Temptation and Self-Control†

Faruk Gul

and

Wolfgang Pesendorfer

Princeton University

October 1999

Abstract

In a two period decision problem, we study individuals who, in the second period, may be tempted by ex ante inferior choices. Individuals have preferences over sets of alternatives that represent the feasible choices in the second period. Our axioms yield a representation that identifies the individual's commitment ranking, her temptation ranking, and her cost of self- control. We provide an axiomatic model of temptation to justify the main assumption of our representation theorem and to analyze second period behavior. An agent has a preference for commitment if she strictly prefers a subset of alternatives to the set itself. An agent has self-control if she resists temptation and chooses an option with higher ex ante utility. We introduce comparative measures of preference for commitment and self-control and relate these measures to our representations.

† Both authors gratefully acknowledge financial support from the National Science Foundation.

1. Introduction

Individuals often choose an alternative deemed inferior ex ante. This "inconsistency", widely documented in experimental settings (see Rabin (1997) for a survey), is attributed to the divergence of preferences reflecting long- run self-interest and behavior motivated by short-run cravings and other visceral factors (Loewenstein (1996)). As an example, consider an individual who must decide what to eat for lunch. She may choose a vegetarian dish or a hamburger. In the morning, when no hunger is felt, she prefers the healthy, vegetarian dish. At lunchtime, the hungry individual experiences a craving for the hamburger.

The decision-maker has two remedies to lessen the conflict between her ex ante ranking of options and her short-run cravings. First, she may engage in activities that limit ex post options. In the extreme case, the individual may be able to commit to her ex ante preferred choice and thereby eliminate all conflict. Second, the individual may exercise *self-control*. Self-control lowers the utility of the individual but allows her to resist the options that are most tempting. Individuals will frequently use both remedies. In the lunch example, the individual may visit a vegetarian restaurant to exclude the hamburger from the option set. However, even the vegetarian restaurant offers unhealthy desserts and self-control may be used to resist that temptation.

To illustrate these ideas more formally, let x denote the vegetarian meal and let y be the hamburger. There are two periods, morning and lunchtime. Consumption takes place at lunchtime when the individual must pick a meal from a menu. In the morning, the individual chooses among menus and hence has preferences over sets of alternatives. The singleton sets $\{x\}$ and $\{y\}$ describe situations where the individual is committed to choose one or the other meal at lunchtime. A situation where the individual must choose between x and y at lunchtime described by $\{x, y\}$.

In the morning, the individual experiences no temptation and ranks x above y. This is captured by a strict preference of $\{x\}$ over $\{y\}$. At lunchtime, the agent may be *tempted* by y. We capture this with a preference that ranks $\{x\}$ strictly above $\{x, y\}$. Thus, temptation creates a preference for commitment. If no commitment is made, then the agent either succumbs to temptation or exercises self-control. In the former case, $\{x, y\}$ is indifferent to the singleton set $\{y\}$ because the menu $\{x, y\}$ leads to the choice y. In the latter case, $\{x, y\}$ is *strictly* preferred to $\{y\}$. This describes a situation where the agent chooses x from the set $\{x, y\}$ but suffers from the availability of y. Self-control at $\{x, y\}$ is therefore identified with $\{x\} \succ \{x, y\} \succ \{y\}$. Self-control enables the agent to choose the vegetarian meal and hence makes her better off than she would be if she had to choose the hamburger, y. On the other hand, the availability of the hamburger creates temptation and therefore, she is worse off than she would be if she were committed to the vegetarian meal, x.

In our model, the agent has preferences over sets of lotteries. Our axioms (described below) allow both a preference for commitment and self-control. We refer to the resulting preferences as *temptation preferences with self- control* or simply *self-control preferences* and show in Theorem 1 that they can be represented by a function U of the form:

$$U(A) := \max_{x \in A} u(x) + v(x) - \max_{y \in A} v(y)$$

Both u and v are von Neumann-Morgenstern utility functions over lotteries. The function v describes the agent's urges in period 2. The function u describes the agents ranking over singleton sets, that is, the ranking when she can avoid temptation through commitment. We refer to v as the agent's *temptation ranking* and u as her *commitment ranking* of lotteries. We interpret $\max_{y \in A} v(y) - v(x)$ as the (utility) cost of self-control. Since this cost is always positive, the presence of temptation always lowers the agent's utility. The representation also suggests a choice behavior in the second period. Choosing a lottery to maximize $u + v$ represents the optimal comprise between the utility that could have been achieved under commitment and the cost of self-control.

Three of the four axioms of Theorem 1 are natural extensions of the standard axioms of expected utility theory: (i) the preferences over sets are complete and transitive, (ii) they satisfy continuity and (iii) the independence axiom. The fourth axiom requires that if A is weakly preferred to B then $A \cup B$ is in fact "between" A and B, i.e., A is weakly preferred to $A \cup B$ which in turn is weakly preferred to B. We call this axiom *Set Betweenness.* A standard decision-maker experiences no temptation and therefore judges sets by their best elements. If such a decision-maker prefers the set A to the set B, it follows that she is indifferent between the sets A and $A \cup B$. In other words, adding the options in the set B to A does not change the best element and therefore does not affect her utility. By contrast, the individual described in this paper may suffer from the availability of additional options. *Set Betweenness* allows for this possibility.

Together the four axioms are equivalent to the representation described above. In our interpretation, temptation may lead to a preference for commitment (i.e., $A \succ A \cup B$). When such a situation arises and commitment is not possible, the agent may either succumb to temptation $A \cup B \sim B$ or use self-control $A \cup B \succ B$. The agent decides which is the best course of action based on the cost of self-control.

In our representation, the cost of self-control is determined by the relative magnitudes of u and v. We also examine the limiting case where self-control is prohibitively costly. In that case, the individual always gives in to her temptation, that is, chooses the lottery that maximizes v but evaluates these choices using u. Such preferences are represented by a utility function of the form,

$$U(A) := \max_{x \in A} u(x) \text{ subject to } \; v(x) \geq v(y) \text{ for all } y \in A$$

We call such behavior *temptation without self-control* or *overwhelming temptation*. To generalize self-control preferences to allow for overwhelming temptation, we weaken the continuity axiom of Theorem 1. In Theorem 2, we show that this weaker continuity axiom together with the remaining assumptions of Theorem 1, are satisfied if and only if preferences can be represented by one of the two representations above.

To this point, we have analyzed and represented period 1 preferences. In section 4 we analyze period 2 behavior. To do this, we extend the individual's preferences to include a choice from the set in period 2. Thus, we assume that the agent has preferences $\succeq^*$ on the set $\mathcal{S} := \{(A, x) : x \in A\}$ where A is the set of lotteries chosen in period 1 and $x \in A$ is the lottery chosen in period 2. As in standard theories of dynamic choice, we assume that period 2 choice maximizes the conditional preference, i.e., the decision- maker chooses $x \in A$ such that $(A, x) \succeq^* (A, y)$ for all $y \in A$. Conversely, $\succeq^*$ induces period 1 preferences over sets denoted by $\succeq_1^*$: $A \succeq_1^* B$ if and only if there is an $x \in A$ with $(A, x) \succeq^* (B, y), \forall y \in B$.

We make the following three assumption on $\succeq^*$. First, making the same choice from a larger set cannot increase the individuals utility ($(A, x) \succeq^* (B, x)$ if $A \subset B$). Second, we say that y tempts x (denoted by $y \succ_T x$) if $(\{x\}, x) \succ^* (\{x, y\}, x)$ and assume that the resulting temptation ranking is complete and transitive. Third, we assume that if $(A, x) \succ^* (A \cup \{y\}, x)$ then y is the most tempting element in $A \cup \{y\}$. This third assumption says that "only the most tempting element in a set matters".

If the extended preference $\succeq^*$ satisfies these three assumptions and a weak form of continuity, then the induced first period preference satisfies *Set Betweenness*. This justifies our interpretation of *Set Betweenness* as capturing an individual who struggles with temptation. If, in addition, the induced preferences $\succeq_1^*$ satisfy the axioms of the representation theorems, then the second period choice behavior is as suggested by the representations: individuals with self-control preferences maximize $u+v$ and individuals with overwhelming temptation preferences lexicographically maximize first v and then u.

Starting with Strotz (1955), the possibility of a preference for commitment has received some attention. This literature assumes that in period 1, the agent has preferences over lotteries that are different from her preferences in period 2. The change of preferences that occurs between period 1 and period 2 is called *dynamic inconsistency.* The benefit of commitment emerges from the first self's desire to "game" her future self.

In the model considered here, the agent's preferences do not change between periods. That is, there is no dynamic inconsistency. Our representation theorems are derived from axioms on the agent's first period preferences over *sets* of lotteries. Our main assumption, *Set Betweenness*, allows us to incorporate temptation and the resulting desire for commitment into a model with dynamically consistent preferences.

Temptation without self-control has the same behavioral implications as Strotz's model of dynamically inconsistent decision-makers. However, Strotz's model allows unambiguous welfare comparisons only when all "selfs" are made better off. Therefore, the elimination of an option can never lead to a clear-cut increase in welfare. By contrast our, dynamically consistent decision-maker is unambiguously better off when ex ante undesirable temptations are no longer available.

As an application of our representation theorems, we develop a measure of preference for commitment and a measure of self-control. We say that $\succeq_1$ has greater preference for commitment than $\succeq_2$, if $\succeq_1$ benefits from commitment whenever $\succeq_2$ benefits from commitment. Our representation theorems assign to each preference relation $\succeq_i$ a pair of utility functions (u_i, v_i). In Theorem 6 we utilize this representation and characterize a greater preference for commitment in terms of the (u_i, v_i). The theorem shows than $\succeq_1$ has a greater preference that $\succeq_2$ if u_1 and v_1 are "further apart" than u_2 and v_2, that is, the indifference for u_2 and v_2 are a convex combination of the indifference curves for u_1 and v_1.

We say that $\succeq_1$ has more self-control than $\succeq_2$ if whenever $\succeq_1$ succumbs to temptation, so does $\succeq_2$. Since overwhelming temptation preferences never exhibit self-control, the measure of self-control applies only to self-control preferences. In Theorem 8 we shows that $\succeq_1$ has more self- control than $\succeq_2$ if the indifference curves for $u_2 + v_2$ and v_2 are a convex combination of the indifference curves for $u_1 + v_1$ and v_1. Recall that individual in period 2 maximizes the utility function $u+v$. If $u+v$ is very different from the temptation ranking v then the agent frequently exercises self-control. Hence, if $u_2 + v_2$ and v_2 are closer together than $u_1 + v_1$ and v_1 then $\succeq_2$ will exercise self-control less frequently than $\succeq_1$.

An individual's preference for commitment increases when the commitment ranking u and the temptation ranking v grow further apart whereas her self-control increases when the utility function defining choice behavior in the second period $u+v$ and the temptation ranking v grow further apart. Hence, it is possible for preference for commitment and self-control to vary independently.

Conceptually, our work relates to the psychology literature on temptation and visceral cues as well as the experimental and theoretical papers on dynamic inconsistency mentioned above. In terms of the formalism, our work is most closely related to the two papers on preference for flexibility. Following Kreps (1979) we study preferences over sets of alternatives. Dekel, Lipman and Rustichini (1999) analyze and extend Kreps' representation of preference for flexibility. The idea of modeling the set of alternatives as lotteries and utilizing the resulting linear structure by imposing the von Neumann-Morgenstern axioms was introduced in an earlier version of their paper.

The paper is organized as follows. In the next section, we provide the model and analyze self-control preferences. In Section 3 we present the general representation theorem that covers temptation with and without self-control and examine the uniqueness of this representation. In Section 4, we demonstrate how our notion of temptation can be used to formulate a dynamically consistent model of preference for commitment. We use this formulation both to analyze second period choice behavior and interpret *Set Betweenness*. Section 5 contains the comparative measures of preference for commitment and self-control. Section 6 relates our approach to preferences for commitment to the one that is offered in the literature on dynamic inconsistency.

2. A Model of Self-Control

We analyze the following two-period decision problem. Consumption only occurs in the second period. At time 2, the agent chooses a consumption (a lottery). Following Kreps (1979) we describe the decision problem at time 1 as the choice of a set of lotteries which constitutes the feasible choices at time 2. Sections 2 and 3 analyze the preferences at time 1. Section 4 extends those preferences to time 2.

Let (Z, d) be a compact metric space, where Z is the set of all prizes. Let Δ denote the set of all measures on the Borel $\sigma-$ algebra of Z. We endow Δ, the set of all lotteries, with the weak topology. Hence, Δ is metrizable. The objects of our analysis are subsets of Δ. Let $\mathcal{A}$ denote the set of compact subsets of Δ. The binary relation $\succeq$ is a subset of $\mathcal{A} \times \mathcal{A}$.[1] We endow $\mathcal{A}$ with the topology generated by the (*Hausdorff*) metric

$$d_h(A, B) = \max\{\max_A \min_B d_p(x, y), \max_B \min_A d_p(x, y)\}$$

where d_p is a metric that generates the weak topology. Define $\alpha A + (1-\alpha)B := \{z = \alpha x + (1-\alpha)y : x \in A, y \in B\}$ for $A, B \subset \Delta, \alpha \in [0, 1]$.

We impose the following axioms:

Axiom 1: *(Preference Relation)* $\succeq$ *is a complete and transitive binary relation.*

Axiom 2: *(Strong Continuity) The sets* $\{B : B \succeq A\}$ *and* $\{B : A \succeq B\}$ *are closed.*

Axiom 3: *(Independence)* $A \succ B$ *and* $\alpha \in (0, 1)$ *implies* $\alpha A + (1-\alpha)C \succ \alpha B + (1-\alpha)C$.

[1] All of the results in this paper would still hold if $\succeq$ were restricted to finite subsets of Δ.

The first two axioms play the same role here as they do in more familiar contexts. To understand the motivation for *Independence* consider an extension of the decision-maker's preferences to the set of lotteries over $\mathcal{A}$. Assume $A \succ B$ and suppose we give the decision-maker the choice between the lottery that yields A with probability α and C with probability $1-\alpha$ (denoted by $\alpha \circ A + (1-\alpha) \circ C$) and the lottery $\alpha \circ B + (1-\alpha) \circ C$. The interpretation is that in either case the randomization occurs prior to the choice in period 2. In period 2, the decision- maker is left with A or C in one case and B or C in the other. In this setting, the standard justification for the independence axiom applies: the fact that C may occur in either case should not interfere with the preference for A versus B and hence $\alpha \circ A + (1-\alpha) \circ C \succeq \alpha \circ B + (1-\alpha) \circ C$. Suppose that the decision-maker satisfies this version of the independence axiom and, in addition, is indifferent as to the timing of the resolution of uncertainty. In that case, the decision-maker is indifferent whether the uncertainty regarding A or C is resolved before or after her choice in period 2. The situation where the uncertainty is resolved after her choice in period 2 is represented by the convex combination $\alpha A + (1-\alpha)C$. Thus, the decision-maker satisfies *Independence* if (1) she satisfies the usual independence axiom and (2) she is indifferent as to when uncertainty is resolved.

A "standard" decision-maker who experiences no temptation and hence has no preference for commitment satisfies Axioms $1-3$ and, in addition, is only interested in the best element of a set. Therefore, such a decision-maker also satisfies the axiom $A \succeq B$ implies $A \sim A \cup B$. It is straightforward to see that this axiom together with Axioms $1-3$ implies that there is a linear utility function u with the property that $A \succeq B$ if and only if $\max_{x \in A} u(x) \geq \max_{x \in B} u(x)$.[2]

By contrast, adding ex ante inferior choices may make matters worse for a decision-maker who experiences temptation since these choices may tempt her in the future. This motivates the following definition.

Definition: *The preference $\succeq$ is has a preference for commitment at A if there exists $B \subset A$ such that $B \succ A$. The preference $\succeq$ has a preference for commitment if $\succeq$ is has a preference for commitment at some $A \in \mathcal{A}$.*

[2] Kreps (1979) observes this result in a finite setting.

The following axiom permits preference for commitment.

Axiom 4: *(Set Betweenness) $A \succeq B$ implies $A \succeq A \cup B \succeq B$.*

Set Betweenness can be understood as an implication of our notion of temptation. An option that is not chosen in period 2 may affect the utility of the decision-maker because it presents a temptation. We assume that a temptation is utility decreasing, that is, an alternative that is not chosen cannot increase the utility of the decision-maker. Furthermore, lotteries can be ranked according to how tempting they are and only the most tempting option available affects the agent's utility.

In section 4 we give a precise statement of these conditions and Theorem 4 demonstrates that they imply *Set Betweenness.* Here, we provide an intuitive explanation of why *Set Betweenness* follows. Consider an agent who is facing the choice set $A \cup B$. Suppose that in period 2 she plans to make the choice x but finds y to be most tempting. Without loss of generality assume $x \in A$. Since alternatives that are not chosen cannot increase the individual's utility, choosing x from A is at least as good as choosing x from $A \cup B$. Since x is a feasible choice in A it follows that $A \succeq A \cup B$. Suppose, without loss of generality, that $y \in B$ and that in the set B the agent would choose z. The lottery y is the most tempting alternative in $A \cup B$. Since the utility cost of temptation only depends on the most tempting alternative choosing z in $A \cup B$ leads to the same utility as choosing z in B and $A \cup B \succeq B$ follows.

One could imagine models of temptation that lead to violations of *Set Betweenness.* For example, a situation where $A \succeq B \succ A \cup B$ may arise if temptation has a cumulative effect so that larger sets are lead to greater temptation. Alternatively, adding options may increase the costs associated with "processing" temptation and hence reduce utility or command a different type of self-control. Finally, temptation may be random so that the agent may end-up with a fifty-fifty gamble between x, y when facing $A = \{x, y\}$ and a fifty- fifty gamble between w, z when facing $B = \{w, z\}$ and a fifty-fifty gamble between $\{x, z\}$ when facing $A \cup B = \{w, x, y, z\}$. Depending on the commitment ranking of the four alternatives, $A \cup B \succ A, A \cup B \succ B$ is plausible as is $A \succ A \cup B, B \succ A \cup B$. More generally, the temptation ranking may fail standard axioms of rational choice such as transitivity. These axioms can be questioned even for agents who display no preference for commitment. The case for imposing these axioms on the description of the agents' temptations is no stronger.

Our objective is to provide a model close enough to the standard model so that the difference in behavior can be attributed solely to the presence of temptation. A narrower definition of temptation is more appropriate for this purpose. We have attempted to rule out any deviation from the standard model that cannot be interpreted as a form of temptation as well as more elaborate formulations of temptation that rely on processing costs, random choice or deviations from the expected utility hypothesis.

We say that the function $U : \mathcal{A} \to I\!R$ represents the preference $\succeq$ when $A \succeq B$ iff $U(A) \geq U(B)$. The function U is linear if $U(\alpha A + (1-\alpha)B) = \alpha U(A) + (1-\alpha)U(B)$ for all $A, B \in \mathcal{A}$. Similarly, the function $u : \Delta \to I\!R$ is linear if $u(\alpha x + (1-\alpha)y) = \alpha u(x) + (1-\alpha)u(y)$ for all $x, y \in \Delta$. Axioms $1-4$ yield the following representation:

Theorem 1: *The binary relation $\succeq$ satisfies Axioms $1-4$ if and only if there are continuous linear functions U, u, v such that $U(A) := \max_{x \in A}\{u(x) + v(x)\} - \max_{y \in A} v(y)$ for all $A \in \mathcal{A}$ and U represents $\succeq$.*

Proof: See Appendix.

It is straightforward to verify that preferences represented by utility functions of the form given in Theorem 1 satisfy Axioms $1-4$. The key steps in the proof of the "only if" part are the following: Lemma 1 shows that under Axioms $1-3$ we can represent the preferences by a continuous, linear utility function U. We define the *commitment* utility as the utility U assigns to singleton sets, i.e., $u(x) = U(\{x\})$. Lemma 1 implies that u is linear, as desired. Lemma 2 says that $U(A) = \max_{x \in A} \min_{y \in A} U(\{x\} \cup \{y\})$ and hence the utility of every set is equal to the utility of subset with at most two elements. Consider any pair (a, b) with $U(\{a\}) > U(\{a, b\}) > U(\{b\})$. Note that if our representation holds then in this case a maximizes $u+v$ and b maximizes v. (If v and $u+v$ had the same maximizer then $U(\{a, b\}) = U(\{a\})$ or $U(\{a, b\}) = U(\{b\})$.) Thus, the change in utility as we vary b must identify v. Formally, fix $\delta > 0$ so that $U(\{a\}) > U(\{a, (1-\delta)b + \delta x\}) > U(\{(1-\delta)b + \delta x\})$ for all x. Then, define

$$v(x) = \frac{1}{\delta}\left(U(\{a, b\}) - U(\{a, (1-\delta)b + \delta x\})\right)$$

Lemma 4 shows that the v so defined is indeed linear, and independent of the particular choice of a and b. Lemma 5 shows that the representation is valid for all two-element sets. Theorem 1 then follows from Lemma 2[3]

When faced with singleton sets, the decision-maker is an expected utility maximizer with utility function u. A singleton set represents the situation where the individual can commit in period 1 to a consumption choice in period 2. Therefore, we say that u represents the commitment ranking of lotteries and refer to $u(x)$ as the commitment utility of the choice x.

The decision-maker's preferences are defined over sets and thus characterize choice behavior only in the first period. Nevertheless, the representation suggests the following choice behavior in the second period: the agent chooses an element in A that maximizes $u + v$. In Section 4 we provide a model of second period preferences and give conditions so that maximizing $u + v$ is indeed the choice behavior in the second period.

We interpret v as representing the temptation ranking and $v(x) - \max_{y \in A} v(y)$ as the disutility of self-control. To motivate this interpretation, consider x, y with $\{x\} \succ \{x, y\}$. In this case, y is tempting the decision-maker and, in particular, y is more tempting than x. But $\{x\} \succ \{x, y\}$ only if $v(y) > v(x)$. This suggests that v represents the temptation ranking. Section 4 analyzes second period preferences and gives a precise model of temptation. For that model, Theorem 6 shows that v indeed represents the temptation ranking. The term $v(x) - \max_{y \in A} v(y)$ is the difference in temptation-utility between the actual choice and the most tempting option available. We therefore interpret it as the disutility of self-control. The compromise between the commitment ranking and the temptation ranking is represented by $u + v$ which determines the individual's choice.

Recall from our discussion of *Set Betweenness* that if x – the planned choice from the set $A \cup B$ – is in A then we have $A \succeq A \cup B$ and if the most tempting alternative, y, is in B we have $A \cup B \succeq B$. For the first preference to be strict it must be that $y \notin A$. For the second preference to be strict it must be that $x \notin B$. Hence, $A \succ A \cup B$ captures the fact that B entails greater temptation than A while $A \cup B \succ B$ captures the fact that the agent resists this temptation. That is, the agent uses self-control.

[3] Dekel, Lipman and Rustichini (1999) consider finite Z and characterize binary relations that satisfy Axioms $1-3$. For finite Z, their Theorem and our Lemma 2 may be used to construct an alternative proof of Theorem 1.

Definition: *The preference $\succeq$ has self-control at C if there exists A, B such that $C = A \cup B$ and $A \succ A \cup B \succ B$. The preference $\succeq$ has self-control if $\succeq$ has self-control at some $C \in \mathcal{A}$.*

Theorem 7 proves that a preference relation has self-control at A if and only if the second-period choice does not maximize the temptation ranking. This characterization of self-control is consistent with the everyday meaning of the term but is different from the one commonly used in the literature. In most studies, what is referred to as self-control is synonymous with commitment. In our model, the benefit of commitment arises from the possibility that an agent may strictly prefer $\{x\}$ to $\{x, y\}$. Such an agent, given the option, would choose to commit to x in period 1. If committing to $\{x\}$ is not possible at time 1, then at time 2, the agent may exercise self-control and choose x from the set $\{x, y\}$, in spite of the fact that he prefers y to x (that is, $v(y) > v(x)$). Thus, commitment refers to behavior at time 1, the planning period, while self-control refers to behavior at time 2, the consumption period.

3. Temptation with and without Self-Control

In this section, we generalize the preferences studied in the previous section to include decision-makers who cannot resist temptation. At time 2, such decision-makers choose according to some preference relation. This choice behavior is anticipated at time 1 and preferences over choice sets reflect the predicted behavior at time 2.

Recall that for preferences that satisfy Axioms $1-4$, the cost of self-control depends on the magnitude of v in comparison to u. If we multiply v by a large positive constant while holding u fixed then it becomes very costly to make choices that are in conflict with v. Thus, the choices made by the individual are ε-optimal choices for the utility function v. We can think of this case as approximating overwhelming temptation: The agent evaluates her options using the utility function u but makes choices that are optimal for the utility function v. Note, however, that there is an important difference: the decision-maker represented in Theorem 1 satisfies *Strong Continuity* whereas temptation preferences without self-control do not. To see this, consider a set A such that $v(x) = v(y)$ for all $x, y \in A$. Then, all choices are equally tempting and both the decision-makers with or without self- control would choose the best alternative for the utility function u from the set A. However, if the set A is slightly perturbed then the decision-maker with self-control would still choose an alternative in A that is close to maximizing u whereas a decision-maker without self-control would maximize v. We therefore must weaken continuity to include decision makers without self- control.

Axiom 2a: *(Upper Semi-Continuity) The sets $\{B \in \mathcal{A} : B \succeq A\}$ are closed.*

Axiom 2b: *(Lower von Neumann-Morgenstern Continuity) $A \succ B \succ C$ implies $\alpha A + (1-\alpha)C \succ B$ for some $\alpha \in (0,1)$.*

Axioms 2c: *(Lower Singleton Continuity) The sets $\{x : \{y\} \succeq \{x\}\}$ are all closed.*

Axioms 2a−c together are weaker than Axiom 2 presented in Section 2. *Strong Continuity* is equivalent to the *Upper Semi-Continuity* and a symmetric lower semi-continuity condition. The latter is strictly stronger than Axioms 2b and 2c.

Axiom 2a implies that if the most tempting alternative is not unique, then an individual with no self-control resolves this indifference to maximize the commitment ranking. This is the natural assumption when we interpret overwhelming temptation as a prohibitively high cost of self-control. To see why Axiom 2a implies this tie-braking rule, consider the decision-maker's preferences over the collections of sets $\{x, y_k\}$ where y_k converges to y. Suppose $\{x\}$ is strictly preferred to $\{y_k\}$ and x more tempting than y_k for all k. Furthermore, suppose that x and y are equally tempting while $\{x\}$ is strictly preferred to $\{y\}$. Hence, $\{x, y_k\} \succeq \{x\}$ for all k. Then, Axiom 2a requires that $\{x, y\} \succeq \{x\}$. This reflects the expectation that the ex ante preferred alternative x will be chosen over y whenever x and y are equally tempting. Axiom 2b is "half" of the familiar von Neumann-Morgenstern continuity axiom and together with Axiom 2a implies von Neumann-Morgenstern continuity. Adding Axiom 2c ensures that u and v are continuous. Axioms 2c can be omitted if the representation is restricted all finite subsets of Δ.

Theorem 2: *The binary relation $\succeq$ satisfies Axioms $1, 2a-c, 3, 4$ if and only if there a linear functions U and continuous linear functions u, v such that either*

$$U(A) := \max_{x \in A}\{u(x) + v(x)\} - \max_{y \in A} v(y) \quad \textit{for all } A \in \mathcal{A} \quad \textit{or}$$

$$U(A) := \max_{x \in A} u(x) \ \textit{subject to} \ \ v(x) \geq v(y) \ \textit{for all } y \in A$$

and U represents $\succeq$.

The proof of Theorem 2 is in the appendix and somewhat indirect. Here, we briefly describe the main steps. First, we consider a preference with a "self- control-pair", i.e., $\{x\} \succ \{x, y\} \succ \{y\}$. In this case, Claim 1 shows that for finite sets Theorem 1 continues to hold - even under the weaker continuity assumption. Second, we consider a preference with no self-control pair, i.e., $\{x\} \sim \{x, y\}$ or $\{y\} \sim \{x, y\}$ for all x, y. We show in Claims 2 and 3 that this implies that for any finite set A, B, $A \sim A \cup B$ or $B \sim A \cup B$. We then define the temptation ranking R for the no-self- control case as follows: Suppose $\{x\} \not\sim \{y\}$. Then xRy if $\{x, y\} \sim \{x\}$, that is, "x is more tempting than y" if $\{x, y\}$ is indifferent to $\{x\}$. Suppose $\{x\} \sim \{y\}$. Then xRy if $\{z, y\} \sim \{y\}$ implies $\{z, x\} \sim \{x\}$, that is, x is more tempting than y if x tempts z whenever y tempts z. Lemmas 6 and 7 show that R can be represented by a linear utility function v. This in turn allows us to get the representation of no self-control preferences for finite sets. The final step is to extend the representations to all elements of $\mathcal{A}$. This is done in Lemma 8 using Axioms 2a and 4.

We call the set of preferences characterized by Theorem 2 *temptation preferences.* We say that (u, v) represents $\succeq$ if u, v are both continuous linear functions and the temptation preference $\succeq$ can be represented in either of the two ways described in Theorem 2. Temptation preferences that have the first type of representation described in Theorem 2 are called *temptation preferences with self-control* or simply *self-control preferences.* Preferences that have the second type of representation are called *temptation preferences without self-control* or *overwhelming temptation preferences.* Overwhelming temptation preferences can be isolated from the general class of preferences characterized in Theorem 2 by imposing no self-control.

A decision-maker has self-control at a set C if $A \succ A \cup B \succ B$ for some A, B such that $C = A \cup B$. Hence, a decision maker has *no self-control* if there exists no A, B such that $A \succ A \cup B \succ B$. Set Betweenness and no-self control are equivalent to the condition

$$A \succeq B \Rightarrow A \sim A \cup B \text{ or } B \sim A \cup B$$

Thus, if $A \succ A \cup B$ and hence adding B to A provides a temptation for the decision-maker, it follows that $B \sim A \cup B$. That is, the individual succumbs whenever a temptation presents itself. The proof of the following corollary is straightforward and therefore omitted.

Corollary 1: *The temptation preference* $\succeq$ *with representation* (u, v) *has no self-control if and only the function*

$$U(A) := \max_{x \in A} u(x) \text{ subject to } v(x) \geq v(y) \text{ for all } y \in A$$

for all $A \in \mathcal{A}$ *represents* $\succeq$*.*

Both self-control preferences and overwhelming temptation preferences include standard preferences with no preference for commitment. It is easy to verify that if (u, v) represents $\succeq$ then $\succeq$ has no preference for commitment if and only if u is constant or v is constant or u is an positive affine transformation of v (i.e. $u = \alpha v + \beta$ for some $\alpha > 0, \beta \in I\!R$). At the other extreme, there are preferences that have *maximal preference for commitment*; that is, either $\{x\} \sim A$ for all $x \in A$ or $\succeq$ has a preference for commitment at A. It is also easy to verify that a preference with representation (u, v) has maximal preference for commitment if and only if $u = \alpha v + \beta$ for some $\alpha \leq 0, \beta \in I\!R$. We say that a preference relation is *regular* if it has some preference for commitment but does not have maximal preference for commitment. Hence, if (u, v) represents $\succeq$ then $\succeq$ is regular if and only if neither u nor v is constant and v is not an affine transformation of u. Note that a regular preference with representation (u, v) is either a self-control preference or an overwhelming temptation preference but not both.

Standard arguments familiar from von Neumann-Morgenstern utility theory can be used to verify that the U representing the preferences in Theorem 3 above is unique up to a positive affine transformation. Theorem 3 below establishes a stronger result for regular preferences: the pair (u, v) representing a regular, self-control preference is unique up to a common, positive affine transformation. If (u, v) represents a regular, temptation preference without self-control then the u and v are both unique up to (possibly different) positive affine transformations.

Theorem 3: *Suppose* (u, v) *represents the regular preference relation* $\succeq$*. If* $\succeq$ *has self-control then* (u', v') *also represents* $\succeq$ *if and only if* $u' = \alpha u + \beta_u$ *and* $v' = \alpha v + \beta_v$ *for some* $\alpha > 0$ *and* β_u*,* $\beta_v \in I\!R$*. If* $\succeq$ *has no self-control then* (u', v') *represents* $\succeq$ *if and only if* $u' = \alpha_u u + \beta_u$ *and* $v' = \alpha_v v + \beta_v$ *for* α_u*,* $\alpha_v > 0$ *and* β_u*,* $\beta_v \in I\!R$*.*

4. Temptation, Set Betweenness and Second Period Choice

This section analyzes behavior in the second period. To motivate our approach, consider a different, but more familiar two-stage decision problem. Suppose in period 1, the agent makes a costly investment decision I and in period 2, she makes a consumption choice c. The agent's overall utility V, depends both on her investment I and her consumption c. The investment decision has three potential consequences. First, it affects the set of feasible choices in period 2. Second, it has a direct effect on utility since there is a cost to investment. That is, for a fixed c, $V(I, c)$ varies as I varies. Finally, the investment decision may influence how the agent ranks various alternatives in period 2. When the agent compares two different investment choices in the first period, she assigns to each investment the maximal utility she can achieve in the second period. In this way, V induces preferences over investment choices and the agent's preferences are *consistent* across time periods.

In our model of temptation an individual is in a situation similar to the one facing the investor above. The set of lotteries, A, chosen in the first period, is analogous to the investment choice. It determines which consumption choices are feasible in the second period. Like the investment decision, the choice of A has a direct effect on overall utility through its impact on the cost of self-control. In addition, the impact of temptation on the decision may depend on the set of options and hence the decision-maker's ranking over alternatives in period 2 may depend directly on A.[4]

Extended preferences are defined over pairs (A, x) where x is a choice from A. Let $\succeq^*$ be a preference relation on $\mathcal{S} := \{(A, x) \in \mathcal{A} \times \Delta : x \in A\}$. We assume that $\succeq^*$ is upper semi-continuous preference relation.

Axiom 1^***:** $\succeq^*$ *is a preference relation.*

Axiom 2^*a**:** *(Upper Semi-Continuity)* $\{(A, x) : (A, x) \succeq^* (B, y)\}$ *is closed for all* (B, y)

Enlarging the domain of preferences from sets of lotteries, $\mathcal{A}$, to pairs in $\mathcal{S}$ permits us to describe the second period behavior of an individual who has a preference for commitment. Just as the "investor" above evaluates consumption in the second period by the utility

[4] At the end of this section, we give an example that shows how such a direct effect of the set A on the ranking of alternatives may emerge in our model of temptation.

function $V(I,\cdot)$, our decision-maker ranks lotteries according to conditional preference induced by her choice A in the first period. Changing first period choice from B to A may change the agent's ranking of x and y, $x, y \in A \cap B$. This change does not reflect a preferences reversal once the effect of the choice set on the extended preference is accounted for. In the first period, the agent evaluates sets anticipating her optimal choice in the second period. For an upper semi-continuous extended preference $\succeq^*$ we define the induced first period preference, denoted by $\succeq_1^*$, as follows: $A \succeq_1^* B$ if and only if there exists $x \in A$ such that $(A, x) \succeq^* (B, y)$ for all $y \in B$.[5]

For example, suppose that the extended preferences $\succeq^*$ can be represented by a utility function

$$U^*(A, x) = u(x) + v(x) - \max_{y \in A} v(y)$$

In period 2, the dynamically consistent agent with utility function U^* chooses $x \in A$ that maximizes $U^*(A, \cdot)$ and therefore, in period 1, this individual evaluates the set A according to $\max_{x \in A}\{u(x) + v(x)\} + \max_{y \in A} v(x)$. Hence, such an agent's induced first period preference over sets, $\succeq_1^*$ is the self-control preference represented by (u, v). Alternatively, define $m := \min_{x \in A} u(x)$ and suppose $\succeq^*$ can be represented by

$$U^*(A, x) = \begin{cases} u(x) & \text{if } v(x) \geq \max_{y \in A} v(y) \\ m - 1 & \text{otherwise} \end{cases}$$

In period 2, an agent with this utility function chooses some $x \in A$ that lexicographically maximizes first v and then u. Hence, in this case $\succeq_1^*$ is the overwhelming temptation preference (u, v).

Clearly, the preference $\succeq^*$ represented by the U^* described above is not the only extended preference that induces a given temptation preference (u, v). A particular temptation preference can be extended to a preference relation on $\mathcal{S}$ in many ways. Moreover, the testable implications of any particular extension $\succeq^*$ are captured by the induced preferences over $\mathcal{A}$ and the associated choice function. For example, if (A, x) is strictly preferred to (A, y) and (A, z) there is no experiment that can determine the agent's ranking of (A, y)

[5] Machina (1989), observes that the dynamic inconsistency associated with non-expected utility preferences can be "resolved" if one re- defines preferences over a suitably large space. We are implementing the same idea in a somewhat different setting.

and (A, z). After all, (A, y) and (A, z) refer to hypothetical events that the agent never experiences. The preference $\succeq^*$ is useful as it allows us to flesh out our theory of temptation and to relate first period preferences over sets of options to second period choices from these options. But, $\succeq^*$ cannot be identified by revealed preference experiments. For this reason, we use preferences on $\mathcal{A}$ in our representation theorems.

The temptation axioms below impose restrictions on how first period choices influence the ranking of alternatives according to $\succeq^*$. The first axiom states that adding new options does not increase the utility of any option that was already available. That is, new alternative may cause temptation but cannot enhance old ones.

T1: $(A, x) \succeq^* (B, x)$ *whenever* $A \subset B$

A lottery y tempts x if introducing the choice y makes x less attractive than choosing x when only x is available. Formally, we say y *tempts* x iff $(\{x\}, x) \succ^* (\{x, y\}, x)$.

T2: $(\{x\}, x) \succ^* (\{x, y\}, x)$ *implies* $(\{y\}, y) \sim^* (\{x, y\}, y)$

T3: $(A, x) \succ^* (A \cup \{y\}, x)$ *implies* $(\{z\}, z) \succ^* (\{y, z\}, z)\ \ \forall z \in A$

The second axiom says that if y tempts x then x does not tempt y. The last axiom states that if the addition of an option y increases the cost of self-control then y must be the most tempting alternative among all available options. In short: only the most tempting option matters. These axioms correspond to the properties of temptation discussed informally in section 2.

In addition, we require $\succeq^*$ to be singleton continuous.

Axioms 2^*c: *(Lower Singleton Continuity) The sets* $\{x : (\{y\}, y) \succeq (\{x\}, x)\}$ *are all closed.*

Note that $\succeq^*$ satisfies Axiom 2^*c if and only if $\succeq^*_1$ satisfies Axiom 2c above. The next result shows that the temptation axioms imply *Set Betweenness.*

Theorem 4: *If* $\succeq^*$ *satisfies* 2^*a,2^*c,*T1−3 then* $\succeq^*_1$ *satisfies Set Betweenness.*

Proof: See Appendix.

To see the intuition for Theorem 4, consider an agent who is facing the choice set $A \cup B$. Suppose that in the set $A \cup B$ the agent chooses x but finds y to be most tempting.

Without loss of generality assume $x \in A$. By Axiom T1, $(A, x) \succeq^* (A \cup B, x)$. Since x is a feasible choice in A it follows that $A \succeq_1^* A \cup B$. Suppose, without loss of generality, that $y \in B$ and that in the set B the agent would choose z. Since lottery y is the most tempting alternative in $A \cup B$ it follows from Axiom T3 that choosing z in $A \cup B$ leads to the same utility as choosing z in B and $A \cup B \succeq B$ follows.

Theorem 5 demonstrates that if an (extended) preference $\succeq^*$ satisfies 2^*a,2^*c and the temptation axioms then y tempts x iff $v(y) > v(x)$. This confirms our earlier interpretation of v as representing the temptation ranking. For this result to hold, v must be unique. Therefore, we require the induced preference $\succeq_1^*$ to be regular, that is, u and v must be linearly independent.

Theorem 5: *Suppose $\succeq^*$ satisfies 2^*a, T 1–3 and $\succeq_1^*$ is a regular temptation preference. If $\succeq_1^*$ can be represented by (u, v) then y tempts x iff $v(y) > v(x)$.*

Proof: See Appendix

To see the intuition for Theorem 5, consider the case where $u(x) > u(y)$. Then, $\{x\}$ is strictly preferred to $\{x, y\}$, that is the decision-maker has a preference for commitment at $\{x, y\}$, if and only if $v(x) < v(y)$. This follows from our representation theorems. On the other hand, $T1$ and the fact that $\succeq_1^*$ is represented by (u, v) imply $(\{x\}, x) \succ^* (\{y\}, y) \succeq^* (\{x, y\}, y)$. We conclude that $\{x\}$ is strictly preferred to $\{x, y\}$ (according to $\succeq_1^*$) if and only if $(\{x\}, x) \succ^* \{x, y\}, y)$. Hence, $v(y) > v(x)$ if and only if y tempts x.

Our next objective is to characterize choice behavior associated with temptation preferences.

Definition: *For any upper semi-continuous function $f : \mathcal{A} \to I\!R$ define the choice function $c(\cdot, f)$ as follows: $c(A, f) := \{x \in A : f(x) \geq f(y) \, \forall y \in A\}$. Similarly, let $c^*(A, \succeq^*) := \{x \in A : (A, x) \succeq^* (A, y) \, \forall y \in A\}$.*

In Theorem 6 we show that if a temptation preference is extended in a manner that satisfies the temptation axioms, then the induced second period choice behavior is as suggested by our representation theorems.

Below we assume that $\succeq^*$ is minimally congruent. Minimal congruence requires that a worst element for the commitment ranking, u, is not tempting. For example, if the set of prizes Z are quantities of "goods" and both u and v satisfy first order stochastic dominance then minimal congruence holds.

Definition: *$\succeq^*$ is minimally congruent if $A \succeq \{x\}$ for all A implies $(\{x, y\}, x) \sim^* (\{x\}, x)$ for all $x \in \Delta$.*

Theorem 6: *Suppose $\succeq^*$ satisfies 2^*a, $T1-3$ and is minimally congruent. Suppose also that $\succeq_1^*$ is a temptation preference with representation (u, v) where u not constant. Then, either $\succeq_1^*$ is a self-control preference and*

$$c^*(A, \succeq^*) = c(A, u + v)$$

for all $A \in \mathcal{A}$, or $\succeq$ has no self-control and

$$c^*(A, \succeq^*) = c(c(A, v), u)$$

for all $A \in \mathcal{A}$.

Proof: See Appendix.

To gain intuition for Theorem 6 suppose $\succeq_1^*$ has self control. Consider a preference (u, v) and a set A with the property that x^* is the unique maximizer of $u + v$ and y^* is the unique maximizer of v with $x^* \neq y^*$. Clearly, $c^*(A, \succeq^*)$ is non-empty and hence we must show that if $\hat{x} \in A$ is not equal to x^* then $\hat{x} \notin c^*(A, \succeq^*)$. Let $\hat{x} \notin c(A, u + v)$ and note that $A \succeq_1^* \{hatx, y^*\}$ since (by construction)

$$u(x^*) + v(x^*) - v(y^*) > \max\{u(y^*), u(\hat{x}) + v(\hat{x}) - v(y^*)\}$$

But since $\succeq_1^*$ is induced by $\succeq^*$ this implies that there is a $z \in A$ such that $(A, z) \succ^* (\{\hat{x}, y^*\}, \hat{x})$. Axiom $T1$ implies that $(\{\hat{x}, y^*\}, \hat{x}) \succeq^* (A, \hat{x})$ and therefore $\hat{x} \notin c^*(A, \succeq^*)$. Continuity and the remaining temptation axioms are used to show that all $u+v$ maximizers are in $c^*(\cdot, \succeq^*)$.

If $\succeq^*$ is not minimally congruent then Theorem 6 still holds for an open and dense subset of $\mathcal{A}$. For any $\succeq^*$, let Δ^o denote all $x \in \Delta$ such that $\{x\} \succ A$ for some $A \in \mathcal{A}$. Thus, Δ^o contains all lotteries x such that x is not a worst element of the commitment ranking u.

If $\succeq^*$ is not minimally congruent then Theorem 6 holds for every set $A \in \mathcal{A}$ with $A \subset \Delta^o$. Under the assumptions of Theorem 5 it can be verified that collection of sets contains an open and dense subset of $\mathcal{A}$ whenever v is not a negative affine transformation of u. Thus, even if minimal congruence does not hold, the choice behavior induced by $\succeq^*$ almost always agrees with that suggested by our representation theorems. We therefore refer to $c(\cdot, u+v)$ as the second period choice function associated with $\succeq$ whenever $\succeq$ is a self-control preference. Similarly, $c(c(\cdot, v), u)$ is the second period choice function associated with the $\succeq$ whenever $\succeq$ has no self-control.

For a self-control preference (u, v), the associated second period choice maximizes $u + v$. Similarly, for overwhelming temptation preferences the second period choice lexicographically maximizes first v and then u. Hence, in the terminology of our investment example the third potential effect of the "investment" A is absent: the ranking over alternatives induced by $\succeq^*$ is independent of A. This feature of second period choice behavior does not follow from weak continuity and the temptation axioms but from *Independence* of the induced preference. Below, we give an example of a utility function that leads to choice behavior that cannot be represented as maximization of a set-independent utility function even though it satisfies Axioms 2^*a,2^*c, $T1-3$.

Consider the preferences represented by

$$U^*(A, x) = u(x) + v_2(x)v_1(x) - v_2(x) \cdot \max_{y \in A} v_1(y)$$

where $v_1 > 0, v_2 > 0, u$ are linear functions. Intuitively, U^* associates three characteristics with each lottery: its commitment ranking u, its temptation ranking v_1 and its susceptibility to temptation v_2. It is easy to verify that U^* is continuous and satisfies the temptation axioms. Hence, the induced preference $\succeq^*_1$ is a continuous preference relation that, by Theorem 4, satisfies *Set Betweenness*. However, for suitable choices of v_1 and v_2 $\succeq^*_1$ does not satisfy *Independence* and the ranking of alternatives depends on the set A. Consequently, an agent with these extended preferences may violate Houthakker's Axiom and choose only x from the set A and only y from some set B even though $x, y \in A \cap B$.

5. Measures of Preference for Commitment and Self-Control

In this section, we provide measures of preference for commitment and of self- control.

Definition: *The preference* $\succeq_1$ *has greater preference for commitment than* $\succeq_2$ *if, for all* $A \in \mathcal{A}$, $\succeq_2$ *has preference for commitment at* A *implies* $\succeq_1$ *has preference for commitment at* A.

Theorem 7 characterizes this comparative measure of preference for commitment in terms of the representation (u, v). For this characterization, we require the preferences $\succeq_i, i = 1, 2$ to be regular. Recall that $\succeq$ with representation (u, v) is regular if and only if neither u nor v is constant and v is not an affine transformation of u.

Theorem 7: *Let* $\succeq_1$, $\succeq_2$ *be two regular temptation preferences and let* (u_1, v_1) *be a representation of* $\succeq_1$. *Then,* $\succeq_1$ *has greater preference for commitment than* $\succeq_2$ *if and only if there exists,* u_2, v_2 *such that* $(u_2, \gamma v_2)$ *represents* $\succeq_2$ *and*

$$u_2 = \alpha u_1 + (1 - \alpha) v_1$$

$$v_2 = \beta u_1 + (1 - \beta) v_1$$

for some $\alpha, \beta \in [0, 1]$ *and some* $\gamma > 0$

Proof: See Appendix.

Consider the indifference curves of u_1, v_1, u_2 and v_2 through a lottery $x \in \Delta$. The theorem says that $\succeq_1$ has more preference for commitment than $\succeq_2$ if and only if the indifference curve of u_2 and v_2 are each a convex combination of the indifference curves of u_1 and v_1. Therefore $\succeq_1$ has greater preference for commitment than $\succeq_2$ if and only if the commitment ranking and temptation ranking associated with $\succeq_2$ are closer together than the commitment ranking and temptation ranking associated with $\succeq_1$. This is to be expected since the benefit of commitment arises from the discrepancy between these two rankings.

To see the sufficiency part of Theorem 7, suppose that (u_2, v_2) has no preference for commitment at A. This means that u_1 and v_1 have a common maximizer x in A. Clearly, the same maximizer continues to work for (u_2, v_2) since each is a positive linear combinations of two utility functions that pick $x \in A$. Thus, whenever $\succeq_1$ has no preference for commitment neither does $\succeq_2$. For the converse we prove that whenever (u_1, v_1) and (u_2, v_2) are not on the same two-dimensional plane we can find a pair of lotteries (x, y) such that u_2 or v_2 is indifferent between x and y but the utility functions u_1 and v_1 rank x and y differently. Clearly, this implies that $\succeq_1$ is dynamically consistent but $\succeq_2$ is not.

Before providing an analogous measure for self-control, we provide a characterization theorem. Three notions of self-control are proven to be equivalent: (i) our earlier definition based on preferences over sets; (ii) an intuitive definition of self-control as an individual's ability to resist temptation; (iii) a revealed preference definition based on the observation that an agent with self-control might prefer set A to B even when the same choice is being made from both sets. Recall that $c(\cdot, u+v)$ is the second period choice behavior corresponding to the self-control preference $\succeq$.

Theorem 8: *Let (u, v) represent the self-control preference $\succeq$. Then, the following three statements are equivalent:*

(i) $\succeq$ has self-control at A

(ii) $c(A, v) \cap c(A, u+v) = \emptyset$

(iii) There exists $B \subset A$ such that $c(A, u+v) = c(B, u+v)$ and $B \succ A$

Proof: To prove (i) implies (ii) let $B \cup C = A$ and $x \in c(A, v) \cap c(A, u+v)$. Clearly, $U(A) = u(x) = U(D)$ if $x \in D \subset A$. So, $B \sim A$ or $C \sim A$ proving that $\succeq$ has no self-control at A. To prove (ii) implies (iii) assume $c(A, v) \cap c(A, u+v) = \emptyset$. Let $B = c(A, u+v)$. Note that $c(B, u+v) = c(A, u+v)$ and $B \succ A$. To prove (iii) implies (i) assume $c(A, u+v) = c(B, u+v)$ and $B \succ A$. Then, let $\beta := \max_{x \in B} v(x)$ and $\alpha := \max_{x \in A} v(x)$. Since $B \succ A$, $\beta < \alpha$. Let $B' := \{x \in A : v(x) \leq (\alpha+\beta)/2\}$ and $C := \{x \in A : v(x) \geq (\alpha+\beta)/2\}$. Then $B' \cup C = A$ and $B' \succ B' \cup C \succ C$. □

Definition: *The preference $\succeq_1$ has more self-control than $\succeq_2$ if, for all $A \in \mathcal{A}$, $\succeq_2$ has self-control at A implies $\succeq_1$ has self-control at A.*

As one might expect, overwhelming temptation corresponds to the minimal level of self-control. Therefore, when characterizing the level of self-control we restrict attention to self-control preferences. Theorem 9 provides a characterization of comparative self-control similar to the characterization of comparative preference for commitment in Theorem 7.

Theorem 9: *Let $\succeq_1$, $\succeq_2$ be two regular self-control preferences and let (u_1, v_1) be a representation of $\succeq_1$. Then, $\succeq_1$ has more self-control than $\succeq_2$ if and only if there exist (u_2, v_2) such that $(u_2, \gamma v_2)$ represents $\succeq_2$ and*

$$u_2 + v_2 = \alpha(u_1 + v_1) + (1 - \alpha)v_1$$
$$v_2 = \beta(u_1 + v_1) + (1 - \beta)v_1$$

for some $\alpha, \beta \in [0, 1], \gamma > 0$.

Proof: Given the characterization of self-control provided in (ii) of Theorem 8, the proof is analogous to the proof of Theorem 7 and therefore omitted.

Consider the indifference curves of $u_1 + v_1$, v_1, $u_2 + v_2$ and v_2 through a lottery $x \in \Delta$. The theorem says that $\succeq_1$ has more self-control than $\succeq_2$ if and only if the indifference curve of $u_2 + v_2$ and v_2 are each a convex combination of the indifference curves of $u_1 + v_1$ and v_1. Theorem 8 therefore establishes that $\succeq_1$ has more self-control than $\succeq_2$ if and only if the preference that describes choice behavior $u + v$ is closer to the temptation ranking for $\succeq_2$ than it is for $\succeq_1$. This is to be expected since self-control refers to the conflict between the agent's second period behavior and temptation.

Recall that for $\succeq_1$ to have greater preference for commitment than $\succeq_2$, (u_2, v_2) must be closer together than (u_1, u_1). In contrast, for $\succeq_1$ to have more self-control than $\succeq_2$, $(u_2 + v_2, v_2)$ must be closer together than $(u_1 + v_1, v_1)$. Hence, it is possible for $\succeq_1$ to be greater preference for commitment than $\succeq_2$ and yet have more self-control than $\succeq_2$.

Let $\succeq_1$ and $\succeq_2$ be two regular self-control preferences. Suppose $\succeq_1$ is represented by (u_1, v_1) and $\succeq_2$ is represented by $(u_1 + \beta v_1, \gamma v_1)$. We can distinguish the following cases:

1. $\beta > 0, \gamma = 1$: In this case $\succeq_1$ has greater preference for commitment than and more self- control than $\succeq_2$.
2. $\beta \in (0, 1), \gamma = (1 - \beta)$: In this case $\succeq_1$ has greater preference for commitment than $\succeq_2$ but both have the same level of self-control.
3. $\beta = 0, \gamma \in (0, 1)$: In this case both have the same level of preference for commitment but $\succeq_2$ has more self-control than $\succeq_1$.
4. $\beta > 0$, $0 < \gamma < 1 - \beta$: In this case $\succeq_1$ has greater preference for commitment than $\succeq_2$ and $\succeq_2$ has more self-control than $\succeq_1$.

6. Temptation versus Change in Preference

In this paper we offer two related conceptual innovations. First, we propose that *temptation* rather than a preference change ("dynamic inconsistency") may be the cause of a preference for commitment. Second, we introduce a model in which agents resist temptation, that is, use self-control. The representation of self-control preferences allows us quantify the cost of self-control as a utility penalty that applies whenever the ultimate choice is not the most tempting one. Our model enables us to distinguish between self-control, which occurs at time of consumption, and commitment, which takes place earlier.

Our model of self-control yields both different behavioral and normative implications than the change in preference approach. An agent with self-control may be worse off when an irrelevant alternative is added to her set of options. That is, we may have $A \succ A \cup \{x\}$ even though x is ultimately not chosen from the set $A \cup \{x\}$. Put differently, removing a non-binding constraint may reduce an agent's utility. Hence, unlike dynamically inconsistent agents or agents with overwhelming temptation, decision-makers with self-control will expend resources to remove tempting alternatives from their choice sets even if they do not expect to succumb to the temptation in the future. To see why this is relevant, consider a representative agent model of an asset market where one asset offers commitment. In a model of changing tastes, the illiquid asset only benefits the agent if its purchase leads to a binding constraint. In other words, the representative agent must hold zero units of the liquid asset in some state of the world to generate a premium for an asset offering commitment. This is in contrast to a model of self-control where such a premium may exist even if the agent is never constrained.

Our model of temptation without self-control leads to the same testable implications as Strotz's model of dynamic inconsistency. Still, the two approaches are not equivalent. A model with dynamic inconsistency rarely leads to clear-cut welfare analysis. When each agent has multiple-selves, the impact of a given policy on a single agent is typically ambiguous. By contrast, removing temptation makes our, dynamically consistent, agents unambiguously better off.

7. Appendix

It is straightforward to verify that preferences represented by utility functions of the form given in Theorem 1 satisfy the Axioms $1-4$. Verifying the "if" part of Theorem 2 is equally straightforward and hence omitted.

7.1 Proof of Theorem 1

Lemma 1: *If Axioms* $1, 2a, 2b$ *and* 3 *hold, then there is a linear function* $U : \mathcal{A} \to R$ *that represents* $\succeq$*. The restriction of* U *to singleton sets is continuous. If, in addition,* U *satisfies Axiom 2, then* U *is continuous on* $\mathcal{A}$*.*

Proof: Let $L(\mathcal{A})$ denote the set of lotteries over $\mathcal{A}$ with finite support. We sometimes use A to denote the degenerate lottery with prize A. Define the following preferences: $\pi \,\dot{\succeq}\, \rho$, $\pi, \rho \in L(\mathcal{A})$ whenever

$$\sum_A \pi(A)A \succeq \sum_A \rho(A)A$$

Note that $L(\mathcal{A})$ is a mixture space. Moreover, Axiom 2a,2b imply that $\dot{\succeq}$ satisfies von Neumann-Morgenstern continuity. That is, for all $\pi, \rho, \mu \in L(\mathcal{A})$ with $\pi \,\dot{\succ}\, \rho \,\dot{\succ}\, \mu$ there are $\alpha, \beta \in (0, 1)$ such that $\alpha\pi + (1-\alpha)\mu \,\dot{\succ}\, \rho \,\dot{\succ}\, \beta\pi + (1-\beta)\mu$. Finally, Axiom 3 implies that $\dot{\succeq}$ satisfies the independence axiom. Therefore, there is a function $W : \mathcal{A} \to I\!R$ such that

$$\pi \,\dot{\succeq}\, \rho \Longleftrightarrow \sum_A W(A)\pi(A) \geq \sum_A W(A)\rho(A)$$

By construction, $A \succeq B$ if and only if $W(A) \geq W(B)$ and $\alpha A + (1-\alpha)B \,\dot{\sim}\, \alpha \circ A + (1-\alpha) \circ B$, therefore

$$W(\alpha A + (1-\alpha)B) = \alpha W(A) + (1-\alpha)W(B)$$

Hence, W restricted to singleton sets is continuous. If $\succeq$ satisfies *Strong Continuity* then W is continuous. Setting $U = W$ yields the result. □

Our second Lemma demonstrates that we may identify the utility of any finite set with an appropriate two-element subset. The two elements can be found by a "maxmin" or "minmax" operation.

Lemma 2: Let U be a function that represents some $\succeq$ satisfying Axiom 4. If $A \in \mathcal{A}$ is a finite set then, $U(A) = \max_{x\in A} \min_{y\in A} U(\{x,y\}) = \min_{y\in A} \max_{x\in A} U(\{x,y\})$. Moreover, there is an x^*, y^* that (x^*, y^*) solves the maxmin and (y^*, x^*) solves the minmax problem.

Proof: Suppose $\bar{u} = \max_{x\in A} \min_{y\in A} U(\{x,y\})$ and (x^*, y^*) is a solutions to this problem. We first show that $U(A) \geq \bar{u}$. Note that by construction $U(\{x^*, y\}) \geq \bar{u}, \forall y \in A$. Therefore, repeated application of *Set Betweenness* implies $U(A) = U(\cup_{y\in A}\{x^*, y\}) \geq \bar{u}$. To see that $U(A) \leq \bar{u}$, observe that for every $x \in A$ there is a y_x such that $U(x, y_x) \leq \bar{u}$. (Otherwise, x^* was not chosen optimally). Hence, by *Set Betweenness*, $U(\cup_{x\in A}\{x, y_x\}) = U(A) \leq \bar{u}$. This proves the first equality. A symmetric argument proves the second equality.

Let (x^*, y) be a solution to the maxmin problem and (y^*, x) be a solution to the minmax problem. Observe that $U(\{x^*, z\}) \geq \bar{u} = U(A)$ for all $z \in A$ and $U(\{z, y^*\}) \leq \bar{u} = U(A)$ for all $z \in A$. Hence $U(\{x^*, y^*\}) = U(A)$, $\{x^*, y^*\}$ solves the maxmin and $\{y^*, x^*\}$ minmax problem. □

Lemma 3: *Let U be a linear function that represents some $\succeq$ satisfying Axiom 4. If*

$$U(\{x\}) > U(\{x,y\}) > U(\{y\}),\ U(\{a\}) > U(\{a,b\}) > U(\{b\}) \qquad \textit{then}$$
$$U(\alpha\{x,y\} + (1-\alpha)\{a,b\}) = U(\{\alpha x + (1-\alpha)a, \alpha y + (1-\alpha)b\})$$

Proof: Let $A = \alpha\{x,y\} + (1-\alpha)\{a,b\}$. By Lemma 2, there exists (w^*, z^*) such that $U(A) = U(\{w^*, z^*\})$ and (w^*, z^*) solves $\max_{w\in A} \min_{z\in A} U(\{w,z\})$ while (z^*, w^*) solves $\min_{z\in A} \max_{w\in A} U(\{w,z\})$. We show that

$$(w^*, z^*) = (\alpha x + (1-\alpha)a, \alpha y + (1-\alpha)b)\,.$$

First, observe that linearity implies

$$U(\alpha\{x\} + (1-\alpha)\{a,b\}) > U(A) > U(\alpha\{y\} + (1-\alpha)\{a,b\})$$
$$U(\alpha\{x,y\} + (1-\alpha)\{a\}) > U(A) > U(\alpha\{x,y\} + (1-\alpha)\{b\})$$

It remains to be shown that $(w^*, z^*) \neq (\alpha x + (1-\alpha)b, \alpha y + (1-\alpha)a) \neq (z^*, w^*)$. Suppose, $w^* = \alpha x + (1-\alpha)b$. Then, since $U(A) = U(\{w^*, z^*\})$ and (x^*, y^*) solves the maxmin problem,

$$U(A) \leq U(\{\alpha x + (1-\alpha)b, \alpha y + (1-\alpha)b\}) < U(A)$$

where the last inequality again follows from linearity. Similarly, if $w^* = \alpha y + (1-\alpha)a$, then,

$$U(A) \leq U(\{\alpha y + (1-\alpha)a, \alpha y + (1-\alpha)b\}) < U(A) \qquad \square$$

We define the function $u : \Delta \to I\!R$ as

$$u(x) := U(\{x\})$$

For any a, b, δ with $a, b \in \Delta, \delta \in (0,1)$, we define the function $v : \Delta \to I\!R$

$$v(x; a, b, \delta) := \frac{U(\{a, b\}) - U(\{a, (1-\delta)b + \delta x\})}{\delta}$$

Observe that u is linear since U is linear.

Lemma 4: *Let U be a linear function that represents some $\succeq$ satisfying Axiom 4. Suppose that $U(\{a\}) > U(\{a, (1-\delta)b + \delta z\}) > U(\{(1-\delta)b + \delta z\})$ for all $z \in \Delta$. Then,*

(i) $\forall z$ such that $U(\{a\}) > U(\{a, z\}) > U(\{z\})$, $v(z; a, b, \delta) = U(\{a, b\}) - U(\{a, z\})$

(ii) $v(a; a, b, \delta) = U(\{a, b\}) - U(\{a\})$

(iii) $v(\alpha z + (1-\alpha)z'; a, b, \delta) = \alpha v(z; a, b, \delta) + (1-\alpha)v(z'; a, b, \delta)$.

(iv) $v(z; a, b, \delta) = v(z; a, b, \delta'), \forall \delta' \in (0, \delta)$.

(v) Suppose that $U(\{x\}) > U(\{x, (1-\delta)y + \delta z\}) > U(\{(1-\delta)y + \delta z\})$ for all $z \in \Delta$. Then $v(z; a, b, \delta) = v(z; x, y, \delta) + v(y; a, b, \delta)$.

Proof:

Part 1: Observe that under the hypothesis of the Lemma we may apply Lemma 3 to conclude

$$(1-\delta)U(\{a, b\}) + \delta U(\{a, z\}) = U(\{a, (1-\delta)b + \delta z\})$$

and hence simplifying the definition of $v(z; a, b, \delta)$ yields the result.

Part 2: Note that

$$(1-\delta)U(\{a, b\}) + \delta U(\{a\}) = U(\{a, (1-\delta)b + \delta a\})$$

by the linearity of U. Hence, (ii) follows from the definition of v.

Part 3: Applying Lemma 3 and the linearity of U we obtain

$$U(\{a,(1-\delta)b+\delta[\alpha z+(1-\alpha)z']\}) = U(\alpha\{a,(1-\delta)b+\delta z\}+(1-\alpha)\{a,(1-\delta)b+\delta z'\})$$
$$= \alpha U(\{a,(1-\delta)b+\delta z\})+(1-\alpha)U(\{a,(1-\delta)b+\delta z'\})$$

The definition of v now implies part (iii).

Part 4: Let $\delta' < \delta$. Observe that

$$\frac{\delta-\delta'}{\delta}b+(1-\frac{\delta-\delta'}{\delta})((1-\delta)b+\delta z) = (1-\delta')b+\delta' z.$$

Moreover, since $(a,(1-\delta)b+\delta z)$ and (a,b) satisfy the hypothesis of Lemma 3 it follows that

$$U(\{a,(1-\delta')b+\delta' z\}) = \frac{\delta-\delta'}{\delta}U(\{a,b\})+(1-\frac{\delta-\delta'}{\delta})U(\{a,(1-\delta)b+\delta z\})$$

Hence, (iv) follows from the definition of v.

Part 5: We need to show that

$$U(\{a,b\})-U(\{a,(1-\delta)b+\delta z\}) =$$
$$U(\{x,y\})-U(\{x,(1-\delta)y+\delta z\})+U(\{a,b\})-U(\{a,(1-\delta)b+\delta y\})$$

which is equivalent to

$$\frac{1}{2}U(\{x,(1-\delta)y+\delta z\})+\frac{1}{2}U(\{a,(1-\delta)b+\delta y\}) = \frac{1}{2}U(\{x,y\})+\frac{1}{2}U(\{a,(1-\delta)b+\delta z\})$$

We apply Lemma 3 to conclude that both sides are equal to

$$U\left(\left\{\frac{x+a}{2},\frac{(1-\delta)b+\delta z+y}{2}\right\}\right) \qquad \square$$

Lemma 5: *Let U be a linear function that represents some $\succeq$ satisfying Axiom 4. Consider $a,y\in\Delta$ such that $U(\{a\})\geq U(\{a,y\})\geq U(\{y\})$. Suppose $b\in\Delta$ and δ satisfy $U(\{a\})>U(\{a,(1-\delta)b+\delta z\})>U(\{1-\delta)b+\delta z\})$ for all $z\in\Delta$. Then,*

$$U(\{a,y\}) = \max_{w\in\{a,y\}} u(w)+v(w;a,b,\delta)-\max_{z\in\{a,y\}} v(z;a,b,\delta)$$

Proof: First, consider the case where $U(\{a\}) > U(\{a,y\}) > U(\{y\})$. By (i) and (ii) of Lemma 4, $v(y;a,b,\delta) = U(\{a,b\}) - U(\{a,y\}) \geq U(\{a,b\}) - U(\{a\}) = v(a;a,b,\delta)$ and

$$\begin{aligned} u(a) + v(a;a,b,\delta) - v(y;a,b,\delta) &= U(\{a\}) + U(\{a,b\}) - U(\{a\}) - U(\{a,b\}) + U(\{a,y\}) \\ &= U(\{a,y\}) > U(\{y\}) = u(y) + v(y;a,b,\delta) - v(y;a,b,\delta) \end{aligned}$$

Second, consider the case where $U(\{a\}) = U(\{a,y\}) > U(\{y\})$. In this case, it is sufficient to show that $v(a;a,b,\delta) \geq v(y;a,b,\delta)$. Since $U(\{a,(1-\delta)b+\delta a\}) = (1-\delta)U(\{a,b\}) + \delta U(\{a\}) = (1-\delta)U(\{a,b\}) + \delta U(\{a,y\})$, $v(a;a,b,\delta) \geq v(y;a,b,\delta)$ if and only if $(1-\delta)U(\{a,b\}) + \delta U(\{a,y\}) \leq U(\{a,(1-\delta)b+\delta y\})$. Let $A = (1-\delta)\{a,b\} + \delta\{a,y\}$. We claim that

$$U(A) = \min_{w \in A} U(\{a,w\})$$

Since $(1-\delta)b + \delta y \in A$ the result follows from this claim. By Lemma 2, there exists $z \in \Delta$ such that $U(A) = \min_{w \in A} U(\{z,w\})$. To prove the claim suppose $z \neq a$. If $z = (1-\delta)a + \delta y$ then setting $w = (1-\delta)b + \delta y$ yields $U(A) = (1-\delta)U(\{a,b\}) + \delta U(\{y\}) < U(A)$ and hence we have a contradiction. Similarly, if $z = (1-\delta)b + \delta a$ setting $w = (1-\delta)b + \delta y$ establishes that $U(A) = (1-\delta)U(b) + \delta U(\{a,y\}) < U(A)$. And finally, if $z = (1-\delta)b + \delta y$ setting $w = z$ yields $U(A) \leq (1-\delta)U(\{b\}) + \delta U(\{y\}) < U(A)$ again, a contradiction.

Third, consider $U(\{a\}) > U(\{a,y\}) = U(\{y\})$. We must show that

$$v(y;a,b,\delta) \geq v(a;a,b,\delta) + u(a) - u(y)$$

By Lemma 4 (ii), $v(a;a,b,\delta) + u(a) - u(y) = U(\{a,b\}) - u(y) = U(\{a,b\}) - U(\{a,y\})$. Then, it follows from the definition of $v(y;a,b,\delta)$ that the above inequality hold iff

$$U(\{a,(1-\delta)b+\delta y\}) \leq (1-\delta)U(\{a,b\}) + \delta U(\{a,y\})$$

Let $A = (1-\delta)\{a,b\} + \delta\{a,y\}$. By Lemma 2,

$$(1-\delta)U(\{a,b\}) + \delta U(\{a,y\}) = U(A) \geq \min_{w \in A} U(\{a,w\})$$

But,

$$U(A) < (1-\delta)U(\{a\}) + \delta U(\{a,y\}) = U(\{a,(1-\delta)a+\delta y\})$$

$$U(A) < (1-\delta)U(\{a,b\}) + \delta U(\{a\}) = U(\{a,(1-\delta)b+\delta a\})$$

$$U(A) < U(\{a\})$$

So,

$$\min_{w \in A} U(\{a,w\}) = U(\{a,(1-\delta)b+\delta y\})$$

and therefore

$$(1-\delta)U(\{a,b\}) + \delta U(\{a,y\}) \geq U(\{a,(1-\delta)b+\delta y\})$$

as desired.

Finally, in the case where $U(\{a\}) = U(\{a,y\}) = U(\{y\})$ it follows that $u(a) = u(y)$ and hence, $\max_{w\in\{a,y\}} u(w) + v(w;a,b,\delta) - \max_{z\in\{a,y\}} v(z;a,b,\delta) = u(a) = U(\{a,y\})$. □

To prove Theorem 1, we first consider all finite subsets of Δ. By Lemma 1, there is a continuous, linear representation U, of $\succeq$. Observe that if $U(\{x\}) = U(\{y\})$ for all $x, y \in \Delta$ then by Lemma 2, $U(A) = U(B)$ for all non-empty finite subsets of Δ. Hence the result follows trivially. Thus, consider the case where $U(\{x\}) > U(\{y\})$ for some $x, y \in \Delta$. We can distinguish three cases:

Case 1: $U(\{x\}) > U(\{x,y\}) > U(\{y\})$ for some pair x, y. Choose $\delta > 0$ so that for all $z \in \Delta$, $U(\{x\}) > U(\{x,(1-\delta)y+\delta z\}) > U(\{(1-\delta)y+\delta z\})$. Continuity of U implies that such a δ exists. Let $u(z) := U(\{z\})$ and $v(z) := v(z;x,y,\delta)$ for all $z \in \Delta$. By Lemma 4, we know that v is linear. Consider the set $A = \{a,b\}$, where a and b are in the relative interior of Δ. Assume wlog that $u(a) \geq u(b)$. We claim that there is a c such that $u(a) > U(\{a,c\}) > u(c)$. Since a is in the relative interior of Δ, there is an a' and an $\alpha \in (0,1)$ such that $\alpha a' + (1-\alpha)x = a$. By linearity, $c = \alpha a' + (1-\alpha)y$ has the desired property. Then, for δ' sufficiently small $U(\{a\}) > U(\{a,(1-\delta')c+\delta' z\}) > U(\{(1-\delta')c+\delta' z\})$ for all $z \in \delta$. Hence, by Lemma 5,

$$U(\{a,b\}) = \max_{w\in\{a,b\}}\{u(w)+v(w;a,c,\delta')\} - \max_{w\in\{a,b\}}\{v(w;a,c,\delta')\}$$

Let $\delta^* = \min\{\delta, \delta'\}$. By Lemma 4 (iv), $v(\cdot; x, y, \delta^*) = v(\cdot; x, y, \delta)$ and $v(\cdot; a, c, \delta^*) = v(\cdot; a, c, \delta')$. By Lemma 4 (v), for an appropriate constant k, $v(\cdot; x, y, \delta^*) = v(\cdot; a, c, \delta^*) + k$ and hence it follows that

$$U(\{a,b\}) = \max_{w \in \{a,b\}} \{u(w) + v(w)\} - \max_{w \in \{a,b\}} \{v(w)\}$$

Since U, u and v are all linear, the above equation holds for *all* $a, b \in \Delta$. Now consider an arbitrary finite set A. We know that

$$\begin{aligned}
U(A) &= \max_{a \in A} \min_{b \in A} U(\{a,b\}) \\
&= \max_{a \in A} \min_{b \in A} \left\{ \max_{w \in \{a,b\}} \{u(w) + v(w)\} - \max_{w \in \{a,b\}} \{v(w)\} \right\} \\
&= \max_{a \in A} \min_{b \in A} \left\{ \max_{w \in \{a,b\}} \{u(w) + v(w)\} + \min_{w \in \{a,b\}} \{-v(w)\} \right\} \\
&= \max_{a \in A} \{u(w) + v(w)\} + \min_{b \in A} \{-v(w)\}
\end{aligned}$$

Case 2: $U(\{x\}) = U(\{x,y\}) > U(\{y\})$ for all x, y with $U(\{x\}) > U(\{y\})$. In this case let $u(x) = U(\{x\})$ and $v(x) = 0$. The result follows from Lemma 2.

Case 3: $U(\{x\}) > U(\{x,y\}) = U(\{y\})$ for all x, y with $U(\{x\}) > U(\{y\})$. In this case let $u(x) = U(\{x\})$ and $v(x) = -U(\{x\})$. Again, the result follows from Lemma 2.

To complete the proof for the case of finite A we show that if x, y satisfies $U(\{x\}) = U(\{x,y\}) > U(\{y\})$ and a, b satisfies $U(\{a\}) > U(\{a,b\}) = U(\{b\})$ then by the continuity of U, there is an α^* such that

$$U(\{\alpha^* x + (1-\alpha^*)a\}) > U(\{\alpha^* x + (1-\alpha^*)a, \alpha^* y + (1-\alpha^*)b\}) > U(\{\alpha^* y + (1-\alpha^*)b\}) \quad (*)$$

and hence the cases are exhaustive. To prove the claim, for all $\alpha \in [0,1]$, let

$$f(\alpha) := \frac{U(\{\alpha x + (1-\alpha)a, \alpha y + (1-\alpha)b\}) - (\alpha U(\{y\}) + (1-\alpha)U(\{b\}))}{\alpha U(\{x\}) + (1-\alpha)U(\{a\}) - (\alpha U(\{y\}) + (1-\alpha)U(\{b\}))}$$

and observe that f is well defined, continuous and $f(0) = 0, f(1) = 1$ and hence takes on the value $1/2$ at some $\alpha^* \in (0,1)$. Hence α^* satisfies inequality $(*)$.

Observe that we can approximate any compact set A, by a sequence of finite sets A_k, as follows: Let $x_k, k = 1, 2, \ldots$ be a countable dense subset of A. Let $A_k := \{x_1, \ldots, x_k\}$.

Both U and $\succeq$ are continuous and hence, the representation can be extended to arbitrary sets. □

7.2 Proof of Theorem 2

By Lemma 1, a linear representation U, of the preferences exists. Therefore, Lemma's 3, 4 and 5 also hold if Axiom 2. Next, we show that one of the two desired representations exists for all finite subsets of Δ.

Claim 1: *Suppose, there exists some pair x, y in the relative interior of Δ such that $U(\{x\}) > U(\{x, y\}) > U(\{y\})$. Then, there is a neighborhood of y, N_y such that for all $z \in N_y$, $U(\{x\}) > U(\{x, z\}) > U(\{z\})$.*

Proof: Since the restriction of U to singleton subsets of Δ is continuous, there exists a neighborhood of y, N'_y such that $U(\{x\}) > U(\{z\})$ for all $z \in N'_y$. By *Set Betweenness*, $U(x) \geq U(\{x, z\}) \geq U(z)$ whenever $\{x\} \succeq \{z\}$. By *Upper Semi-Continuity*, if $y_k \to y$ and $U(\{x, y_k\}) = U(\{x\})$ for all k then $U(\{x, y\}) \geq U(\{x\})$. Hence, if the claim is false, there exists a sequence $y_k \to y$ such that $U(\{x, y_k\}) = U(\{y_k\})$ for all k. Since y is in the relative interior of Δ, we can choose $z_k \in \Delta$, $\delta_k \in (0, 1)$ with $\delta_k \to 0$ such that $y_k = (1 - \delta_k)y + \delta_k z_k$. Let $x_k := (1 - \delta_k)x + \delta_k z_k$. By *Independence*, $U(\{x_k, y_k\}) = (1-\delta_k)U(\{x, y\})+\delta_k U(\{z_k\})$. Define $U_k = \frac{1}{2}U(\{x, y_k\})+\frac{1}{2}U(\{x_k, y\})$. Since $U(\{x, y_k\}) = U(\{y_k\})$ *Upper Semi-Continuity* implies that $\limsup U_k \leq \frac{U(\{y\})+U(\{x,y\})}{2}$. On the other hand, *Set Betweenness* implies that

$$
\begin{aligned}
U_k &= U\left(\frac{1}{2}\{x, y_k\} + \frac{1}{2}\{x_k, y\}\right) \\
&\geq \min\left\{(1 - \frac{\delta_k}{2})U(\{x, y\}) + \frac{\delta_k}{2}U(\{z\}), U\left(\left\{\frac{1}{2}x + \frac{1}{2}y\right\}\right), U\left(\left\{\frac{1}{2}x_k + \frac{1}{2}y_k\right\}\right)\right\}
\end{aligned}
$$

Since U is continuous when restricted to singleton sets, $U\left(\{\frac{1}{2}x_k + \frac{1}{2}y_k\}\right) \to \frac{U(\{x\})+U(\{y\})}{2}$. But $U(\{x, y\}) > U(\{y\})$ and $U(\{x\}) > U(\{x, y\})$, establishing the desired contradiction. □

Whenever $U(\{x\}) > U(\{x, y\}) > U(\{y\})$ for some x, y we use *Independence* to find x, y in the relative interior of Δ satisfying the same inequalities. Moreover, Claim 1 establishes that we may find a $\delta > 0$ such that $U(\{x\}) > U(\{x, (1 - \delta)y + \delta z\}) > U(\{(1 - \delta)y + \delta z\})$

for all $z \in \Delta$. Thus, we may proceed as in the proof of Case 1 of Theorem 1 to show that for some linear u and v, $U(A) := \max_{z\in A}\{u(z) + v(z) - \max_{w\in A} v(w)\}$ represents $\succeq$ on all finite subset of Δ.

Claim 3 below establishes that if no x, y such that $U(\{x\}) > U(\{x, y\}) > U(\{y\})$ exists then the agent has no self-control. Then, we prove that no self-control implies that the second type of representation holds for all finite subsets of Δ.

The triple (w, x, y) is a *cycle* if $\{w\} \succ \{x\} \succ \{y\}$ and either $\{w, x\} \sim \{w\}$, $\{x, y\} \sim \{x\}$, $\{w, y\} \sim \{y\}$ or $\{w, x\} \sim \{x\}$, $\{x, y\} \sim \{y\}$, $\{w, y\} \sim \{w\}$.

Claim 2: *If $\{x, y\} \sim x$ or $\{x, y\} \sim y$ for all $x, y \in \Delta$ then no cycle exists.*

Proof: We prove that no cycle (w, x, y) such that $\{w, x\} \sim \{w\}$, $\{x, y\} \sim \{x\}$, $\{w, y\} \sim \{y\}$ exists. A symmetric argument yields the impossibility of a cycle such that $\{w, x\} \sim \{x\}$, $\{x, y\} \sim \{y\}$, $\{w, y\} \sim \{w\}$

Suppose (w, x, y) is the a cycle such that $\{w, x\} \sim \{w\}$, $\{x, y\} \sim \{x\}$, $\{w, y\} \sim \{y\}$. Hence, $U(\frac{1}{2}\{w, x\} + \frac{1}{2}\{x, y\}) = (U(\{w\}) + U(\{x\}))/2$. But $\frac{1}{2}\{w, x\} + \frac{1}{2}\{x, y\} = \{\frac{1}{2}w + \frac{1}{2}x, \frac{1}{2}y + \frac{1}{2}x\} \cup \{x\} \cup \{\frac{1}{2}w + \frac{1}{2}y\}$. But, $U(\{\frac{1}{2}w + \frac{1}{2}x, \frac{1}{2}y + \frac{1}{2}x\}) = \frac{U(\{y\})+U(\{x\})}{2}$ and $U(\{\frac{1}{2}w + \frac{1}{2}y\}) = \frac{1}{2}U(\{w\}) + \frac{1}{2}U(\{y\})$. Hence, by *Set Betweenness*, $U(\frac{1}{2}\{w, x\} + \frac{1}{2}\{x, y\}) \leq \max\left\{\frac{U(\{y\})+U(\{x\})}{2}, U(\{x\}), \frac{U(\{w\})+U(\{y\})}{2}\right\} < \frac{U(\{w\})+U(\{x\})}{2}$, a contradiction. □

Claim 3: *If $\{x, y\} \sim x$ or $\{x, y\} \sim y$ for all $x, y \in \Delta$ then $A \cup B \sim A$ or $A \cup B \sim B$ for all finite $A, B \in \mathcal{A}$.*

Proof: Assume $A \succ A \cup B \succ B$. By Lemma 2,

$$U(A \cup B) = \max_{x\in A\cup B} \min_{y\in A\cup B} U(\{x, y\}) = \min_{y\in A\cup B} \max_{x\in A\cup B} U(\{x, y\})$$

Let $S \subset \Delta^2$ denote the set of pairs (x^*, y^*) such that (x^*, y^*) is a solution to the maxmin and (y^*, x^*) is a solution to the minmax problem above. By Lemma 2, The set S is non-empty.

Step 1: $(x^*, y^*) \in S$ implies $x^* \in A\backslash B$ and $y^* \in B\backslash A$. If $x^* \in B$ then $U(B) \geq \min_{y\in B} U(\{x^*, y\}) \geq \min_{y\in A\cup B} U(\{x^*, y\}) = U(A \cup B)$, a contradiction. Similarly, if $y^* \in A$ then $U(A) \leq \max_{x\in A} U(\{x, y^*\}) \leq \max_{y\in A\cup B} U(\{x, y^*\}) = U(A \cup B)$, again a contradiction. Hence, $x^* \notin B$ and $y^* \notin A$. It follows that $x^* \in A$, $y^* \in B$.

Step 2: There exists $(x^*, y^*) \in S$ such that $\{x^*\} \succ \{y^*\}$. Take $(x,y) \in S$. Either $\{x,y\} \sim \{x\}$ or $\{x,y\} \sim \{y\}$. If $\{x,y\} \sim \{x\}$ then since x maximizes $U(\{\cdot, y\})$, we have $\{x\} \succeq \{y\}$. If $\{x,y\} \sim \{y\}$ then since y minimizes $U(\{x, \cdot\})$, we have $\{x\} \succeq \{y\}$. Hence, $\{x\} \succeq \{y\}$. If $\{x\} \succ \{y\}$ we are done. So assume $\{x\} \sim \{y\}$. Therefore $A \succ A \cup B \sim \{y\} \succ B$. By Step 1, $(y,y) \notin S$. So, there exists $z \in A \cup B$ such that $\{y\} \sim A \cup B \succ \{y,z\} \sim \{z\}$. If $(x,z) \in S$ we are done. If not, there exists $w \in A \cup B$ such that $\{w\} \sim \{w,z\} \succ A \cup B \succ \{z\}$. But then (w,y,z) is a cycle, a contradiction.

By Step 2, we can choose $(x^*, y^*) \in S$ such that $\{x^*\} \succ \{y^*\}$. By our hypothesis, $U(\{x^*, y^*\}) = U(\{x^*\})$ or $U(\{x^*, y^*\}) = U(\{y^*\})$. Assume the first equality holds. Then, $U(A \cup B) = U(\{x^*\})$. Since $U(A) > U(A \cup B)$, by Lemma 2, there exists $w \in A$ such that $U(\{w\}) = U(\{w, x^*\}) > U(\{x^*\})$. Since $(x^*, y^*) \in S$, $U(\{y^*\}) = U(\{w, y^*\}) \leq U(\{x^*\}) < U(\{w\})$. Then, (w, x^*, y^*) is a cycle, contradicting Claim 2 above. A symmetric argument yields a contradiction if $U(\{x^*, y^*\}) = U(\{y^*\})$ and completes the proof. □

Next, we define the binary relation over lotteries that captures the temptation ranking for an agent without self-control.

Definition: *xRy if $\{x\} \not\sim \{y\}$ and $\{x,y\} \sim \{x\}$, or if $\{x\} \sim \{y\}$ and $\{z,y\} \sim \{y\}$ implies $\{z,x\} \sim \{x\}$.*

To interpret the above definition, note that whenever $\{x\} \not\sim \{y\}$ then $\{x,y\} \sim \{x\}$ implies that the temptation ranking of x higher than that of y. If the decision-maker is indifferent between the singleton sets x and y then we cannot infer their temptation rankings by comparing $\{x,y\}$ and $\{x\}$. However, if for some z, $\{x\} \sim \{x,z\}$ and $\{y\} \not\sim \{y,z\}$ then the temptation ranking of x must be higher than that of y.

Lemma 6: *If $\succeq$ satisfies Axiom 4 and has no self-control then R is a complete and transitive binary relation on Δ.*

Proof: First we demonstrate that R is complete. If $\{x\} \not\sim \{y\}$ then xRy or yRx by Axiom 4 and no self-control. Suppose that $\{x\} \sim \{y\}, \{x,w\} \sim \{x\}$ and $\{y,w\} \not\sim \{y\}$ for some w. We need to show that $\{y,z\} \sim \{y\}$ implies $\{x,z\} \sim \{x\}$. If $\{z\} \sim \{y\}$ then the result follows from Axiom 4 and the transitivity of $\succeq$. Hence assume that $\{z\} \not\sim \{y\}$. We know that $\{w\} \not\sim \{y\}$. We also know that $\{w,x,y,z\} \sim \{y\}$ since both $\{w,x\} \sim \{y\}$

and $\{y,z\} \sim \{y\}$. But then it must be that either $\{w,y\} \sim \{y\}$ or $\{x,z\} \sim \{y\}$. Since the former indifference does not hold the latter holds and we have $\{x,z\} \sim \{y\} \sim \{x\}$ as desired.

To prove transitivity assume that xRy and yRz. If $\{x\} \not\sim \{y\} \not\sim \{z\} \not\sim \{x\}$ then by Axiom 4 and no self-control we have $\{x,y,z\} \sim \{x\}$ or $\{x,y,z\} \sim \{y,z\}$. But we must also have $\{x,y,z\} \sim \{x,y\}$ or $\{x,y,z\} \sim \{z\}$. Therefore, $\{x,y,z\} \sim \{x\}$. Again, by Axiom 4 and no self-control, we observe that $\{x,y,z\} \sim \{x,z\}$ or $\{x,y,z\} \sim \{y\}$. Since $\{x,y,z\} \sim \{x\}$ we may rule out the latter case and obtain the desired conclusion. If $\{x\} \sim \{y\} \sim \{z\}$ then $\{w,z\} \sim \{z\}$ implies $\{w,y\} \sim \{y\}$ which in turn implies $\{w,x\} \sim \{x\}$ (since $xRyRz$). From the transitivity of $\succeq$ we conclude that $\{x\} \sim \{z\}$ and hence xRz.

If $\{x\} \sim \{y\} \not\sim \{z\}$ then since yRz we have $\{y,z\} \sim \{y\}$ and since xRy we conclude $\{x,z\} \sim \{x\}$. Therefore, xRz.

If $\{x\} \not\sim \{y\} \sim \{z\}$ then again $\{x\} \not\sim \{z\}$ and hence it is sufficient to show that $\{x,z\} \sim \{x\}$. But yRz and $\{x,y\} \not\sim \{y\}$ implies $\{x,z\} \not\sim \{z\}$ and hence $\{x,z\} \sim \{x\}$.

If $\{x\} \sim \{z\} \not\sim \{y\}$ then $\{x,y\} \sim \{x\}$ and $\{y,z\} \not\sim \{z\}$ and hence z not$R$ x and by the completeness xRz. □

Lemma 7: *There exists a linear function $v : \Delta \to I\!R$ that represents R.*

Proof: We prove the Lemma by establishing that R satisfies the standard von Neumann-Morgenstern assumptions. By Lemma 6, R is complete and transitive. Next, we prove that R satisfies the independence axiom. Let P denote the strict part of R; that is, xPy iff not yRx. By Lemma 1, we know that there is a linear function $U : \mathcal{A} \to \mathcal{R}$ that represents $\succeq$. Moreover, by *Upper Semi Continuity* and *Lower Singleton Continuity*, the restriction of U to singleton sets is continuous. To prove *Independence*, assume that xPy and $\alpha \in (0,1)$.

Step 1: Let $\{x\} \not\sim \{y\}$. Hence $U(\{x\}) = U(\{x,y\})$. By linearity, $U(\{\alpha x + (1-\alpha)z\}) = U(\{\alpha x + (1-\alpha)z, \alpha y + (1-\alpha)z\})$ and $U(\{\alpha x + (1-\alpha)z\}) \neq U(\{\alpha y + (1-\alpha)z\})$. That is, $[\alpha x + (1-\alpha)z]P[\alpha y + (1-\alpha)z]$ as desired.

Step 2: Let $\{x\} \sim \{y\}$. Since xPy, there exists w such that $\{x\} \sim \{x,w\}$ and $\{y\} \not\sim \{y,w\}$. Clearly, $\{x\} \not\sim \{w\} \not\sim \{y\}$. Hence, applying Step 1, we conclude that $[\alpha x + (1-\alpha)z]P[\alpha w + (1-\alpha)z]$ and $[\alpha w + (1-\alpha)z]P[\alpha y + (1-\alpha)z]$ and since R is complete and transitive, we conclude that $[\alpha x+(1-\alpha)z]P[\alpha y+(1-\alpha)z]$. This completes the proof of *Independence*.

To complete the proof, we will show that R satisfies von Neumann-Morgenstern continuity. That is, there exists α, β such that $\alpha x + (1-\alpha)z \succ y \succ \beta x + (1-\beta)z$ whenever $xPyPz$, Assume that $xPyPz$.

Step 3: If $\{z\} \succ \{y\}$ then there exists $\bar{\beta} > 0$ such that $\beta < \bar{\beta}$ implies $yP[\beta x+(1-\beta)z]$. Similarly, if $\{y\} \succ \{x\}$ then there exists $\bar{\alpha} < 1$ such that $\alpha > \bar{\alpha}$ implies $[\alpha x + (1-\alpha)z]Py$. To prove the first assertion, assume that $U(\{y, \beta_k x + (1-\beta_k)z\}) = U(\{\beta_k x + (1-\beta_k)z\})$ for some sequence $\beta_k \in (0,1)$ converging to 0. Then, by *Upper Semi-Continuity* and the continuity of U on singleton sets, we obtain $\{y, z\} \succeq \{z\}$ and hence $\{y, z\} \not\succ \{y\}$, a contradiction. The proof of the second assertion follows from a similar argument.

Step 4: If $\{y\} \succ \{z\}$ then there exists $\bar{\beta} > 0$ such that $\beta < \bar{\beta}$ implies $yP[\beta x+(1-\beta)z]$. Similarly, if $\{x\} \succ \{y\}$ then there exists $\bar{\alpha} < 1$ such that $\alpha > \bar{\alpha}$ implies $[\alpha x + (1-\alpha)z]Py$. To prove the first assertion, assume that $U(\{y, \beta_k x + (1-\beta_k)z\}) = U(\{\beta_k x + (1-\beta_k)z\})$ for some sequence $\beta_k \in (0,1)$ converging to 0. Since U is linear, we can assume wlog that $U(\{y\}) > U(\{\beta_k x + (1-\beta_k)z\})$ for all k. Therefore, $[\beta_k x + (1-\beta_k)z]Py$ for all k. Applying Step 3 to $[\beta_k x + (1-\beta_k)z]PyPz$ yields $\bar{\beta} > 0$ such that $\beta < \bar{\beta}$ implies yPw_k, where $w_k := \beta[\beta_k x + (1-\beta_k)z] + (1-\beta)z$, a contradiction. The second assertion follows from a similar argument.

Step 5: If $\{y\} \sim \{z\}$ then there exists $\beta > 0$ such that $yP[\beta x + (1-\beta)z]$. Similarly, if $\{x\} \sim \{y\}$ then there exists $\alpha < 1$ such that $[\alpha x + (1-\alpha)z]Py$. Again, we will prove only the first assertion and omit the similar proof of the second. By the definition of R, there exists w such that $yPwPz$ and $\{w\} \not\sim \{z\}$. Applying Step 3 or Step 4 to $xPwPz$ yields $\beta \in (0,1)$ such that $wP[\beta x + (1-\beta)z]$. Then yPw and the fact that R is complete and transitive yields the desired conclusion: $yP[\beta x + (1 - \beta z]$.

Clearly, Steps $3-5$ establish von Neumann-Morgenstern continuity of R and complete the proof. □

Define $u(x) := U(\{x\})$ for all $x \in \Delta$. Since U is linear, so is u. Now, we show that the representation holds for all finite sets $A \in \mathcal{A}$. Pick $x^* \in A$ such that $v(x^*) \geq v(y)$ for all $y \in A$ and $v(x^*) = v(y)$ implies $u(x^*) \geq u(y)$. Hence, $u(x^*) = \max_{x \in A} u(x)$ subject to $v(x) \geq$

$v(y)$ for all $y \in A$. Note that $A = \bigcup_{y \in A} \{x^*, y\}$ and since U represents $\succeq$ and has no self-control, we have $U(A) = U(\{x^*, y\})$ for some $y \in A$. Since v represents R we have $U(\{x^*, y\}) = u(x^*)$ and hence $U(A) = u(x^*) = \max_{x \in A} u(x)$ subject to $v(x) \geq v(y)$ for all $y \in A$ as desired.

Hence, we have show that for all finite subsets, a self-control representation exist if $\{x\} \succ \{x, y\} \succ \{y\}$ and an overwhelming temptation representation exists if no such x, y can be found. Next, we show that in both cases, u and v are continuous. As noted earlier, the continuity of u follows immediately from the *Upper Semi-Continuity* and *Lower Singleton Continuity* of $\succeq$.

To prove the continuity of v first assume $\{x\} \succ \{x, y\} \succ \{y\}$ for some x, y. That is, a self-control representation exists and $u(x) + v(x) > u(y) + v(y)$ and $v(x) < v(y)$. If v is not continuous, since Δ is compact, there exists z_n converging to some z such that $\lim v(z_n) = \beta \neq v(z)$. Suppose $\beta > v(z)$. Let $x_n = \alpha x + (1 - \alpha) z_n$ and $\bar{x} = \alpha x + (1 - \alpha) z$. Choose $\alpha \in (0, 1)$ such that $u(x_n) + v(x_n) > u(y) + v(y)$ and $v(x_n) < v(y)$ for all n. Pick $\gamma \in (0, 1)$ such that $\alpha u(x) + (1 - \alpha)\beta - v(y) > \gamma u(x) + (1 - \gamma) u(y) > \alpha u(x) + (1 - \alpha) v(z) - v(y)$. Then, for n sufficiently large $\{x_n, y\} \succeq \{\gamma x + (1 - \gamma) y\}$ but $\{\gamma x + (1 - \gamma) y\} \succ \{\bar{x}, y\}$, contradicting *Upper Semi-Continuity.* If $\beta < v(z)$ then define y_n and $\bar{y}$ by replacing x with y in the corresponding definitions of x_n and $\bar{x}$. Again, choose $\alpha \in (0, 1)$ such that $u(x) + v(x) > u(y_n) + v(y_n)$ and $v(x) < v(y_n)$ for all n. Pick $\gamma \in (0, 1)$ such that $u(x) + v(x) - \alpha v(y) - (1 - \alpha)\beta > \gamma u(x) + (1 - \gamma) u(y) > u(x) + v(x) - \alpha v(y) - (1 - \alpha) v(z)$. Then, again for n sufficiently large $\{x, y_n\} \succeq \{\gamma x + (1 - \gamma) y\}$ but $\{\gamma x + (1 - \gamma) y\} \succ \{x, \bar{y}\}$, contradicting *Upper Semi-Continuity.*

If there exists no x, y such that $\{x\} \succ \{x, y\} \succ \{y\}$, then we have an overwhelming temptation representation over finite sets. If there exists no x, y such that $\{x\} \succ \{y\}$ it follows from our definition of R above that v is a constant and we are done. So, assume $\{x\} \succ \{y\}$. If v is constant or an affine transformation of u we are done. Otherwise, by linearity, we can find a, b such that $u(a) > u(b)$ and $v(a) < v(b)$. Without loss of generality, assume $a = x$ and $b = y$. Define $z_n, z, \beta, x_n, y_n, \bar{x}, \bar{y}$ as above. If $\beta > v(z)$ then pick $\alpha \in (0, 1)$ such that $\alpha v(x) + (1 - \alpha)\beta > v(\bar{y}) > \alpha v(x) + (1 - \alpha) v(z)$. Then, for n sufficiently large $\{x_n, \bar{y}\} \succeq \{.5\bar{x} + .5\bar{y}\} \succ \{\bar{x}, \bar{y}\}$, contradicting *Upper Semi-Continuity.*

If $\beta < v(z)$ then pick $\alpha \in (0,1)$ such that $v(\bar{y}) > v(\bar{x}) > \alpha v(x) + (1-\alpha)\beta$. Then, for n sufficiently large $\{\bar{x}, y_n\} \succeq \{.5\bar{x} + .5\bar{y}\} \succ \{\bar{x}, \bar{y}\}$, once again contradicting *Upper Semi-Continuity.*

The following lemma enables us to extend the either representation to arbitrary compact sets and completes the proof.

Lemma 8: *Let $\succeq$ be a preference relation on $\mathcal{A}$ that satisfies Axioms 2a and 4. For any $A \in \mathcal{A}$, if there exists a non-empty, finite set $B \subset A$ such that $B \subset C \subset A$ implies $C \sim B$ for all finite C, then $A \sim B$.*

Proof: Let $B \subset A$ be a non-empty, finite set such that $B \subset C \subset A$ implies $C \sim B$ for all finite C. Define A_k, as follows: Let $x_k, k = 1, 2, \ldots$ be a countable dense subset of A. Let $A_k := \{x_1, \ldots, x_k\} \cup B$. Since $A_k \sim B$ for all k, by *Upper Semi-Continuity* $A \succeq B$. Next, let $B_\varepsilon(x)$ denote the open ε ball around $x \in A$. Since A is compact, for any $\varepsilon > 0$ there exists a finite number of such open balls that cover A. Let $\bar{B}_\varepsilon(x)$ denote the closure of $B_\varepsilon(x)$. Since a finite collection of sets $B_\varepsilon(x)$ cover A, by Axiom 4, there exists some $x, \varepsilon > 0$ such that $(B \cup \bar{B}_\varepsilon(x)) \cap A \succeq A$. Choose a sequence of $\varepsilon_k > 0$ converging to 0 to obtain a sequence $(B \cup \bar{B}_{\varepsilon_k}(x_k)) \cap A \succeq A$, where x_k converges to some $x \in A$. It follows that $(B \cup \bar{B}_{\varepsilon_k}(x)) \cap A$ converges to $B \cup \{x\}$. Hence, by Axiom 2a, we have $B \cup \{x\} \succeq A$. By our hypothesis, $B \sim B \cup \{x\}$ and therefore $B \succeq A$ establishing $A \sim B$ as desired. □

To see how Lemma 8 extends the representation from finite sets to the entire $\mathcal{A}$, assume that a self-control representation holds on finite sets. Let x^* maximize u in A, y^* maximize v in A and set $B = \{x^*, y^*\}$ and apply Lemma 8. If an overwhelming temptation representation holds of finite sets, then let x^* be a solution to the maximization problem $\max u(x)$ subject to $v(x) \geq v(y)$ for all $y \in A$, let $B := \{x^*\}$ and apply Lemma 8. (Note that the continuity of u and v ensures that x^* and y^* are well-defined for both of the above cases.) □

7.3 A Lemma on Linear Independence

The following Lemma is used in the proof of Theorems 3 and 7. The proofs of the two facts stated after the Lemma follow from the argument used in proving Lemma 9. Let $e := (1, 1, \ldots, 1)$.

Lemma 9: *Let w^1, w^2, w^3, e be linearly independent vectors in $I\!R^n$ and x be any vector in $S^{n-1} := \{z \in I\!R^n_+ : \sum z_i = 1\}$. Then, there exists y in the relative interior of Δ such that $w^1 \cdot (x - y) > 0 = w^2 \cdot (x - y) > w^3 \cdot (x - y)$.*

Proof: Choose $w^4, \ldots, w^{n-1}$ so that $w^1, w^2, w^3, \ldots, w^{n-1}, e$ are n linearly independent vectors in $I\!R^n$. The smallest subspace containing the vectors $w^1 + w^3, w^2, w^4, \ldots, w^{n-1}, e$ is a hyperplane and this hyperplane does not contain w^1 or w^3. Let ζ be the normal to this hyperplane. Hence, $w^1 \cdot \zeta \neq 0$, $w^3 \cdot \zeta \neq 0$ and $(w^1 + w^3) \cdot \zeta = 0$. Then, either $w^1 \cdot \zeta > 0 > w^3 \cdot \zeta$ or $w^1 \cdot \zeta < 0 < w^3 \cdot \zeta$. Without loss of generality assume $w^1 \cdot \zeta > 0 > w^3 \cdot \zeta$ (otherwise use $-\zeta$ instead of ζ). Note that for $\epsilon > 0$ small, $y = x - \epsilon\zeta$ is in the relative interior of Δ and $w^1 \cdot (x - y) > 0 = w^2 \cdot (x - y) > w^3 \cdot (x - y)$ as desired. □

Fact 1: *Let w^1, w^2, e be a linearly independent vectors in $I\!R^n$ and x be any vector in the relative interior of S^{n-1}. Then, there exists y in the relative interior of S^{n-1} such that $w^1 \cdot (x - y) > 0 > w^2 \cdot (x - y)$.*

Fact 2: *Let w^1, w^2, e be a linearly independent vectors in $I\!R^n$ and x be any vector in the relative interior of S^{n-1}. Then, there exists y in the relative interior of S^{n-1} such that $w^1 \cdot (x - y) > 0 = w^2 \cdot (x - y)$.*

7.4 Proof of Theorem 3

Proof: First, we will prove the statement regarding preferences without self-control. Clearly, if (u, v) represents $\succeq$ and $u' = \alpha_u u + \beta_u$, $v' = \alpha_v v + \beta_v$ for some α_u, $\alpha_v > 0$ then (u', v') represents $\succeq$. To prove the converse, assume both (u, v) and (u', v') represent $\succeq$. The standard uniqueness argument of von Neumann-Morgenstern utility theory applied to singleton sets yields that u' is a positive affine transformation of u. Since $\succeq$ is regular, neither v nor v' is constant. Hence if v' is not a positive affine transformation of v, then there exists three lotteries $x^1, x^2, x^3 \in \Delta$ such that $(v(x^1), v(x^2), v(x^3)), (v'(x^1), v'(x^2), v'(x^3))$ and $e := (1, 1, 1)$ are linearly independent. Then, by Fact 1 in Section 7.4, there exists $x, y \in \Delta$ such that $v(x) > v(y)$ and $v'(x) < v'(y)$. Since $\succeq$ is regular and u, v, v' are linear, we can assume that $\{x\} \not\sim \{y\}$. If $\{x\} \succ \{y\}$ then, according to the representation (u, v), $\{x\} \sim \{x, y\}$ while according to the representation (u', v'), $\{x\} \succ \{x, y\}$, a contradiction. A symmetric argument yields a contradiction if $\{y\} \succ \{x\}$.

For self-control preferences, again it is easily verified that if (u', v') is a common positive affine transformation of (u, v) and (u, v) represents $\succeq$, then so does (u', v'). In proving the converse, note that the argument used in the no self-control case still applies to establish that $u' = \alpha_u u + \beta_u$, $v' = \alpha_v v + \beta_v$ for some α_u, $\alpha_v > 0$. To complete the proof we need to show that $\alpha_u = \alpha_v$. Since $\succeq$ is regular, it follows from our representation that there exists x, y such that $\{x\} \succ \{x, y\} \succ \{y\}$. Hence, for some $\gamma \in (0, 1)$, $\{\gamma x + (1-\gamma)y\} \sim \{x, y\}$. That is, $u(x) > u(y)$, $v(x) < v(y)$ and $\gamma u(x) + (1 - \gamma)u(y) = u(x) + v(x) - v(y)$. Hence, $(1-\gamma)[u(x)-u(y)] = v(y)-v(x)$. Similar calculations for the representation (u', v') yield $(1 - \gamma)\alpha_u[u(x) - u(y)] = \alpha_v[v(y) - v(x)]$. Therefore, $\alpha_u = \alpha_v$ as desired. □

7.5 Proof of Theorem 4

Definition: $x \succeq_T y$ *iff* $(\{x\}, x) \sim^* (\{x, y\}, x)$.

Lemma 10: *If $\succeq^*$ satisfies Axioms 2^*a, 2^*c, T1–3 then $\succeq_T$ can be represented by a continuous utility function v^*. Furthermore, if y^* maximizes v^* in A then*

$$(\{x, y^*\}, x) \sim^* (A, x)$$

for all $x \in A$.

Proof: By $T1$ and $T2$, $\succeq_T$ is complete. To prove transitivity, assume $x \succeq_T y \succeq_T z$. Then, $(\{x, z\}, x) \succeq^* (\{x, y, z\}, x)$ by $T1$. By $T3$, $(\{x, y, z\}, x) \succeq^* (\{x, y\}, x)$. But since $x \succeq_T y$ transitivity follows. Hence, $\succeq_T$ is a preference relation.

It follows from $T2$ and the upper semi-continuity of $\succeq^*$ that $\{y : x \succeq_T y\}$ is closed for all $x \in \Delta$. Also, since $\succeq^*$ is upper semi-continuous and satisfies 2^*c, $\{y : y \succeq_T x\}$ is closed. Hence, $\succeq_T$ is continuous. Therefore, by Debreu's celebrated theorem, $\succeq_T$ can be represented by a continuous function v^*.

Let y^* be any maximizer of $\succeq_T^*$ in A and choose any $x \in A$. Note that by $T1$, $(\{x, y^*\}, x) \succeq^* (A, x)$. Let B be any finite set such that $\{x, y^*\} \subset B \subset A$. If $(\{x, y^*\}, x) \succ^* (B, x)$ there exists a minimal set B^* such that $\{x^*, y^*\} \subset B^* \subset A$ and $(\{x, y^*\}, x) \succ^* (B^*, x)$. Hence, there exists $y \in B^* \backslash \{x, y^*\}$ such that $B^* = A^* \cup \{y\}$, $A^* \neq B^*$ and $(A^*, x) sp^* (B^*, x)$. By $T3$, $y \succ_T y^*$ a contradiction. Hence, $(\{x, y^*\}, x) \sim^* (B, x)$ for every finite B such that $\{x, y^*\} \subset B \subset A$. But since the set of all finite subsets of A is dense in the set of all subsets of A, $(\{x^*, y^*\}, x) \sim^* (A, x)$ follows from the upper semi-continuity of $\succeq^*$. $\square$

To prove Theorem 4, take any $A, B \in \mathcal{A}$. Let x^* satisfy $(\{x^*, y^*\}, x^*) \succeq^* (\{x, y^*\}, x)$ among x in $A \cup B$ and y^* be any maximizer of v^* in $A \cup B$. (Clearly, by Axiom 2^*a such an x^* exists.) Without loss of generality, assume $y^* \in B$. Let $\hat{x}$ satisfy $(B, \hat{x}) \succeq^* (B, x)$ for all $x \in B$. Then, it follows from the Lemma 10 that

$$\begin{aligned}(A \cup B, x^*) &\sim^* (\{x^*, y^*\}, x^*) \succeq^* (\{\hat{x}, y^*\}, \hat{x}) \succeq^* (B, \hat{x}) \\ &\succeq^* (B, x), \forall x \in B\end{aligned}$$

Hence, $A \cup B \succeq B$. Again, without loss of generality, assume, $x^* \in A$. Let $\hat{y}$ be a maximizer of v^* in A. Note that by $T1$ and Lemma 10, $(\{x^*, \hat{y}\}, x^*) \succeq^* (\{x^*, \hat{y}, y^*\}, x^*) \sim^* (\{x^*, y^*\}, x^*)$. Using Lemma 10 again, we obtain

$$\begin{aligned}(A, x^*) &\sim^* (\{x^*, \hat{y}\}, x^*) \succeq^* (\{x^*, y^*\}, x^*) \\ &\sim^* (A \cup B, x^*) \succeq^* (A \cup B, x),\end{aligned}$$

for all $x \in A \cup B$ and hence, $A \cup B \succeq B$. $\square$

7.6 Proof of Theorem 5

Proof: Let v^* be the continuous function that represents $\succeq_T$. We must show that v^* and v represent the same preference. Take any $x, y \in \Delta$.

Step 1: If x, y are interior then $v(x) \geq v(y)$ implies $v^*(x) \geq v^*(y)$.

Assume $u(x) > u(y)$. Since (u, v) represents $\succeq_1^*$, $\{x\} \sim_1^* \{x, y\}$. This implies $(\{x\}, x) \sim^* (\{x, y\}, x)$ and hence $v^*(x) \geq v^*(y)$. Assume $u(x) < u(y)$ and $v(x) > v(y)$. Then, $\{y\} \succ_1^* \{x, y\}$ and hence $(\{y\}, y) \succ^* (\{x, y\}, y)$ and therefore $v^*(x) \geq v^*(y)$. Assume $u(x) < u(y)$ and $v(x) = v(y)$. Assume $v^*(x) < v^*(y)$. Then, by regularity and interiority, there is an x' such that $v(x') > v(y)$ and $u(x') < u(y)$ and $v^*(x') < v^*(y)$. Then, since (u, v) represents $\succeq_1^*$ we have $\{y\} \succ_1^* \{x', y\}$ and hence $v^*(x') > v^*(y)$, a contradiction. Assume $u(x) = u(y)$. Since u is regular, it is not constant and we can find an interior z such that $u(z) > u(x)$. Then, we can apply the above argument to x_α, y where $x_\alpha := \alpha x + (1 - \alpha)z$ and $y_\alpha := \alpha y + (1 - \alpha)z$ to conclude that $v^*(x_\alpha) \geq v^*(y)$. Then, the continuity of v^* and v yields $v^*(x) \geq v^*(y)$.

Step 2: If x, y are interior then $v^*(x) \geq v^*(y)$ implies $v(x) \geq v(y)$.

Assume $u(x) > u(y)$. Since $(\{x\}, x) \sim^* (\{x, y\}, x)$ and since the extended preference induces (u, v) it follows that $v(x) \geq v(y)$.

Assume $u(x) < u(y)$ and $v^*(x) > v^*(y)$. Then, $(\{y\}, y) \succ^* (\{x, y\}, y)$ and since $(\{y\}, y) \succ^* (\{x\}, x) \succeq^* (\{x, y\}, x)$ therefore $v(x) \geq v(y)$.

Assume $u(x) < u(y)$ and $v^*(x) = v^*(y)$. Assume $v(x) < v(y)$. By regularity and interiority we can find a y' with $u(y') > u(y)$ and $v(x) < v(y') < v(y)$. Then, since $\succeq^*$ induces (u, v), $(\{y'\}, y') \succ^* (\{y, y'\}, y')$ and hence $v^*(y) > v^*(y')$. By Step 1, $v^*(x) \leq v^*(y')$ a contradiction.

Assume $u(x) = u(y)$. Since u is regular, it is not constant and we can find an interior z such that $u(z) > u(x)$. Then, we can apply the above argument to x_α, y where $x_\alpha := \alpha x + (1 - \alpha)z$ and $y_\alpha := \alpha y + (1 - \alpha)z$ to conclude that $v(x_\alpha) \geq v(y)$. Then, the continuity of v^* and v yields $v(x) \geq v(y)$.

Step 3. Assume x or y are arbitrary. First observe that by continuity and Step 1 $v^*(x) > (<) v^*(y)$ implies $v(x) \geq (\leq) v(y)$ and conversely $v(x) > (<) v(y)$ implies $v^*(x) \geq (\leq) v^*(y)$. Thus, we must show that $v^*(x) = v^*(y)$ iff $v(x) = v(y)$.

If $v(x) = v(y)$ then by linearity $v(\alpha x + (1 - \alpha)z) = v((\alpha y + (1 - \alpha)z)$. If z is interior then this implies $v^*(\alpha x + (1 - \alpha)z) = v^*((\alpha y + (1 - \alpha)z)$ by Step 1. Continuity of v^* now gives the desired result.

It remains to show that $v^*(x) = v^*(y)$ implies $v(x) = v(y)$. Assume (wlog) $v(x) > v(y)$ then $v(x) > v(x^k) > v(y^k) > v(y)$ for some interior x^k, y^k. Continuity and Case 1 implies $v^*(x) \geq v^*(x^k) > v^*(y^k) \geq v(y)$ contradicting $v^*(x) = v^*(y)$. □

7.7 Proof of Theorem 6

Since $\succeq^*$ is minimally congruent and induces (u, v) with u not constant, either (u, v) is regular, or, (u, v) has no preference for commitment. Assume the latter. Then, $c(\cdot, u + v) = c(\cdot, u) = c(c(\cdot, v), u)$. Suppose $x \in c(A, u)$. Suppose there exists $y \in A$ such that $v^*(y) > v^*(x)$. If $u(y) < u(x)$, then (u, v) has a preference for commitment at $\{x, y\}$, a contradiction. If $u(x) = u(y)$, then by minimal congruence we can find a x' such that $u(x') > u(x)$ and $v^*(x') < v^*(y)$. Then (u, v) has a preference for commitment at $\{x', y\}$, a contradiction. Hence $v^*(x) \geq v^*(y), \forall y \in A$. Therefore, by T3 $(\{x\}, x) \sim^* (A, x)$. Since

$\succeq_1^*$ is represented by (u,v) we have $(\{x\},x) \succeq^* (\{y\},y) \succeq^* (A,y)$ for all $y \in A$. That is, x is in $c(A,\succeq^*)$. Next suppose $x \in c(A,\succeq^*)$. Therefore $(\{x\},x) \sim^* (A,x) \succeq^* (A,y)$ for all $y \in A$. Hence $\{x\} \succeq_1^* A$. But, since $\succeq_1^*$ has no preference for commitment, this implies $x \in c(A,u)$.

Next assume (u,v) is regular. Also observe that Axiom 2c holds and hence $\succeq^*$ satisfies Axiom 2^*c. We may conclude that Theorem 5 holds.

Step 1: Suppose $\succeq$ is represented by (u,v) and has self- control. Then $c^*(A,\succeq^*) \subset c(A,u+v)$.

Assume $\hat{x} \notin c(A,u+v)$. If $\hat{x} \notin A$ then obviously $\hat{x} \notin c^*(A,\succeq^*)$. So, let $\hat{x} \in A$ and choose $y^* \in c(A,v)$ and $x^* \in c(A,u+v)$. If $u(y^*) > u(\hat{x})$ then (by Theorem 1) $A \succ_1^* \{\hat{x}\}$ and hence it follows that for some $x \in A$, $(A,x) \succ^* (\{\hat{x}\},\hat{x}) \succeq^* (A,\hat{x})$. Hence, $\hat{x} \notin c^*(A,\succeq^*)$, as desired.

Assume $u(y^*) \leq u(\hat{x})$. If $y^* \notin c(A,u+v)$ then $\{x^*,y^*\} \succ_1^* \{\hat{x},y^*\}$. But since (u,v) represents $\succeq_1^*$ there exists $z \in A$ with $(A,z) \succeq^* (A,x^*)$. Lemma 10 implies that $(A,x^*) \sim^* (\{x^*,y^*\},x^*) \succ^* (\{\hat{x},y^*\},\hat{x}) \succeq^* (A,\hat{x})$. Hence, $\hat{x} \notin c^*(A,\succeq^*)$ as desired. If $y^* \in c(A,u{+}v)$ then $v(y^*) > v(\hat{x})$ and by minimal congruence, we can find y close to y^* such that $u(y) < u(y^*)$ and $v(y) < v(y^*)$. If y is sufficiently close to y^* then $\{\hat{x},y,y^*\} \succ \{\hat{x},y\}$. It follows that there exists a $z \in A$ such that

$$(A,z) \succeq^* (\{\hat{x},y,y^*\},y^*) \succ^* (\{\hat{x},y\},\hat{x}) \succeq^* (\{\hat{x},y^*,\hat{y}\},\hat{x}) \sim^* (A,\hat{x})$$

where the last assertion follows from Lemma 10.

Step 2: Suppose $\succeq$ is represented by (u,v) and has self- control. Then $c(A,u+v) \subset c^*(A,\succeq^*)$.

Let $\hat{x} \in c(A,u+v)$. Clearly $\{\hat{x}\} \succeq A$ and $(\{\hat{x}\},\hat{x}) \succeq^* (A,x), \forall x \in A$. If $\hat{x} \in c(A,v) = c(A,v^*)$ then it follows from Lemma 10 that $(\{\hat{x}\},\hat{x}) \sim^* (A,\hat{x})$. Hence, $\hat{x} \in c^*(A,\succeq^*)$. So, assume $\hat{x}$ does not maximize v^* and pick some $y^* \in A$ that does. Again, by Lemma 10, to show that $\hat{x} \in c^*(A,\succeq^*)$ it is sufficient to prove $(\{\hat{x},y^*\},\hat{x}) \succeq^* (\{z,y^*\},z)$ for all $z \in A$. Since, $u(\hat{x}) + v(\hat{x}) \geq u(y^*) + v(y^*)$ and $v(\hat{x}) < v(y^*)$, we conclude $u(\hat{x}) > u(y^*)$. First, we demonstrate that $(\{\hat{x},y^*\},\hat{x}) \succeq^* (\{y^*\},y^*)$.

Case 1: If $u(\hat{x}) + v(\hat{x}) > u(y^*) + v(y^*)$ then $\{\hat{x}, y^*\} \succ \{y^*\}$. Since $\succeq_1^*$ is represented by (u, v), and satisfies $T1$, $(\{x, y^*\}, \hat{x}) \sim^* (\{x, y^*\}, x) \succ^* (\{y^*\}, y^*) \succeq^* (\{\hat{x}, y^*\}, y^*)$, for all $x \in \{\hat{x}, y^*\}$ as desired.

Case 2: If $u(\hat{x}) + v(\hat{x}) = u(y^*) + v(y^*)$, since u is not constant and $\succeq^*$ is minimally congruent, there exists $y \in \Delta$ sufficiently close to y^* such that $u(y) < u(y^*)$, $v(y) < v(y^*)$ and $v(\hat{x}) < v(y)$. Then, $(\{\hat{x}, y\}, \hat{x}) \succeq^* (\{\hat{x}, y\}, y) \sim^* (\{y\}, y)$ follows from the analysis of case 1 above. Axiom 2^*a and Axiom $2c$ now imply the desired conclusion if we let y converge to y^*.

Finally, we show that $(\{\hat{x}, y^*\}, \hat{x}) \succeq^* (\{z, y^*\}, z)$ for all $z \in A, z \neq y^*$. Since (u, v) represents $\succeq_1^*$, $\{\hat{x}, y^*\} \succeq_1^* \{z, y^*\}$ and there exist $x \in \{\hat{x}, y^*\}$ such that $(\{\hat{x}, y^*\}, x) \succeq^* (\{\hat{x}, y^*\}, z)$ for all $z \in \{\hat{x}, y^*\}$. Therefore, $(\{\hat{x}, y^*\}, \hat{x}) \succeq^* (\{z, y^*\}, z)$ as desired.

Step 3: Suppose $\succeq_1^*$ is represented by (u, v) and has no self-control. Then $c^*(A, \succeq^*) \subset c(c(A, v), u)$.

Assume $\hat{x} \notin c(c(A, v), u)$. If $\hat{x} \notin A$ we are done. Therefore, let $\hat{x} \in A$ and choose $y^* \in c(c(A, v), u)$. If $u(\hat{x}) < u(y^*)$ then $A \sim_1^* \{y^*\} \sim_1^* \{\hat{x}, y^*\} \succ_1^* \{\hat{x}\}$ and hence there exists $x \in A$ such that $(A, x) \succ^* (\{\hat{x}\}, \hat{x}) \succeq^* (A, \hat{x})$ by T1. Hence, $\hat{x} \notin c^*(A, \succeq^*)$ as desired.

If $u(\hat{x}) \geq u(y^*)$ then, since $\hat{x} \notin c^*(A, \succeq^*)$, $v(\hat{x}) < v(y^*)$. Therefore, since $\succeq^*$ satisfies minimal congruence we can find y close to y^* such that $u(y) < u(y^*)$, $v(y) < v(y^*)$ and $v(\hat{x}) < v(y)$. For y sufficiently close to y^*, $\{\hat{x}, y, y^*\} \succ \{\hat{x}, y\}$. It follows that there exists $x \in A$ such that

$$(A, x) \succeq^* (\{\hat{x}, y, y^*\}, y^*) \succ^* (\{\hat{x}, y\}, \hat{x}) \succeq^* (\{\hat{x}, y^*, \hat{y}\}, \hat{x}) \succeq^* (\{\hat{x}, y^*\}, \hat{x})$$

Step 4: Suppose $\succeq$ is represented by (u, v) and has no self- control. Then $c(c(A, v), u) \subset c^*(A, \succeq^*)$.

Let $\hat{x} \in c(c(A, v), u)$. Since (u, v) represents $\succeq$, $\{x\} \sim A$. And, since $\succeq^*$ induces $\succeq$, $(\{\hat{x}\}, \hat{x}) \succeq^* (A, x), \forall x \in A$. But we have shown that v and v^* represent the same preference. Therefore, it follows from Lemma 10 that $(\{\hat{x}\}, \hat{x}) \sim^* (A, \hat{x})$. □

7.8 Proof of Theorem 7

Lemma 11: *Let $z^1, z^2, e := (1, 1, \ldots, 1)$ be linearly independent vectors in $\mathbb{R}^n$. Suppose $\alpha < 0$ or $\beta < 0$. Then, for x in the relative interior of Δ, there exists y in the relative interior of Δ such that $z^1 \cdot (x - y) < 0$, $z^2 \cdot (x - y) < 0$ and $(\alpha z^1 + \beta z^2) \cdot (x - y) > 0$.*

Proof: Without loss of generality assume that $\alpha < 0$. If $\beta \leq 0$ set $w^1 = -z^1$ and $w^2 = z^2$ and apply Fact 1 in Section 7.4 to obtain the desired y.

If $\beta > 0$ then set $w^1 = -z^1$, $w^2 = (\alpha/2)z^1 + \beta z^2$ and apply Fact 2 in Section 7.4 to obtain the desired y. □

Lemma 12: *Suppose (u_1, v_1) represents $\succeq_1$ and (u_2, v_2) represents $\succeq_2$. Then $\succeq_1$ has greater preference for commitment than $\succeq_2$ iff there exist non-negative, full rank matrix $\mathbf{\Theta}$ and a $\lambda \in I\!R^2$ such that*

$$\begin{pmatrix} \hat{u}_2(x) \\ v_2(x) \end{pmatrix} = \mathbf{\Theta} \cdot \begin{pmatrix} u_1(x) \\ v_1(x) \end{pmatrix} + \lambda$$

for all $x \in \Delta$.

Proof: We prove the only if part of the Lemma in 2 steps. Assume the hypothesis of the lemma holds. Let T^n denote any n element subset of Δ. Define $\Delta(T^n) := \{x \in \Delta : x = \sum_{i=1}^n \alpha_i x^i\}$ and let $\mathcal{A}(\Delta(T^n))$ denote the set of all compact subsets of $\Delta(T^n)$. For $j = 1, 2$, let u^j, v^j denote the row vectors $(u_j(x^1), \ldots, u_j(x^n))$, $(v_j(x^1), \ldots, v_j(x^n))$, respectively.

Step 1: If the restriction of $\succeq_i$ for $i = 1, 2$ is regular then there exist non-negative, full rank matrix $\mathbf{\Theta}$ and a $\lambda \in I\!R^2$ such that

$$\begin{pmatrix} u^2 \\ v^2 \end{pmatrix} = \mathbf{\Theta} \cdot \begin{pmatrix} u^1 \\ v^1 \end{pmatrix} + \lambda$$

Proof of Step 1: First, we prove that the vectors u^2, u^1, e, v^2 are not linearly independent. If they were, by Lemma 9 we could find x, y in the relative interior of Δ such that $u^2 \cdot x > u^2 \cdot y, u^1 \cdot x = u^1 \cdot y, v^2 \cdot x < v^2 \cdot y$. But then we would have $\succeq_2$ display preference for commitment at $\{x, y\}$ and $\succeq_1$ have no preference for commitment $\{x, y\}$, a contradiction. A similar argument establishes that u^2, v^1, e, v^2 and e are linearly dependent. Since $\succeq_2$ is regular, u^2, v^2, e are linearly independent. And since $\{u^2, u^1, e, v^2\}$, $\{u^2, v^1, e, v^2\}$ are two linearly dependent sets of vectors, we can write $u^1 = m_{1,1}u^2 + m_{1,2}v^2 + m_{1,3}e$ and $v^1 = m_{2,1}u^2 + m_{2,2}v^2 + m_{2,3}e$ for $m_{k,l} \in I\!R$. Since $\succeq_1$ is regular the matrix $\mathbf{M} := \begin{pmatrix} m_{1,1}, m_{1,2} \\ m_{2,1}, m_{2,2} \end{pmatrix}$ is non-singular. Let $\mathbf{\Theta} = \mathbf{M^{-1}}$ and $\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = -\mathbf{M^{-1}} \cdot \begin{pmatrix} m_{1,3} \\ m_{2,3} \end{pmatrix}$. Obviously, $\mathbf{\Theta}$ is non-singular.

It remains to be shown that $\mathbf{\Theta}$ is non-negative. If $\mathbf{\Theta}$ is not non-negative, then, by Lemma 11, there exist x, y in the relative interior of Δ such that $u^1 \cdot (x - y) < 0, v^1 \cdot (x - y) <$

0 and $w \cdot (x-y) > 0$ for $w = u^2$ or $w = v^2$. If ($u^2 \cdot (x-y) > 0$ and $v^2 \cdot (x-y) < 0$) or ($u^2 \cdot (x-y) < 0$ and $v^2 \cdot (x-y) > 0$) then $\succeq_2$ has a preference for commitment at $\{x, y\}$ while $\succeq_1$ does not, a contradiction. So, we assume either ($u^2 \cdot (x-y) > 0$ and $v^2 \cdot (x-y) \geq 0$) or ($u^2 \cdot (x-y) \geq 0$ and $v^2 \cdot (x-y) > 0$). In the former case, by Fact 1, there exists z in the relative interior of Δ such that $u^2 \cdot (x-z) > 0 > v^2 \cdot (x-z)$. Then, for $\alpha \in (0,1)$ sufficiently small $\{x, \alpha z + (1-\alpha)x, y\} \sim_1 \{y\} \succeq_1 A$ for all $A \subset \{x, \alpha z + (1-\alpha)x, y\}$. However, $\{x\} \succ_2 \{x, \alpha z + (1-\alpha)x, y\}$ contradicting the hypothesis that $\succeq_1$ is greater preference for commitment than $\succeq_2$. Similarly, if $u^2 \cdot (x-y) \geq 0$ and $v^2 \cdot (x-y) > 0$ then choose z such that $v^2 \cdot (x-z) > 0 > u^2 \cdot (x-z)$ and choose $\alpha \in (0,1)$ sufficiently small so that $\{x, \alpha z + (1-\alpha)x, y\} \sim_1 \{y\} \succeq_1 A$ for all $A \subset \{x, \alpha z + (1-\alpha)x, y\}$. But $\{\alpha z + (1-\alpha)x\} \succ_2 \{x, \alpha z + (1-\alpha)x, y\}$ which again yields a contradiction.

To conclude the only if part, note that since $\succeq$ is regular, there exists a three element subset T_1^0 of Z such that $\succeq_1$ restricted to $\mathcal{A}(\Delta(T_1^0))$, is regular. Similarly, there exists some T_2^0 such that $\succeq_2$ restricted to $\mathcal{A}(\Delta(T_1^0))$ is regular. Let $T^0 := T_1^0 \cup T_2^0$ and construct a nested sequence of finite sets T^n such that $\cup_n T^n$ is a dense subset of Z. Apply Step 1 to the restrictions of $\succeq_1$ and $\succeq_2$ to the set of all probability distributions with prizes in T^n to get the $\mathbf{\Theta^n}, \lambda^n$ that satisfy the conditions of the Lemma, for every $x \in \mathcal{A}(\Delta(T^n))$. But then, $\mathbf{\Theta^1} = \mathbf{\Theta^n}$ and $\lambda^1 = \lambda^n$ for all n. Since, u_2, v_2, u_1, v_1 are all continuous and $\cup_n \mathcal{A}(\Delta(T^n))$ is dense in $\mathcal{A}$, it follows that the equation in the statement of Step 1 holds for all $x \in \Delta$.

To prove the if part of the Lemma, assume that there is a, non-negative matrix $\mathbf{\Theta}$ and $\lambda \in \Re^2$ such that $\mathbf{\Theta}$ for all $x \in \Delta$

$$\begin{pmatrix} u_2(x) \\ v_2(x) \end{pmatrix} = \mathbf{\Theta} \cdot \begin{pmatrix} u_1(x) \\ v_1(x) \end{pmatrix} + \lambda$$

Suppose $\succeq_1$ has no preference for commitment at A. Then, there exists $x \in A$ such that $u_1(x) \geq u_1(y)$ and $v_1(x) \geq v_1(y)$ for all $y \in A$. Hence, $u_2(x) = \theta_{1,1} \cdot u_1(x) + \theta_{1,2} \cdot v_1(x) \geq \theta_{1,1} \cdot u_1(y) + \theta_{1,2} \cdot v_1(y) = u_2(y)$ and $v_2(x) = \theta_{2,1} \cdot u_1(x) + \theta_{2,2} \cdot v_1(x) \geq \theta_{2,1} \cdot u_1(y) + \theta_{2,2} \cdot v_1(y) = v_2(y)$ for all $y \in A$. Therefore, $\succeq_2$ has no preference for commitment at A. □

To conclude the proof, suppose the desired α, β and γ exists. Then, clearly $\mathbf{\Theta}$ and λ as specified by Lemma 2 exist and hence by Lemma 12, $\succeq_1$ has greater preference for commitment than $\succeq_2$. Conversely, assume that $\succeq_1$ has greater preference for commitment than $\succeq_2$ and take any representation u_2, v_2 of $\succeq_2$. By Lemma 12, $\mathbf{\Theta}$ and λ with the desired properties exist. Let $\theta_{i,j}$ denote the i, j'the entry of $\mathbf{\Theta}$. Let $\alpha := \frac{\theta_{1,1}}{\theta_{1,1}+\theta_{1,2}}$, $\beta := \frac{\theta_{2,1}}{\theta_{2,1}+\theta_{2,2}}$ $\gamma := \frac{\theta_{2,1}+\theta_{2,2}}{\theta_{1,1}+\theta_{1,2}}$, $u_2 = \alpha u_1 + (1-\alpha)v_1$ and $v_2 = \beta u_1 + (1-\beta)v_1$. By Theorem 3, since (u_2, v_2) represents $\succeq_2$ so does $(u_2, \gamma v_2)$. $\square$.

References

1. Akerlof, G. and W. Dickens, “The Economic Consequences of Cognitive Dissonance,” *American Economics Review* 72 (1982), pp. 543–569.

2. Brehm, J., “Postdecision Changes in the Desirability of Alternatives,” *Journal of Abnormal and Social Psychology* 52 (1956), pp. 384–389.

3. Dekel, E., B. Lipman and A. Rustichini, “A Unique Subjective State Space for Unforeseen Contingencies,” Discussion Paper, Northwestern University (1997).

4. Kreps, D., “A Preference for Flexibility,” *Econometrica* 47 (1979) 565–576.

5. Kreps, D. and E. Porteus, “Temporal Resolution of Uncertainty and Dynamic Choice Theory,” *Econometrica* 46 (1978), 185–200.

6. Kirby, K. and R. J. Herrnstein, “Preference Reversals Due to Myopic of Delayed Reward,” *Psychological Science* 6 (1995) 83–89.

7. Laibson, D., “Golden Eggs and Hyperbolic Discounting,” *Quarterly Journal of Economics* 112 (1997), 443–477.

8. Loewenstein G., “Out of Control: Visceral Influences on Behavior,” *Organizational Behavior and Human Decision Processes*, 65 (1996), 272–292.

9. Loewenstein G. and D. Prelec, “Anomalies in Intertemporal Choice: Evidence and an Interpretation,” *Quarterly Journal of Economics* 107 (1992), 573–598.

10. Rabin M., “Psychology and Economics,” working paper, Department of Economics, University of California , Berkeley, January 1997.

11. Strotz, R. H., “Myopia and Inconsistency in Dynamic Utility Maximization,” *Review of Economic Studies* 1955 23 (3): 165–180.