TEMPTATION AND SELF-CONTROL

By Faruk Gul and Wolfgang Pesendorfer[1]

We study a two-period model where ex ante inferior choice may tempt the decision-maker in the second period. Individuals have preferences over sets of alternatives that represent second period choices. Our axioms yield a representation that identifies the individual's commitment ranking, temptation ranking, and cost of self-control. 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 them to our representations.

Keywords: Dynamic choice, preference for commitment, set betweenness.

1. INTRODUCTION

Consider an individual who must decide what to eat for lunch. She may choose a vegetarian dish or a hamburger. In the morning, when she feels no hunger, she prefers the healthy, vegetarian dish but at lunchtime, she experiences a craving for a hamburger.[2] To lessen the impact of her lunchtime craving, the decision-maker may seek to limit the options at lunchtime. For example, she may choose a vegetarian restaurant. When this is not possible and the individual is confronted with a menu that includes both the vegetarian meal and the hamburger, she may exercise *self-control*, that is, resist the craving for the hamburger and choose the vegetarian meal. Individuals will frequently use both remedies. In the example, the individual may visit a vegetarian restaurant to exclude the hamburger from the option set. However, even the vegetarian restaurant offers an unhealthy dessert 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 she must choose between x and y at lunchtime is described by $\{x, y\}$.

[2] See Loewenstein (1996) for a discussion of the impact of short-run cravings and other visceral factors on behavior.

When there is no choice at lunchtime, the individual ranks x above y and therefore $\{x\}$ is strictly preferred to $\{y\}$. The alternative y is tempting when it is on the menu. We capture this with a strict preference for $\{x\}$ over $\{x, y\}$. Confronted with the menu $\{x, y\}$ some agents succumb to the temptation and choose y. We describe this by a preference for which $\{x, y\}$ is indifferent to $\{y\}$. Alternatively, an individual may exercise self-control at $\{x, y\}$. Self-control is captured by a strict preference for $\{x, y\}$ over $\{y\}$. Self-control at $\{x, y\}$ is therefore identified with $\{x\} \succ \{x, y\} \succ \{y\}$. A decision-maker with self-control chooses the vegetarian meal and is better off than she would be if only the hamburger was on the menu. On the other hand, the availability of the hamburger creates a temptation and therefore the individual is worse off than she would be if she were committed to the vegetarian meal, x.

In our model, the decision-maker 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 agent's 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 compromise 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 "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 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, she chooses the lottery that maximizes v but evaluates these choices using u. Such preferences are represented by a utility function of the form

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

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 3, 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.

Up to this point, we have analyzed and represented period 1 preferences. In Section 4 we analyze period 2 behavior. To do this, we assume that the agent has preferences $\succeq^*$ on the set $\mathscr{S} := \{(A, x): x \in A\}$ 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, the extended preference $\succeq^*$ induces a period 1 preference $\succeq_1^*$ as follows: $A \succeq_1^* B$ if and only if there is an $x \in A$ with $(A, x) \succeq^* (B, y)$ for all $y \in B$.

The extended preference $\succeq^*$ allows us to formalize our notion of temptation. We say that y *tempts* x if $(\{x\}, x) \succ^* (\{x, y\}, x)$. We assume the following: if y tempts x, then x does not tempt y; choosing x from $A \cup B$ is not better than choosing x from A and may be worse only if something in B tempts every option in A. These axioms are equivalent to the assumption that "tempts" is a preference relation and that the utility of a fixed choice is affected by the set A only through the most tempting element in A.

Theorem 5 shows that if the extended preference $\succeq^*$ satisfies these assumptions and a weak form of continuity, then the induced first period preference satisfies *Set Betweenness*. This justifies our interpretation of *Set Betweenness* as describing an individual who struggles with temptation. Theorem 7 shows that if, in addition, the induced period 1 preferences, $\succeq_1^*$, satisfy our earlier axioms, then the second period choice behavior is as suggested by the representation theorems: 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 "selves" 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 $\succeq_1$ a pair of utility functions (u_i, v_i). In Theorem 8 we utilize this representation and characterize greater preference for commitment in terms of the (u_i, v_i). The theorem shows that $\succeq_1$ has greater preference for commitment than $\succeq_2$ if u_1 and v_1 are "further apart" than u_2 and v_2, that is, the indifference curves 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 9 we show 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 in period 2, the individual 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 preferences 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 two papers on preference for flexibility, Kreps (1979) and Dekel, Lipman, and Rustichini (2001). Following Kreps, we study preferences over sets of alternatives. Dekel, Lipman, and Rustichini 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.

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 that constitute the feasible choices at 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 σ-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 $\mathcal{A} \times \mathcal{A}$.[3] We endow $\mathcal{A}$ with the topology generated by the (*Hausdorff*) metric

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

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* $\alpha A + (1-\alpha)C \succ \alpha B + (1-\alpha)C$.

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 preferences to the set of lotteries over $\mathcal{A}$. Assume $A \succ B$ and suppose the decision-maker has the choice between the lottery that yields A with probability α and C with probability $1-\alpha$ (denoted $\alpha \circ A + (1-\alpha) \circ C$) 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 $\alpha \circ A + (1-\alpha) \circ C \succ \alpha \circ B + (1-\alpha) \circ C$. 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 this 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 (i) she satisfies

---

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

the usual independence axiom and (ii) she is indifferent as to when uncertainty is resolved.[4]

A "standard" decision-maker who experiences no temptation and therefore 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 such that U defined by $U(A) := \max_{x \in A} u(x)$ represents $\succeq$.[5]

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 $\succeq$ 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 $\succeq$ has a preference for commitment at some $A \in \mathscr{A}$.

The following axiom permits a 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 causes temptation. We assume that temptation is utility decreasing, that is, an alternative that is not chosen cannot increase the utility of the decision-maker. Furthermore, we assume that lotteries can be ranked according to how tempting they are and that only the most tempting option available affects the agent's utility.

In Section 4 we give a precise statement of these conditions and Theorem 5 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. Since the utility cost of temptation only depends on the most tempting alternative, y, choosing z in $A \cup B$ leads to the same utility as choosing z in B and $A \cup B \succeq B$ follows.

One can 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

---

[4] Dekel, Lipman, and Rustichini (2001) provide a related justification for *Independence* in an earlier version of their paper.

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

has a cumulative effect so that larger sets 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 $A \succ A \cup B$, $B \succ A \cup B$. We rule out these more elaborate formulations of temptation, as well as other deviations from the standard model, to stay close enough to the standard model so that the difference in behavior can be attributed solely to the presence of temptation.

We say that the function $U: \mathscr{A} \to \mathbb{R}$ 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 $A, B \in \mathscr{A}$. Similarly, the function $u: \Delta \to \mathbb{R}$ 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 $A \in \mathscr{A}$ 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. For the only if part, Lemma 1 in the Appendix shows that under Axioms 1–3 we can represent the preferences by a continuous, linear utility function U. The main steps in the remainder of the proof are (i) identifying u, (ii) identifying v, (iii) proving that the representation holds for two-element sets, and (iv) proving that the representation holds for arbitrary sets if it holds for two-element sets.

We define u as the utility U assigns to singleton sets, i.e., $u(x) = U(\{x\})$. Since U is linear, so is u. 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

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

Lemma 4 shows that the v so defined is indeed linear, and independent of the particular choice of a and b.

With the u and v above our representation holds for two element sets. The following Lemma demonstrates that each finite set is indifferent to an appro-

priate two-element subset. Thus, we can extend the representation to arbitrary finite sets, and by continuity, to arbitrary compact sets.[6]

LEMMA: *Let U be a function that represents some $\succeq$ satisfying Axiom 4. If $A \in \mathscr{A}$ is a finite set, then $U(A) = \max_{x \in A} \min_{y \in A} U(\{x, y\})$.*

PROOF: Suppose $\bar{u} = \max_{x \in A} \min_{y \in A} U(\{x, y\})$ and (x^*, y^*) is a solution to this problem. We first show that $U(A) \geq \bar{u}$. 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*, $U(\cup_{x \in A} \{x, y_x\}) = U(A) \leq \bar{u}$.   Q.E.D.

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 extend preferences to include choices from the set and show that behavior in the second period is indeed as suggested by the representation. In anticipation of this result, we refer to the maximization of $u + v$ as the choice behavior in period 2.

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

---

[6] Dekel, Lipman, and Rustichini (2001) 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.

while $A \cup B \succ B$ captures the fact that the agent resists this temptation. That is, the agent uses self-control.

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 $C \in \mathscr{A}$.

In the following theorem three notions of self-control are proven to be equivalent: (i) our definition based on preferences over sets; (ii) an intuitive definition of self-control as the 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. For an upper semicontinuous function $f: \mathscr{A} \to \mathbb{R}$, we define $c(A, f) := \{x \in A : f(x) \geq f(y) \; \forall \, y \in A\}$. Thus, $c(\cdot, u + v)$ describes the second period choice behavior whereas $c(\cdot, v)$ denotes the lotteries that are most tempting.

THEOREM 2: *Suppose U, u, v are continuous linear functions and $\succeq$ is represented by U where $U(A) := \max_{x \in A}\{u(x) + v(x)\} - \max_{y \in A} v(y)$. Then, the following three statements are equivalent*:
  (i) *$\succeq$ has self-control at A.*
 (ii) *$c(A, v) \cap c(A, u + v) = \varnothing$.*
(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 $x \in c(A, v) \cap c(A, u + v)$. It follows from the representation that $U(A) = u(x) = U(D)$ 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 $c(A, v) \cap c(A, u + v) = \varnothing$. Let $B = c(A, u + v)$. Clearly, $c(B, u + v) = c(A, u + v)$ and it follows from the representation that $B \succ A$. 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 $B' \succ B' \cup C \succ C$.   *Q.E.D.*

The characterization of self-control in Theorem 2 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 (see, for example, Rabin (1998)). 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 she is tempted by y (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. PREFERENCES WITH AND WITHOUT SELF-CONTROL

In this section, we generalize the preferences studied previously to include decision-makers who cannot resist temptation. At time 2, such decision-makers choose according to their temptation rankings. 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 l–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 k 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. In the limit, as $k \to \infty$, the agent does not resist temptation. However, in this limit, preferences may not satisfy *Strong Continuity*. To see this, consider the sets $\{x, y\}$, $\{x, y'\}$ where y and y' are very close. Assume that $u(x) > u(y)$ and $u(x) > u(y')$. Further $v(x) > v(y)$ and $v(y') > v(x)$. When the agent cannot resist temptation, it follows that $\{x, y\} \sim \{x\} \succ \{y'\} \sim \{x, y'\}$ and even though the sets $\{x, y\}$ and $\{x, y'\}$ are close, their utilities are not. We therefore must weaken continuity to include decision-makers without self-control.

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

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

AXIOM 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.

We interpret individuals without self-control as having a prohibitively high cost of self control. Therefore, when the most tempting alternative is not unique, a decision-maker with no self-control resolves this indifference to maximize the commitment ranking. Axiom 2a implies this tie-breaking rule. To see this, 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 is 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. Axiom 2c can be omitted if the representation is restricted to all finite subsets of Δ or if the set Z is finite.

THEOREM 3: *The binary relation $\succeq$ satisfies Axioms* 1, 2a–c, 3, 4 *if and only if there are continuous linear functions* u, v *such that either the function* U *defined as*

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

for all $A \in \mathcal{A}$ *or the function* U *defined as*

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

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

PROOF: See Appendix.

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

$$A \sim A \cup B \quad \text{or} \quad B \sim A \cup B.$$

Thus, if $A \succ A \cup B$ and hence adding B to A provides temptation for the decision-maker, it follows that $B \sim A \cup B$. That is, the individual succumbs whenever 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 if the function* U *defined as*

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

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

To see how the proof of Theorem 3 works, first assume the agent has self-control at some $\{x, y\}$, i.e., $\{x\} \succ \{x, y\} \succ \{y\}$. Claim 1 in the Appendix shows that in this case the representation of Theorem 1 continues to hold for finite sets. Next, assume that the agent has no self-control at any set $\{x, y\}$, i.e., $\{x\} \sim \{x, y\}$ or $\{y\} \sim \{x, y\}$ for all x, y. Then, Claims 2 and 3 establish that the agent has no self-control. We define the temptation ranking R for the agent without self-control as follows: Suppose $\{x\} \not\succ \{y\}$. 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 3 *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 3. Temptation preferences that have the first type of representation 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*.

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 a positive affine transformation of v (i.e. $u = \alpha v + \beta$ for some $\alpha > 0$, $\beta \in \mathbb{R}$). 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 $\alpha \leq 0$, $\beta \in \mathbb{R}$. 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 4 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 4: *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, $\beta_v \in \mathbb{R}$. 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 $\alpha_u, \alpha_v > 0$ and $\beta_u, \beta_v \in \mathbb{R}$.*

4. EXTENDED PREFERENCES, TEMPTATION AND SECOND PERIOD CHOICE

In this section, we study the second period choice behavior associated with the first period preferences over sets analyzed above. Let $\mathscr{S} := \{(A, x) \in \mathscr{A} \times \Delta : x \in A\}$ and denote by $\succeq^*$ the agent's preference relation on $\mathscr{S}$. For a given first period choice A, the relation $\succeq^*$ induces a preference over lotteries $x \in A$. Observe that the first period choices A and B may lead to different conditional preferences

and therefore may change the agent's ranking of x and y for $x, y \in A \cap B$.[7] In the first period, the agent evaluates sets anticipating her choice in the second period. For an extended preference $\succeq^*$, the induced first period preference, denoted $\succeq_1^*$, is defined as follows: $A \succeq_1^* B$ if and only if there exists $x \in A$ such that $(A, x) \succeq^* (B, y)$ for all $y \in B$.

We use extended preferences to analyze the Set Betweenness Axiom. Extended preferences also enable us to relate the first period preferences over sets characterized by Theorems 1 and 3 to the second period choices from these sets. To see how this can be done, suppose $\succeq^*$ can be represented by the utility function

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

In period 2, this individual chooses a lottery $x \in A$ to maximize $U^*(A, \cdot)$ and in period 1, she evaluates the set A according to $\max_{x \in A}\{u(x) + v(x)\} - \max_{y \in A} v(x)$. Hence, the induced first period preference over sets is the self-control preference represented by (u, v).

For the case when the agent has no self-control, 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, this agent 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 U^* 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 $\mathscr{S}$ in many ways. Moreover, the testable implications of the extension $\succeq^*$ are captured by the induced preferences over $\mathscr{A}$ 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) and (A, z). For this reason, we use preferences on $\mathscr{A}$ in our representation theorems. Nevertheless, 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 sets. We assume that $\succeq^*$ satisfies the following axioms.

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

AXIOM 2*a (*Upper Semi-Continuity*): *The sets $\{(A, x): (A, x) \succeq^* (B, y)\}$ are closed.*

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

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

AXIOM T2: $(\{x\}, x) \succ^* (\{x, y\}, x)$ *implies* $(\{x, y\}, y) \succeq^* (\{y\}, y)$.

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

Axioms T1–3 describe how the set of alternatives influences the agent's preference when the choice in the second period is fixed. In our model, this influence comes about because the agent may be tempted by alternatives that are not chosen. Axiom T1 says that eliminating temptations cannot make the agent worse off. To understand Axiom T2, observe that $(\{x\}, x) \succ^* (\{x, y\}, x)$ is a situation where the choice x is less attractive when y is an option than when x is the only possible choice. Therefore, y *tempts* x. Hence, T2 says that if y tempts x, then x does not tempt y. Finally, T3 requires that unless y tempts every option in A, adding y to A does not make the agent worse off. T1–3 imply that the relation $\succeq_T$ defined by $x \succeq_T y$ iff $(\{x\}, x) \not\succ^* (\{x, y\}, x)$ is a preference relation. Thus, we can interpret T3 as saying that, for a fixed choice, the set A matters only through its most tempting ($\succeq_T$-maximal) element.

Theorem 5 demonstrates that Axioms 1, 2*a, T1–3 imply *Set Betweenness* and hence provides a foundation for our interpretation of *Set Betweenness* as the implication of temptation on the induced period one preferences.

THEOREM 5: *If $\succeq^*$ satisfies* 1*, 2*a, *T1–3, then* $\succeq_1^*$ *satisfies Set Betweenness.*

PROOF: See Appendix.

To see the intuition for Theorem 5, 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 (i.e., $\succeq_T$ maximal). 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 $A \cup B \succeq_1^* B$ follows.

Theorem 6 shows that if the induced first period preference is a temptation preference, then $v(y) > v(x)$ if and only if y tempts x. This confirms our earlier interpretation of v as representing the temptation ranking. To ensure uniqueness of v, we require the induced preference $\succeq_1^*$ to be regular, that is, u and v must be linearly independent. In addition, 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 $A \succeq \{y\}$ for all A implies $(\{x, y\}, x) \succeq^* (\{x\}, x)$ for all $x \in \Delta$.

THEOREM 6: *Suppose $\succeq^*$ satisfies 2^*a, T1–3 and $\succeq_1^*$ is a regular temptation preference that satisfies minimal congruence. If $\succeq_1^*$ can be represented by (u, v), then $(\{x\}, x) \succ^* (\{x, y\}, x)$ iff $v(y) > v(x)$.*

PROOF: See Appendix.

In Theorem 6, we can replace the assumption of minimal congruence with the hypothesis that x or y is not u maximal. Thus, minimal congruence is used only in the case where x and y are best elements for the commitment ranking.

To gain intuition for Theorem 6, 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, the fact that $\succeq_1^*$ is represented by (u, v) and T1 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 $(\{x\}, x) \succ^* (\{x, y\}, x)$. Hence, $v(y) > v(x)$ if and only if y tempts x.

In period 2, the individual maximizes the conditional preference induced by $\succeq^*$. Thus, the second period choice function for the preference $\succeq^*$, denoted by $c^*(\cdot, \succeq^*)$, satisfies

$$c^*(A, \succeq^*) := \{x \in A: (A, x) \succeq^* (A, y) \ \forall \ y \in A\}.$$

Theorem 7 shows that an individual whose induced first period preference is a temptation preference indeed behaves as suggested by the representation theorems: if she has self control, she maximizes $u + v$; if she has no self-control, she lexicographically maximizes v and then u.

THEOREM 7: *Suppose $\succeq^*$ satisfies 2^*a, T1–3 and is minimally congruent. Suppose also that $\succeq_1^*$ is a temptation preference with representation (u, v) with u not constant. Then, either $\succeq_1^*$ is a self-control preference and*

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

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

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

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

PROOF: See Appendix.

Suppose minimal congruence fails but the other assumptions in Theorem 7 are satisfied. If the set of prizes, Z, is finite, then the conclusion of Theorem 7 holds for every A that is a subset of the relative interior of Δ. Thus, even if minimal congruence is violated, the agent behaves as suggested by the representation theorems for a typical A.

To gain intuition for Theorem 7 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 nonempty. Hence we must show that if $\hat{x} \in A$ is not equal to x^*, then $\hat{x} \notin c^*(A, \succeq^*)$. Suppose $\hat{x} \neq x^*$ and hence $\hat{x}$ does not maximize $u + v$ in A. Then, $A \succ_1^* \{\hat{x}, y^*\}$ since (by construction)

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

The fact that $A \succ_1^* \{\hat{x}, y^*\}$ implies that there is a $z \in A$ such that $(A, z) \succ^* (\{\hat{x}, y^*\}, \hat{x})$. Axiom T1 implies $(\{\hat{x}, y^*\}, \hat{x}) \succeq^* (A, \hat{x})$ and therefore $\hat{x} \notin c^*(A, \succeq^*)$, as desired. Continuity and the axioms T1, T3 are used to show that all $u + v$ maximizers are in $c^*(\cdot, \succeq^*)$.

In the choice behavior characterized in Theorem 7, the individual maximizes the same utility function independent of the set, A, chosen in the first period. This does not follow from Axioms 2^*a, T1–3 alone but also requires the assumption that the induced first period preference satisfies *Independence*. To see this, 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 define comparative measures of preference for commitment and of self-control. We provide a characterization of these measures in terms of our representations.

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

Theorem 8 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 8: *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 curves 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 8, 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 combination 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 $\succeq_1$ has preference for commitment but $\succeq_2$ does not.

DEFINITION: *The preference* $\succeq_1$ *has more self-control than* $\succeq_2$ *if, for all* $A \in \mathscr{A}$, $\succeq_2$ *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 8.

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 2, the proof is analogous to the proof of Theorem 8 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 curves 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 9 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, v_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 have greater preference for commitment than $\succeq_2$ and yet have more self-control than $\succeq_2$.

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 to quantify the cost of self-control as a utility penalty that applies whenever the ultimate choice is not the most tempting one. We can therefore 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. It may be that $A \succ A \cup \{x\}$ even though x is ultimately not chosen from the set $A \cup \{x\}$ and hence, removing a nonbinding constraint may reduce an agent's utility. 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, that is, the representative agent must hold zero units of the liquid asset in some state of the world. In our model 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. In a model with 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.

Department of Economics, Princeton University, Princeton, NJ 08544, U.S.A.; fgul@princeton.edu; http://www.princeton.edu/~fgul/

and

Department of Economics, Princeton University, Princeton, NJ 08544, U.S.A.; pesendorfer@princeton.edu; http://www.princeton.edu/~pesendor/

APPENDIX

The following observation regarding compact subsets of metric spaces will be used throughout the Appendix. The proof is simple and omitted.

LEMMA 0: *If X is a compact set in a metric space, then there exists a sequence of subsets X_k of X such that each X_k has k elements and the sequence X_k converges to X in the Hausdorff metric.*

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 3 is equally straightforward and hence omitted.

A. PROOF OF THEOREM 1

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

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

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

Note that $L(\mathscr{A})$ is a mixture space. Moreover, Axiom 2a and 2b imply that $\dot{\succeq}$ satisfies von Neumann-Morgenstern continuity. That is, for all $\pi, \rho, \mu \in L(\mathscr{A})$ 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 $W: \mathscr{A} \to \mathbb{R}$ such that

$$\pi \dot{\succeq} \rho \iff \sum_A W(A)\pi(A) \geq \sum_A W(A)\rho(A).$$

By construction, $A \succeq B$ if and only if $W(A) \geq W(B)$ and $\alpha A + (1-\alpha)B \dot{\sim} \alpha \circ A + (1-\alpha) \circ B$, where $\alpha \circ A + (1-\alpha) \circ B$ is the element of $L(\mathscr{A})$ that yields A with probability α and B with probability $1 - \alpha$. Therefore

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

By 2c, W restricted to singleton sets is continuous. To see this, note that Axioms 1, 2a, 2c, 3 imply that there is a continuous linear representation of the restriction of $\succeq$ to singleton sets. W restricted to singleton sets represents this preference relation and is therefore an affine transformation of a continuous function. Hence, W restricted to singleton sets is continuous. Setting $U = W$ then yields the result. *Q.E.D.*

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 $A \in \mathscr{A}$ 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^* such 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 solution 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 $\{y^*, x^*\}$ the minmax problem.   Q.E.D.

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

$$U(\{x\}) > U(\{x, y\}) > U(\{y\}), \quad U(\{a\}) > U(\{a, b\}) > U(\{b\}), \quad then$$

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

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 (w^*, z^*) 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

$$U(A) \leq U(\{\alpha y + (1-\alpha)a, \alpha y + (1-\alpha)b\}) < U(A). \qquad Q.E.D.$$

We define the function $u: \Delta \to \mathbb{R}$ as

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

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

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

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* (i): 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 (i) follows from the definition of $v(z; a, b, \delta)$.

Part (ii): 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 (iii): 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 (iv): Let $\delta' < \delta$. Observe that

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

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

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

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

Part (v): 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

$$U\left(\left\{\frac{x+a}{2}, \frac{(1-\delta)b + \delta z + y}{2}\right\}\right). \qquad\qquad \textit{Q.E.D.}$$

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 $\delta \in (0, 1)$ satisfy $U(\{a\}) > U(\{a, (1-\delta)b + \delta z\}) > U(\{(1-\delta)b + \delta z\})$ for all $z \in \Delta$. Then*

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

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

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

$$= U(\{a, y\}) > U(\{y\})$$

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

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 $U(A) \leq (1-\delta)U(\{a, b\}) + \delta U(\{y\}) < U(A)$ and hence we have a contradiction. Similarly, if $z = (1-\delta)b + \delta a$, setting $w = (1-\delta)b + \delta y$ establishes that $U(A) \leq (1-\delta)U(\{b\}) + \delta U(\{a, y\}) < U(A)$. 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 $U(\{a\}) > U(\{a, y\}) = U(\{y\})$. It is enough to 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 holds if and only if

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

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

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 nonempty 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, without loss of generality, 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

$$U(\{a\}) > U(\{a, (1-\delta')c + \delta'z\}) > U(\{(1-\delta')c + \delta'z\}) \quad \text{for all } z \in \Delta.$$

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

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

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

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

$$= \max_{a \in A} \{u(w) + v(w)\} + \min_{b \in A} \{-v(w)\}.$$

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

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

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 (∗).

Note that both U and $\max_{x \in A}(u(x) + v(x)) - \max_{y \in A} v(y)$ are continuous functions on $\mathcal{A}$. Hence, if they agree on any set, they must agree on the closure of that set. We have shown that these two functions agree on the set of all finite subsets of Δ. Then, by Lemma 0, they must agree on $\mathcal{A}$.   *Q.E.D.*

B. PROOF OF THEOREM 3

By Lemma 1, a linear representation U of the preferences exists. Moreover, the restriction of U to singleton sets is continuous and therefore Lemmas 3, 4, and 5 also hold. We first show that the theorem holds for all finite subsets of Δ.

CLAIM 1: *If there exists some pair $x, y \in \Delta$ such that $U(\{x\}) > U(\{x, y\}) > U(\{y\})$, then there is a $\delta > 0$ such that $U(\{x\}) > U(\{x, (1-\delta)y + \delta z\}) > U(\{(1-\delta)y + \delta z\})$ for all $z \in \Delta$.*

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 exist sequences $\delta_k \in (0, 1)$ and $z_k \in \Delta$ such that $\delta_k \to 0$ and $U(\{x, y_k\}) = U(\{y_k\})$ for all k, where $y_k := (1-\delta_k)y + \delta_k z_k$. 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 $\limsup U_k \leq (U(\{y\}) + U(\{x, y\}))/2$. On the other hand, *Set Betweenness* implies that

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

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

Whenever $U(\{x\}) > U(\{x, y\}) > U(\{y\})$ for some $x, y \in \Delta$, Claim 1 enables us to 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 subsets 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 a cycle such that $\{w, x\} \sim \{w\}, \{x, y\} \sim \{x\}, \{w, y\} \sim \{y\}$. Hence, $U(\frac{1}{2}\{w, x\} + \frac{1}{2}\{x, y\}) = (U(\{w\}) + U(\{x\}))/2$. Note that

$$\frac{1}{2}\{w, x\} + \frac{1}{2}\{x, y\} = \left\{\frac{1}{2}w + \frac{1}{2}x, \frac{1}{2}y + \frac{1}{2}x\right\} \cup \{x\} \cup \left\{\frac{1}{2}w + \frac{1}{2}y\right\}.$$

But $U(\{\frac{1}{2}w + \frac{1}{2}x, \frac{1}{2}y + \frac{1}{2}x\}) = (U(\{y\}) + U(\{x\}))/2$ and $U(\{\frac{1}{2}w + \frac{1}{2}y\}) = \frac{1}{2}U(\{w\}) + \frac{1}{2}U(\{y\})$. Hence, by *Set Betweenness*,

$$U\left(\frac{1}{2}\{w, x\} + \frac{1}{2}\{x, y\}\right) \leq \max\left\{\frac{U(\{y\}) + U(\{x\})}{2}, U(\{x\}), \frac{U(\{w\}) + U(\{y\})}{2}\right\}$$

$$< \frac{U(\{w\}) + U(\{x\})}{2},$$

a contradiction. *Q.E.D.*

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 $A, B \in \mathscr{A}$.*

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 nonempty.

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 $(x^*, y^*) \in S$, $U(\{y^*\}) = U(\{w, y^*\}) < U(\{x^*\}) < U(\{w\})$. 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. *Q.E.D.*

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

DEFINITION: xRy if $\{x\} \nsim \{y\}$ 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 $\{x\} \nsim \{y\}$, then $\{x, y\} \sim \{x\}$ implies that the temptation ranking of x is 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 $\{y\} \nsim \{y, z\}$, 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 $\{x\} \nsim \{y\}$, then xRy or yRx by Axiom 4 and no self-control. Suppose that $\{x\} \sim \{y\}$, $\{x, w\} \sim \{x\}$, and $\{y, w\} \nsim \{y\}$ 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 $\{z\} \nsim \{y\}$. We know that $\{w\} \nsim \{y\}$. By Axiom 4 $\{w, x, y, z\} \sim \{y\}$ 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 $\{x\} \nsim \{y\} \nsim \{z\} \nsim \{x\}$, 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 $\{x\} \sim \{y\} \nsim \{z\}$, then since yRz we have $\{y, z\} \sim \{y\}$ and since xRy, we conclude $\{x, z\} \sim \{x\}$. Therefore, xRz.

If $\{x\} \nsim \{y\} \sim \{z\}$, then again $\{x\} \nsim \{z\}$ and hence it is sufficient to show that $\{x, z\} \sim \{x\}$. But yRz and $\{x, y\} \nsim \{y\}$ implies $\{x, z\} \sim \{z\}$ and hence $\{x, z\} \sim \{x\}$.

If $\{x\} \sim \{z\} \nsim \{y\}$, then $\{x, y\} \sim \{x\}$ and $\{y, z\} \nsim \{z\}$ and hence z not$R$ x and by the completeness xRz. *Q.E.D.*

LEMMA 7: *There exists a linear function $v : \Delta \to \mathbb{R}$ 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 there is a linear function $U: \mathscr{A} \to \mathscr{R}$ that represents $\succeq$ and the restriction of U to singleton sets is continuous. To prove *Independence*, assume that xPy and $\alpha \in (0, 1)$.

Step 1: Let $\{x\} \nsim \{y\}$. 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 $\{y\} \nsim \{y, w\}$. Clearly, $\{x\} \nsim \{w\} \nsim \{y\}$. 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*.

Claims 1–3 below complete the proof by demonstrating that R satisfies von Neumann-Morgenstern continuity. That is, there exist α, β such that $\alpha x + (1-\alpha)z \succ y \succ \beta x + (1-\beta)z$ whenever $xPyPz$. Assume that $xPyPz$.

CLAIM 1: *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 $\{y, z\} \sim \{y\}$, a contradiction. The proof of the second assertion follows from a similar argument.

CLAIM 2: *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 Claim 1 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.

CLAIM 3: *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 $\{w\} \sim \{z\}$. Applying Claim 1 or Claim 2 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: $yP[\beta x + (1-\beta z]$. Q.E.D.

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 $A \in \mathscr{A}$. 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 shown that for all finite subsets a self-control representation exists 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 Lemma 1.

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 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 nonempty, 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 nonempty, finite set such that $B \subset C \subset A$ implies $C \sim B$ for all finite C. By Lemma 0, there exists a sequence of finite subsets A_k of A that converges to A. Since $A_k \cup B \sim B$ for all k, by *Upper Semi-Continuity* $A = A \cup B \succeq B$. 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.   Q.E.D.

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 on 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.)   Q.E.D.

C. A LEMMA ON LINEAR INDEPENDENCE

The following Lemma is used in the proof of Theorems 4 and 8. The proofs of the two facts stated after the Lemma follow from the argument used in proving the lemma. Let $e := (1, 1, \ldots, 1)$.

LEMMA 9: *If w^1, w^2, w^3, e are linearly independent vectors in $\mathbb{R}^n$ and x is any vector in the relative interior of $S^{n-1} := \{z \in \mathbb{R}^n_+ : \sum z_i = 1\}$, then there exists y in the relative interior of S^{n-1} such that $w^1 \cdot (x - y) > 0 = w^2 \cdot (x - y) > w^3 \cdot (x - y)$.*

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 $\mathbb{R}^n$. 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 $e \cdot \zeta = 0$ and hence for $\epsilon > 0$ small, $y = x - \epsilon\zeta$ is in the relative interior of S^{n-1} and $w^1 \cdot (x - y) > 0 = w^2 \cdot (x - y) > w^3 \cdot (x - y)$ as desired.   Q.E.D.

FACT 1: *Let w^1, w^2, e be linearly independent vectors in $\mathbb{R}^n$ 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 linearly independent vectors in $\mathbb{R}^n$ 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 $w^1 \cdot (x - y) > 0 = w^2 \cdot (x - y)$.*

D. PROOF OF THEOREM 4

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 $\alpha_u, \alpha_v > 0$, 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 exist 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 Appendix Section C, there exist $x, y \in \Delta$ 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 $\{x\} \nsim \{y\}$. 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 $\alpha_u, \alpha_v > 0$. 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. $\hspace{1cm}$ *Q.E.D.*

E. PROOF OF THEOREM 5

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

LEMMA 10: *If $\succeq^*$ satisfies T1–3, then $\succeq_T$ is a preference relation*.

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, $(\{x, y, z\}, x) \succeq^* (\{x, y\}, x) \sim^* (\{x\}, x)$ since $x \succeq_T y$. Hence transitivity follows. $\hspace{1cm}$ *Q.E.D.*

To prove Theorem 5, take any $A, B \in \mathscr{A}$ with $A \succeq_1^* B$. Let x^* satisfy $(A \cup B, x^*) \succeq^* (A \cup B, x)$ for all x in $A \cup B$. (Clearly, by Axiom 2^*a such an x^* exists.) If $x^* \in A$, then T1 implies $(A, x^*) \succeq^* (A \cup B, x)$ for all x. Similarly, if $x^* \in B$, then $(B, x^*) \succeq^* (A \cup B, x)$ for all x. In either case, we conclude $A \succeq_1^* A \cup B$.

To prove $A \cup B \succeq_1^* B$ assume there exists no $y^* \in A$ such that $y^* \succ_T y$ for all $y \in B$ and let $x^* \in B$ satisfy $(B, x^*) \succeq^* (B, x)$ for all x in $A \cup B$. By Lemma 0, there exists a sequence A_k of nested subsets of A such that each A_k has k elements and the sequence A_k converges to A. By T3 $(A_k \cup B, x^*) \succeq^* (B, x^*)$ for all k and hence by Axiom 2^*a, $(A \cup B, x^*) \succeq^* (B, x^*)$ and therefore $A \cup B \succeq_1^* B$. If there exists some $y^* \in A$ such that $y^* \succ_T y$ for all $y \in B$, then it follows from Lemma 10 that there exists no $\hat{y} \in B$ such that $\hat{y} \succ_T y$ for all $y \in A$. Then reversing the roles of A and B in the preceding argument yields $A \cup B \succeq_1^* A \succeq_1^* B$ and completes the proof. $\hspace{1cm}$ *Q.E.D.*

F. PROOF OF THEOREM 6

PROOF: By Lemma 10, $\succeq_T$ is a preference relation. Claims 1–3 establish that v represents $\succeq_T$.

CLAIM 1: *If $u(x) > u(y)$, then $v(x) \geq v(y)$ iff $x \succeq_T y$.*

Since (u, v) represents $\succeq_1^*, v(x) \geq v(y)$ implies $x \sim_1^* \{x, y\} \succ_1^* \{y\}$. Therefore, $(\{x\}, x) \sim^* (\{x, y\}, x)$ and hence $x \succeq_T y$. Conversely, $x \succeq_T y$ implies $(\{x\}, x) \sim^* (\{x, y\}, x) \succ_1^* (\{y\}, y)$ and therefore $v(x) \geq v(y)$, proving Claim 1.

CLAIM 2: *If $u(x) < u(y)$, then $v(x) \geq v(y)$ iff $x \succeq_T y$.*

First observe that Claim 1 implies $v(y) \geq v(x)$ iff $y \succeq_T x$ and hence to prove Claim 2 it suffices to show that $v(y) = v(x)$ iff $y \sim_T x$. Suppose $v(x) = v(y)$ and x is not v maximal. Then (by continuity of u and v) there is a sequence $x_k \to x$ with $v(x_k) > v(x), u(x_k) > u(x)$. Then $x_k \succ_T y$ by Claim 1

and by T2 $(\{x_k, y\}, x_k) \succeq^* (\{x_k\}, x_k) \succeq^* (\{x\}, x)$ for all k and by Axiom 2*a $x \succeq_T y$. Since $y \succeq_T x$ by Claim 1, it follows that $x \sim_T y$. Suppose $v(x) = v(y)$ and x is v maximal. Then, there is a sequence $y_k \to y$ with $v(y_k) < v(y)$. By Claim 1, $x \succ_T y_k$ and hence $(\{x, y_k\}, x) \succeq^* (\{x\}, x)$ by T2. By 2*a, $(\{x, y\}, x) \succeq^* (\{x\}, x)$ and hence $x \sim_T y$.

For the converse, by Claim 1 $x \sim_T y$ implies $v(x) \geq v(y)$. To conclude the proof, we show that $v(y) > v(x)$ implies $y \succ_T x$. By regularity, there exists $z \in \Delta$ such that $(u(x), u(y), u(z)), (1, 1, 1)$ and $(v(x), v(y), v(z))$ are linearly independent. Let

$$\Delta(\{x, y, z\}) := \{w = a_1 \cdot x + a_2 \cdot y + a_3 \cdot z : (a_1, a_2, a_3) \in S^2\}.$$

We will use $a \in S^2$ and the corresponding element of $\Delta(\{x, y, z\})$ interchangeably. Let w be such that $u(y) > u(w) > u(x), v(y) > v(w) > v(x)$ and w is in the relative interior of S^2. Now Fact 1 of Appendix Section C implies that there is a w' such that $u(y) > u(w') > u(w)$ and $v(w) > v(w') > v(x)$. By Claim 1, $w' \succ_T w \succeq_T x$. Since (again by Claim 1) $y \succeq_T w'$ and since $\succeq_T$ is transitive, we conclude $y \succ_T x$.

CLAIM 3: *If $u(x) = u(y)$, then $v(x) \geq v(y)$ iff $x \succeq_T y$.*

Assume $x \succeq_T y$. If $v(y) > v(x)$, then by linearity we can choose z such that $u(z) \neq u(x)$ and $v(y) > v(z) > v(x)$. Applying Claims 1 and 2 above yield $y \succ_T z \succ_T x$, a contradiction.

For the converse, assume first x is not u maximal. Then, there is a z with $u(z) > u(x)$. If $v(z) \geq v(x)$, since $\succeq_1^*$ is represented by (u, v), it follows that $\{\lambda z + (1 - \lambda)x, y\} \succ_1^* \{y\}$. Since, by T1 $(\{y\}, y) \succeq^* (\{\lambda z + (1 - \lambda)x, y\}, y)$ we conclude that $(\{\lambda z + (1 - \lambda)x, y\}, \lambda z + (1 - \lambda)x) \succ^* (\{y\}, y) \sim^* (\{x\}, x)$ and hence by 2*a, $x \succeq_T y$. If $v(z) < v(x)$, define $(\{\lambda z + (1 - \lambda)y, x\}, x) \succeq (\{x\}, x)$ and hence $x \succeq_T y$ by Axiom 2*a.

If x is u maximal, then by minimal congruence we can find a z such that $u(x) > u(z), v(x) > v(z)$. Then $(\{\alpha z + (1 - \alpha)y, x\}, x) \succeq (\{x\}, x)$ and hence $x \succeq_T y$ by Axiom 2*a. *Q.E.D.*

G. PROOF OF THEOREM 7

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 $x \in c(A, u)$ and there is a $z \in \Delta$ with $u(z) > u(x)$. Then, since (u, v) represents $\succeq_1^*$, $A \cup \{(1 - \alpha)x + \alpha z\} \succ_1^* A$ for $\alpha \in (0, 1)$. By T1, $(A, y) \succeq^* (A \cup \{(1 - \alpha)x + \alpha z\}, y)$ for $y \in A$ and therefore $(A \cup \{(1 - \alpha)x + \alpha z\}, (1 - \alpha)x + \alpha z) \succ^* (A, y)$ for all $y \in A$. Axiom 2*a then implies $(A, x) \succeq^* (A, y)$ for all $y \in A$.

If x is u maximal, choose z with $u(z) < u(x)$. For $\alpha \in (0, 1)$, let $A_\alpha := (1 - \alpha)A + \alpha\{z\}$. Since (u, v) represents $\succeq_1^*$, $A_\alpha \cup \{x\} \sim_1^* A \succ_1^* A_\alpha$. By T1, $(A_\alpha, y) \succeq^* (A_\alpha \cup \{x\}, y)$ for $y \in A_\alpha$ and therefore $(A_\alpha \cup \{x\}, x) \succeq^* (A, y)$ for all $y \in A$. Axiom 2*a then implies $(A, x) \succeq^* (A, y)$ for all $y \in A$.

For the converse, suppose $x \in c(A, \succeq^*)$. This implies $(\{x\}, x) \sim^* (A, x) \succeq^* (A, y)$ for all $y \in A$. Hence $\{x\} \succeq_1^* A$. But, since $\succeq_1^*$ has no preference for commitment, this implies $x \in c(A, u)$. This completes the proof when (u, v) is not regular. For the remainder of the proof, assume (u, v) is regular. In that case, by Theorem 6, v represents $\succeq_T$.

CLAIM 1: *If $y^* \in c(A, v)$, then $(A, x) \sim^* (\{x, y^*\}, x)$ for all $x \in A$.*

Suppose $y^* \in c(A, v)$. By Lemma 0, there exists a sequence A_k of nested subsets of A such that each A_k has k elements and the sequence A_k converges to A. Apply T3 to each successive pair in $\{x, y^*\}, \{x, y^*\} \cup A_1, \ldots, \{x, y^*\} \cup A_k, \ldots$ to conclude that $(\{x, y^*\}, x) \sim^* (\{x, y^*\} \cup A_k, x)$ for all k, then use 2*a to establish the desired conclusion.

CLAIM 2: *If $\succeq$ 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 since (u, v) represents $\succeq_1^*$, we have $A \succ_1^* \{\hat{x}\}$ and

hence it follows from T1 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. So, assume $u(y^*) \leq u(\hat{x})$. It follows that $v(y^*) > v(\hat{x})$. If $y^* \notin c(A, u+v)$, then $\{x^*, y^*\} \succ_1^* \{\hat{x}, y^*\}$ and therefore for some $z \in \{x^*, y^*\}$, we have $(\{x^*, y^*\}, z) \succ^* (\{\hat{x}, y^*\}, z')$ for all $z' \in \{\hat{x}, y^*\}$. By Claim 1, $(\{\hat{x}, y^*\}, y^*) \sim^* (\{x^*, y^*\}, y^*)$. So, $(\{x^*, y^*\}, x^*) \succ^* (\{\hat{x}, y^*\}, \hat{x})$. Again by Claim 1, $(A, x^*) \sim^* (\{x^*, y^*\}, x^*)$ and $(\{\hat{x}, y^*\}, \hat{x}) \sim^* (A, \hat{x})$. Hence, $(A, x^*) \succ^* (A, \hat{x})$ and $\hat{x} \notin c^*(A, \succeq^*)$ as desired. If $y^* \in c(A, u+v)$, then since (u, v) represents $\succeq_1^*$ we have $A \sim_1^* \{y^*\}$. Since $u(y^*) + v(y^*) > u(x) + v(x)$ we have $v(y^*) > v(x)$ and by minimal congruence, we can find y such that $v(y) < v(y^*)$ and $u(y) < u(y^*)$. Since u, v are linear we can choose this y sufficiently close to y^* so that $u(y) + v(y) > u(\hat{x}) + v(\hat{x})$. Then, $A \succeq_1^* A \cup \{y\} \succeq_1^* \{\hat{x}, y, y^*\} \succ_1^* \{\hat{x}, y\}$ and since (u, v) represents $\succeq_1^*$, T1 implies $(\{\hat{x}, y, y^*\}, y^*) \succ^* (\{\hat{x}, y\}, \hat{x})$. Then, it follows from Claim 1 and T1 that for some $z \in A$, we have

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

hence $\hat{x} \notin c^*(A, \succeq^*)$ as desired.

CLAIM 3: *If $\succeq$ 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 hence $(\{\hat{x}\}, \hat{x}) \succeq^* (A, x)$, for all $x \in A$. If $\hat{x} \in c(A, v)$, then it follows from Claim 1 that $(\{\hat{x}\}, \hat{x}) \sim^* (A, \hat{x})$. Hence, $\hat{x} \in c^*(A, \succeq^*)$. So assume $\hat{x} \notin c(A, v)$ and pick some $y^* \in c(A, v)$. By minimal congruence there exists y such that $v(y) < v(y^*)$ and $u(y) < u(y^*)$. Let $A_\alpha := \{\hat{x}\} \cup \{\alpha\{y\} + (1-\alpha)A\}$. Note that $\{\hat{x}\} = c(A_\alpha, u+v)$ for all $\alpha < 1$ and $A_\alpha \succeq_1^* A$. Hence by Claim 2, for all $\alpha < 1$, $\hat{x} \in c^*(A_\alpha, \succeq^*)$. That is, $(A_\alpha, \hat{x}) \succeq^* (A, x)$ for all $x \in A$. Therefore, by Axiom 2*a, $(A, \hat{x}) \succeq^* (A, x)$ for all $x \in A$ and hence $\hat{x} \in c^*(A, \succeq^*)$.

CLAIM 4: *If $\succeq_1^*$ 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 $(A, x) \succ^* (\{\hat{x}\}, \hat{x})$ and by T1 $(\{\hat{x}\}, \hat{x}) \succeq^* (A, \hat{x})$. Hence, $(A, x) \succ^* (A, \hat{x})$ and therefore $\hat{x} \notin c^*(A, \succeq^*)$ as desired. If $u(\hat{x}) \geq u(y^*)$, then since $\hat{x} \notin c(c(A, v), u), v(y^*) > v(\hat{x})$. Therefore, by minimal congruence we can find y such that $u(y) < u(y^*)$ and $v(y) < v(y^*)$. Since u, v are linear we can choose this y sufficiently close to y^* so that $v(\hat{x}) < v(y) < v(y^*)$. Then $A \sim_1^* A \cup \{y\} \sim_1^* \{\hat{x}, y, y^*\} \succ_1^* \{\hat{x}, y\}$ and since (u, v) represents $\succeq_1^*$, T1 implies $(\{\hat{x}, y, y^*\}, y^*) \succ^* (\{\hat{x}, y\}, \hat{x})$. It follows from Claim 1 and T1 that for some $z \in A$ we have

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

hence $\hat{x} \notin c^*(A, \succeq^*)$ as desired.

CLAIM 5: *If $\succeq$ 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 $\succeq_1^*$, $\{\hat{x}\} \sim_1^* A$. By Claim 1, $(\{\hat{x}\}, \hat{x}) \sim^* (A, \hat{x})$. So, $(A, \hat{x}) \succeq^* (A, x)$ for all $x \in A$ and $x \in c^*(A, \succeq^*)$ as desired.   Q.E.D.

H. PROOF OF THEOREM 8

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 S^{n-1}, there exists y in the relative interior of S^{n-1} 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 Appendix Section C 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 C to obtain the desired y.   Q.E.D.

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 a nonnegative, full rank matrix Θ and a $\lambda \in \mathbb{R}^2$ such that*

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

for all $x \in \Delta$.

PROOF: We prove the only if part of the Lemma in two 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 $\mathscr{A}(\Delta(T^n))$ 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 $i = 1, 2$ to $\mathscr{A}(\Delta(T^n))$ is regular, there exists a nonnegative, full rank matrix Θ and a $\lambda \in \mathbb{R}^2$ such that

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

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 $a, b \in S^{n-1}$ such that $u^2 \cdot a > u^2 \cdot b, u^1 \cdot a = u^1 \cdot b, v^2 \cdot a < v^2 \cdot b$. But then we would have $\succeq_2$ display preference for commitment at $\{\sum a_i x_i, \sum b_i x_i\}$ and $\succeq_1$ have no preference for commitment at $\{\sum a_i x_i, \sum b_i x_i\}$, 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 $m_{k,l} \in \mathbb{R}$. 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 nonsingular. Let $\Theta = \mathbf{M}^{-1}$ and

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = -\mathbf{M}^{-1} \cdot \begin{pmatrix} m_{1,3} \\ m_{2,3} \end{pmatrix}.$$

Obviously, Θ is nonsingular.

It remains to be shown that Θ is nonnegative. If Θ is not nonnegative, then, by Lemma 11, there exist a, b in the relative interior of S^{n-1} such that $u^1 \cdot (a - b) < 0, v^1 \cdot (a - b) < 0$, and $w \cdot (a - b) > 0$ for $w = u^2$ or $w = v^2$. Let $x := \sum a_i x_i, y := \sum b_i x_i$. If $(u^2 \cdot (a - b) > 0$ and $v^2 \cdot (a - b) < 0)$ or $(u^2 \cdot (a - b) < 0$ and $v^2 \cdot (a - b) > 0)$ then $\succeq_2$ has a preference for commitment at $\{x, y\}$ while $\succeq_1$ does not, a contradiction. So, we assume either $(u^2 \cdot (a - b) > 0$ and $v^2 \cdot (a - b) \geq 0)$ or $(u^2 \cdot (a - b) \geq 0$ and $v^2 \cdot (a - b) > 0)$. In the former case, by Fact 1, there exists c in the relative interior of S^{n-1} such that $u^2 \cdot (a - c) > 0 > v^2 \cdot (a - c)$. Let $z := \sum c_i x_i$. Then, for $\alpha \in (0, 1)$ sufficiently small, $\{x, \alpha z + (1 - \alpha)x, x\} \sim_1 \{y\} \succeq_1 A$ 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 $\succeq_1$ has 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 $\mathscr{A}(\Delta(T_1^0))$ is regular. Similarly, there exists some T_2^0 such that $\succeq_2$ restricted to $\mathscr{A}(\Delta(T_2^0))$ 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 Θ^n, λ^n that satisfy the conditions of the Lemma, for every $x \in \Delta(T^n)$. But then, $\Theta^1 = \Theta^n$ and $\lambda^1 = \lambda^n$ for all n. Since, u_2, v_2, u_1, v_1 are all continuous and $\cup_n \mathscr{A}(\Delta(T^n))$ is dense in $\mathscr{A}$, 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 nonnegative matrix Θ and $\lambda \in \Re^2$ such that for all $x \in \Delta$

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

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 $y \in A$. Let $\theta_{i,j}$ denote the i,jth entry of Θ and

$$\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}.$$

Hence,

$$u_2(x) = \theta_{1,1} \cdot u_1(x) + \theta_{1,2} \cdot v_1(x) + \lambda_1 \geq \theta_{1,1} \cdot u_1(y) + \theta_{1,2} \cdot v_1(y) + \lambda_1 = u_2(y)$$

and

$$v_2(x) = \theta_{2,1} \cdot u_1(x) + \theta_{2,2} \cdot v_1(x) + \lambda_2 \geq \theta_{2,1} \cdot u_1(y) + \theta_{2,2} \cdot v_1(y) + \lambda_2 = v_2(y)$$

for all $y \in A$. Therefore, $\succeq_2$ has no preference for commitment at A.   Q.E.D.

To conclude the proof, suppose the desired α, β, and γ exist. Then, clearly Θ 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, Θ and λ with the desired properties exist. Let

$$\alpha := \frac{\theta_{1,1}}{\theta_{1,1} + \theta_{1,2}}, \qquad \beta := \frac{\theta_{2,1}}{\theta_{2,1} + \theta_{2,2}}, \qquad \gamma := \frac{\theta_{2,1} + \theta_{2,2}}{\theta_{1,1} + \theta_{1,2}},$$

$$u_2 = \alpha u_1 + (1 - \alpha)v_1, \quad \text{and} \quad v_2 = \beta u_1 + (1 - \beta)v_1.$$

By Theorem 4, since (u_2, v_2) represents $\succeq_2$, so does $(u_2, \gamma v_2)$.   Q.E.D.

REFERENCES

DEKEL, E., B. LIPMAN, AND A. RUSTICHINI (2001): "Representing Preferences with a Unique Subjective State Space," *Econometrica,* 69, 891–934.

KREPS, D. (1979): "A Preference for Flexibility," *Econometrica,* 47, 565–576.

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

MACHINA, M. J. (1989): "Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty," *Journal of Economic Literature,* 27, 1622–1668.

RABIN, M. (1998): "Psychology and Economics," *Journal of Economic Literature,* 36, 11–46.

STROTZ, R. H. (1955): "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies,* 23, 165–180.