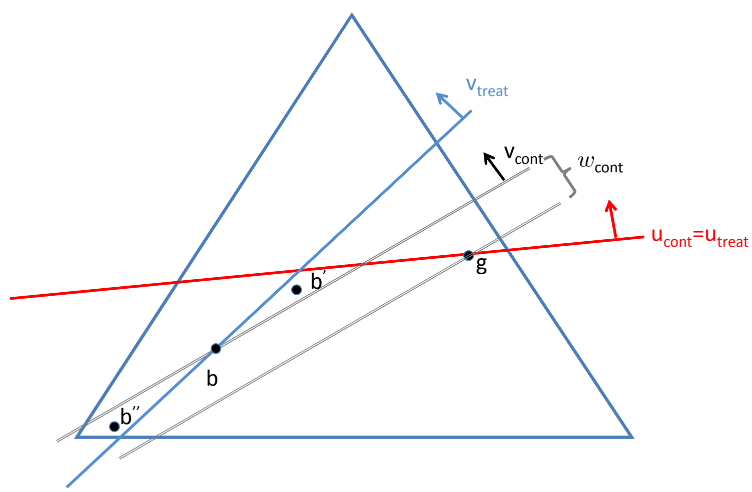


Figure 9: Comparative Statics

Consider two lotteries $b' = b + \varepsilon\alpha$ and $b'' = b - \varepsilon\alpha$, both of which are element of Δ as long as $\varepsilon > 0$ is small enough. Then, $v_{cont}(b') > v_{cont}(b) > v_{cont}(b'')$ so $c_{cont}(g, b') = b'$ and $c_{cont}(g, b'') = g$. In contrast, it is $v_{treat}(b'') > v_{treat}(b) > v_{treat}(b')$. Hence, we can find $\beta \in (0, 1]$ such that $v_{treat}(g\beta b'') - v_{treat}(g) > w_{treat} > v_{treat}(g\beta b') - v_{treat}(g)$, which implies $c_{treat}(g, g\beta b') = g$ but $c_{treat}(g, g\beta b'') = g\beta b''$. This contradicts the second requirement of the definition. $\square$

Proof of Theorem 6

Since $(\succsim, c)$ satisfies Axioms 1-11, it has a limited willpower representation (u, v, w).

Suppose $y \succsim x$ but $x \succ^c y$. Then by definition of $\succ^c$ either (i) $x = c(x, y)$ and there exists no $\alpha \in (0, 1)$ such that $y \in c(x\alpha y, y)$ or (ii) $y = c(x, y)$ and there exists some $\alpha \in (0, 1)$ such that $x\alpha y = c(x\alpha y, y)$.

Since $y \succsim x$, $u(y) \geq u(x)$. In case (i), $x = c(x, y)$, thus we have $v(x) - v(y) > w$. Since v is linear, $v(x\alpha y) - v(y) = \alpha (v(x) - v(y))$. For α small enough, we have $\alpha (v(x) - v(y)) < w$. In addition $u(y) \geq u(x\alpha y)$. Thus for α small, it must be that $y \in c(x\alpha y, y)$ which is a contradiction. In case (ii), $y = c(x, y)$, thus we have $v(x) - v(y) \leq w$. For all $\alpha \in (0, 1)$, $v(x) - v(x\alpha y) = (1 - \alpha)(v(x) - v(y)) \leq w$. Hence we must have $x\alpha y \in c(x\alpha y, y)$, which is a contradiction. Thus $y \succsim x$ implies $y \succsim^c x$.

Next suppose $y \succsim^c x$, but $x \succ y$. Since $u(x) > u(y)$, for any $\alpha \in (0, 1]$, $u(x\alpha y) > u(y)$. There are two cases: If $v(y) - v(x) \leq w$ then $v(y) - v(x\alpha y) = \alpha (v(y) - v(x)) \leq w$, and $x\alpha y = c(x\alpha y, y)$ for any $\alpha \in (0, 1]$. If $v(y) - v(x) > w$, $v(y) - v(x\alpha y) = \alpha (v(y) - v(x)) < w$ and $x\alpha y = c(x\alpha y, y)$ for α close to zero. In either case $x \succ^c y$, a contradiction. Thus $y \succeq x$ if and only if $y \succsim^c x$. $\square$