American Economic Journal: Microeconomics 2009, 1:2, 148–167
http://www.aeaweb.org/articles.php?doi=10.1257/mic.1.2.148

A "Dual Self" Representation for Stochastic Temptation†

By Kalyan Chatterjee and R. Vijay Krishna*

We consider the following two-period problem of self-control. In the first period, an individual has to decide on the set of feasible choices from which she will select one in the second period. In the second period, the individual might choose an alternative that she would find inferior in the first period, an eventuality that need not occur with certainty. We propose a model for this problem and axioms for first-period preferences, in which the second-period choice could be interpreted as being made by an "alter ego" who appears randomly. We provide a discussion of the behavioral implications of our model as compared with existing theories. (*JEL* D11, D80)

Casual observation and introspection (especially about unhealthy but tasty food items), as well as stories like that of Ulysses and the Sirens, suggest that otherwise rational individuals act to constrain their future choices. At first sight, this seems inconsistent with standard notions of rational choice, where increasing the set of available choices cannot make anyone strictly worse off. Robert H. Strotz (1955, 168), who discusses intertemporal choice, calls this constraining of one's future choices the "strategy of precommitment" and suggests that it results from the agent today being a "different person with a different discount function from the agent of the past." Thomas C. Schelling (1978, 1984) gives several examples of this kind of behavior and also speaks of "multiple selves," though he also writes that he is somewhat hesitant to use this terminology outside the circle of professional economists. Peter J. Hammond (1976), in a paper on coherent dynamic choice, discusses an example of addiction as a reflection of changing tastes.

It is important here to distinguish between changing tastes and choices being made that maximize a different utility function. An agent whose tastes might change would never (strictly) prefer a smaller set of alternatives to a larger set of alternatives. An example of this is a person who knows in the morning that she might be in the mood for either seafood or vegetarian food for dinner (depending on her changed taste in the evening), even though she would pick seafood if she were forced to choose a menu item in the morning. It is clear that she would prefer making a

* Chatterjee: Pennsylvania State University, 609 Kern Building, University Park, PA 16802 (e-mail: kchaterjee@psu.edu); Krishna: University of North Carolina, 107 Gardner Hall, CB 3305, Chapel Hill, NC 27599 (e-mail: rvk@e-mail.unc.edu). An earlier version of this paper was circulated under the title "Menu Choice, Environmental Cues and Temptation: A 'Dual Self' Approach to Self-control." We wish to thank Sophie Bade, Siddhartha Bandyopadhyay, Susanna Esteban, Kfir Eliaz, Drew Fudenberg, Ed Green, Faruk Gul, John Hardman Moore, Jawwad Noor, József Sákovics, Rani Spiegler, Jonathan Thomas, an anonomous referee, and especially Bart Lipman for helpful comments, and Fernando Chague and Avantika Chowdhury for research assistance.

† To comment on this article in the online discussion forum, or to view additional materials, visit the articles page at: http://www.aeaweb.org/articles.php?doi=10.1257/mic.1.2.148.

reservation at a restaurant that serves both rather than one that serves only seafood. If instead, the agent perceives that future choices might be made that maximize a utility function different from hers, she might prefer to limit her options. For instance, if someone had to choose between going to a party where tea and coffee are being served, versus another one where tea, coffee, and cocaine[1] would be served, the fact that this person might choose cocaine in the evening would possibly lead her to precommit in the morning to the party where cocaine would not be available. We are interested in the second notion mentioned above, namely, that of choices in the future possibly being made to maximize a different utility function.

One way to conceptualize both types of problems is to consider, after David M. Kreps (1979), an implicit two-stage choice problem. In the morning, an individual chooses a menu of objects. In the evening she chooses an item from the previously chosen menu. As mentioned above, we are interested in the following behavior, which the individual considers a possibility. The menu chosen triggers temptation with some probability in the next period. Temptation is thought of as the (utility maximizing) choice of a virtual alternate self or alter ego.[2] Whichever self is in charge of making a choice from the menu in the second period makes its own most preferred choice. Each self does not particularly care about the utility of the other self, so this is not an interdependent utilities model, but does care about the choice made (although we shall assume that the alter ego breaks ties in favor of the decision maker).

We therefore interpret the presence of systematic "mistakes" as being due to a virtual alter ego who (systematically) chooses items from a menu that are not preferred by the "long-run" self. This suggests a model with three ingredients: a long-run self's utility function $u(\cdot)$, a virtual self's utility function $v(\cdot)$, and a probability of getting tempted ρ. Since the objects of choice in the (unmodelled) second period are objective lotteries, $u(\cdot)$ and $v(\cdot)$ must be von Neumann-Morgenstern utility functions. Note that $v(\cdot)$ is supposed to model choices that tempt the long-run self, so it must also represent temptation in a way to be specified by the axioms.

Let x be a typical menu and β a typical member of the menu. The decision-maker's utility from a singleton is given by u; the alter ego's utility function is v; and $B_v(x)$ is the set of v-maximizers in x. The decision maker believes she will be tempted with probability ρ, i.e., with probability ρ, the alter ego will make a choice. This induces a preference, indeed a utility function, over menus, given by

$$U(x) = (1 - \rho) \max_{\beta \in x} u(\beta) + \rho \max_{\beta \in B_v(x)} u(\beta). \tag{1}$$

The aforementioned asymmetry between the selves requires that the alter ego, if he has to make a choice, will choose, among his most preferred alternatives, that which

[1] For those who are made uncomfortable with the examples from Hammond, Schelling, and Amartya Sen, and the later discussion of broccoli and chocolate, Manuel Amador, Iván Werning (2006) and George-Marios Angeletos discuss the macroeconomic implications of problems that feature both a preference for flexibility and self-control.

[2] We can think of this as a generalization of Strotz (1955), in that the alternate self appears here with some probability instead of with certainty.

is most preferred by the decision maker. Since we are characterizing the decision maker's utility, how she breaks ties does not really matter to us.

We shall say that a utility function over menus that takes such a form admits a *dual self* representation. We provide axioms for first-period preferences over menus so that the decision maker's utility from a menu is given by equation (1) above. Thus, a decision maker who satisfies our axioms behaves as if there is a probability of her being tempted when a choice has to be made from a menu, which is represented as the choice being made by an alter ego. It should be emphasized that the alter ego (and his utility function v) is subjective, as is the probability, ρ, of getting tempted. The only observables are first-period choices over menus, and v and ρ must be inferred from these choices. The inference works as follows.

Consider two prizes: $\{a\}$ and $\{b\}$. If the decision maker's preferences are given by $\{b\} \succ \{a,b\} \succcurlyeq \{a\}$, then it must be the case that (i) (obviously) $u(b) > u(a)$, (ii) $v(a) > v(b)$, i.e., the alter ego prefers b to a, and (iii) $\rho > 0$, i.e., the decision maker subjectively assesses that there is a positive probability that the alter ego will make the choice from the menu $\{a,b\}$. We may say that a is a revealed temptation for b. If requirements (ii) and (iii) do not hold, then temptation would not be an issue.

In our Dual Self Theorem, we provide a characterization of utility functions (over menus) that satisfies equation (1). Such a utility function has the property that it depends solely on the best and the maximally tempting elements on a menu (i.e., the u- and v-maximizers, respectively) in a very simple way, namely with a constant ρ. However, it is natural to think that ρ could vary with the menu (for example, the presence of sorbet on a menu makes it easier to be tempted by rich ice cream). We do not consider such preferences in this paper, but for a model that captures such behavior see Chatterjee and Krishna (2005).

We emphasize that this paper seeks to characterize behavior that exhibits temporary loss of self-control by determining a set of axioms on a choice of menus that yields this representation. Note that since we are interested in choice under temptation, the various components of the representation above cannot be unrestricted. If we restrict ourselves to menus consisting of a single alternative each (an alternative is an "objective lottery" like the ones in the classic von Neumann-Morgenstern formulation of expected utility), there is no scope for temptation and the utility of a singleton "menu" $\{\alpha\}$ is just the von Neumann-Morgenstern utility $u(\alpha)$.[3]

The remainder of the paper is structured as follows. In Section I, we introduce our model. In Section II, we introduce the axioms, our representation theorem, and a sketch of its proof. In Section III, we compare our model with that of Faruk Gul and Wolfgang Pesendorfer (2001) (henceforth GP). We also describe a dual self model where the probability of temptation depends on the menu and compare it to the "random Strotz representation" of Eddie Dekel and Barton L. Lipman (2007) (henceforth DL), and then review other related literature. Section IV concludes and proofs are in the Appendix.

[3] This implies that the axioms in our paper should (and do) reduce to those of von Neumann and Morgenstern for singleton menus. These latter axioms are not uncontroversial. Nevertheless, since our focus is on issues relating to self-control, we think they are acceptable assumptions in our framework.

I. The Model

We have in mind a decision maker who faces a two-period decision problem. In the first period, the agent chooses the set of alternatives from which a consumption choice will be made in the second period. Nevertheless, as in Kreps (1979), Dekel, Lipman, and Aldo Rustichini (2001) (henceforth DLR), and GP, we shall only look at first-period choices. Let us now describe the ingredients more formally. (The basic objects of analysis are exactly the same as in GP.)

The set of all prizes is Z, where (Z, d) is a compact metric space. The space of probability measures on Z is denoted by Δ (with generic elements being denoted by $\alpha, \beta, \dots$) and is endowed with the topology of weak convergence (which is metrisable). The objects of analysis are subsets of Δ. Let $\mathcal{A}$ be the set of all closed, *essentially finite* subsets of Δ (with generic elements, called *menus*, denoted by $x, y, \dots$) endowed with the Hausdorff metric. A set $x \subset \Delta$ is essentially finite if it is finite, or if it is the convex hull of finitely many points.

Convex combinations of elements x, $y \in \mathcal{A}$ are defined as follows. We let $\lambda x + (1-\lambda)y := \{\gamma = \lambda\alpha + (1-\lambda)\beta : \alpha \in x, \beta \in y\}$, where $\lambda \in [0,1]$. (This is the so-called Minkowski sum of sets.) We are interested in binary relations $\succcurlyeq$ which are subsets of $\mathcal{A} \times \mathcal{A}$.

Before we impose axioms on $\succcurlyeq$, it may be worthwhile to dwell on the implications of the model. The use of subsets of lotteries over Z as the domain for preferences instead of subsets of Z itself was first initiated by DLR in this context and is reminiscent of the approach pioneered by Frank J. Anscombe and Robert J. Aumann (1963). From a normative point of view, this approach should not be troublesome as long as our decision makers are able to conceive of the lotteries they consume and agree with the axioms we impose on them. But from a revealed preference perspective, are decision makers faced with menus of lotteries? As noted by Kreps (1988, 101) (in the context of the Anscombe-Aumann theory), if decision makers are not faced with choices of lotteries, our assumption, that they are, can be quite burdensome, especially from a descriptive point of view.

Nevertheless, it could be argued that such menus of lotteries are, in fact, objects of choice. A patient who chooses to go to a hospital is, arguably, choosing a menu of lotteries with the level of pain being an uncontrolled random event. Similarly, someone who fancies seafood and goes to a restaurant not knowing the quality of the shrimp, is doing the same. It is also possible that the menu of lotteries could arise from a nondegenerate, mixed strategy played by an opponent, for instance, in determining the set of objects available for sale by, say, a car dealer. There is, of course, the analytical benefit of our approach, which is the use of the additional structure a linear space provides.

II. Axioms and Representations

We first define the linear functionals relevant to a dual self representation. As is standard, we shall say that $U: \mathcal{A} \to \mathbb{R}$ is *linear* if $U(\lambda x + (1-\lambda)y) = \lambda U(x) + (1-\lambda) U(y)$ for all $x, y \in \mathcal{A}$ and $\lambda \in (0,1)$, and that it *represents* $\succcurlyeq$ if it is the case that $U(x) \geq U(y)$ if and only if $x \succcurlyeq y$. The functions u, $v: \Delta \to \mathbb{R}$ are linear if similar conditions hold. Let $B_v(x) = \arg\max_{\beta \in x} v(\beta)$ be the set of v-maximizers in x (with a similar definition for B_u). We are interested in preferences that admit the following representation.

A. *Definition*

A *dual self representation* of a preference relation $\succcurlyeq$ is a triple (u, v, ρ), where u, $v : \Delta \to \mathbb{R}$ are continuous and linear, and $\rho \in [0, 1]$ such that $U : \mathcal{A} \to \mathbb{R}$ given by

$$U(x) := (1 - \rho) \max_{\beta \in x} u(\beta) + \rho \max_{\beta \in B_v(x)} u(\beta)$$

represents $\succcurlyeq$.

Here, u represents the agent's preferences over singletons, her *normative preference*, and the function v represents the alter ego's preferences over lotteries. The probability that the alter ego makes a choice is ρ, so that this is the probability that an agent is *tempted* when a choice must be made from the menu. We now turn to the behavioral postulates that ensure the existence of such a representation.

AXIOM 1: (Preferences) $\succcurlyeq$ is a complete and transitive binary relation.

AXIOM 2A: (Upper Semicontinuity) The sets $\{y \in \mathcal{A} : y \succcurlyeq x\}$ are closed.

AXIOM 2B: (Lower von Neumann–Morgenstern Continuity) $x \succ y \succ z$ implies $\lambda x + (1 - \lambda) z \succ y$ for some $\lambda \in (0, 1)$.

AXIOM 2C: (Lower Singleton Continuity) The sets $\{\alpha : \{\beta\} \succcurlyeq \{\alpha\}\}$ are closed.

AXIOM 3: (Independence) $x \succ y$ and $\lambda \in (0, 1]$ implies $\lambda x + (1 - \lambda) z \succ \lambda y + (1 - \lambda) z$.

The first axiom is standard. Axioms 2A–C are identical to Axioms 2A–C in Section 3 of GP. They ensure that we have enough continuity to enable us to have a linear utility representation while allowing for discontinuous preferences. The motivation for *independence* is the familiar one, and some normative arguments in its favor are given in DLR and GP. It basically says that our decision maker does not distinguish between simple and compound lotteries and all that matters to her are the prizes. (Nevertheless, as noted by Drew Fudenberg and David K. Levine 2006, this may not be an innocuous assumption.) Clearly, Axioms 1–3 deliver a continuous linear functional $U : \mathcal{A} \to \mathbb{R}$ that represents $\succcurlyeq$. Our next axiom captures the essence of temptation.

AXIOM 4: (Temptation) There exist α, $\beta \in x$ such that $\{\alpha\} \succcurlyeq x \succcurlyeq \{\beta\}$.

Temptation says that insofar as the presence of alternatives different from the best alternative on the menu affects the decision maker, it does not make the decision maker worse off than her worst choice on the menu. It also says that the cost of temptation (i.e., the cost of not being able to choose the best alternative) is bounded. In other words, it is never the case that "analysis is paralysis."

Consider two prizes: broccoli (b) and rich chocolate cake (c). Suppose the decision maker has preferences over menus as follows: $\{b\} \succ \{b, c\} \succcurlyeq \{c\}$. We can then conclude that the presence of c on the menu, which makes the decision maker worse off, is the source of temptation. Formally, let $\{\beta\}$ be any lottery. Say that $\{\alpha\}$ *tempts* $\{\beta\}$ if $\{\beta\}$ is superior to $\{\alpha\}$, and the addition of $\{\alpha\}$ to $\{\beta\}$ makes the agent strictly worse off, i.e., $\{\beta\} \succ \{\alpha, \beta\} \succcurlyeq \{\alpha\}$. More generally, say that *y tempts β* if $\{\beta\} \succ \{\beta\} \cup y$, and $\{\beta\} \succcurlyeq \{\alpha\}$ for all $\alpha \in y$. Thus, at the very least, y contains some elements that tempt β.

The next axiom says that a decision maker finds lotteries over tempting alternatives to be tempting. To see why this might be the case, think of a decision maker for whom $\{m\}$ and $\{c\}$ tempt $\{b\}$. Let us also suppose that before being offered a ½ – ½ lottery over the two tempting alternatives, she is allowed a whiff of the two temptations. To say that the decision maker also finds this lottery tempting says that she does not use fortune as a cover for her frailty. We now formally state the axiom which says that the set of lotteries that tempts any $\{\beta\}$ is well behaved.[4]

AXIOM 5: (Regularity) $\{\alpha_1\}$ and $\{\alpha_2\}$ tempt $\{\beta\}$ implies, $\{\lambda\alpha_1 + (1 - \lambda)\alpha_2\}$ tempts $\{\beta\}$ for all $\lambda \in [0, 1]$.

The first part of our next axiom is an excision axiom in that it allows us to *excise* elements from a menu without affecting the value of the menu to the decision maker. Consider three prizes: broccoli (b), rich chocolate cake (c), and deep-fried Mars bars (m). Our decision maker has the following preferences over the prizes in the morning: $\{b\} \succ \{c\} \succ \{m\}$, and both m and c tempt b, i.e., $\{b\} \succ \{b, c\}, \{b, m\}$. Thus, the presence of c and m make the decision maker strictly worse off. Now, suppose that adding m to the menu $\{c\}$ does not affect the decision maker, i.e., $\{c\} \sim \{c, m\}$. This implies that the "real" temptation comes from the rich chocolate cake and the addition of m to the menu $\{b, c\}$ should leave the agent indifferent, (i.e, $\{b, c\} \sim \{b, c, m\}$). Toward this end, say that $\beta \in x$ is *untempted in x* if there is no $\alpha \in x$ such that $\{\beta\} \succ \{\beta, \alpha\} \succcurlyeq \{\alpha\}$.

AXIOM 6: (Additivity of Menus: AoM) For x, y finite, $\beta \in \{\beta\} \cup x \cup y$ untempted in $\{\beta\} \cup x \cup y$ and y such that $\{\beta\} \succcurlyeq \{\alpha\}$ for all $\alpha \in y$, $\{\beta\} \sim \{\beta\} \cup y$ implies $\{\beta\} \cup x \cup y \sim \{\beta\} \cup x$.

To recap, the axiom says that if y is dominated, elementwise, by β, and there is no $\alpha \in y$ that tempts β, then removing y from $\{\beta\} \cup x \cup y$ does not affect the value of $\{\beta\} \cup x$.[5] This has the flavor of additivity, namely adding the same set to both sides of an expression does not change the relation between the left- and right-hand sides.

We now proceed toward the axiom that isolates the fact that ρ is a constant across all menus. We have already seen one kind of excision in *AoM*. Another kind of excision is the notion that the only items on a menu that matter to a decision maker is the

[4] This rules out the possibility that there are multiple alter egos or selves who might each possess a different utility function and arise with some probability. More on this in the conclusion.

[5] This is similar to the axiom used by Kreps (1979) who requires that $x \sim x \cup x'$ implies $x \cup x'' \sim x \cup x' \cup x''$.

alternative she would have chosen were she not tempted and the item in the menu that causes her maximal temptation. For instance, suppose the decision maker's preferences are as follows: $\{b\} \succ \{b, c\} \succ \{c\} \succ \{c, m\} \succ \{m\}$. Thus, although c tempts b, c itself is tempted by m. Then, whenever both are present, we will require that the decision maker is unaffected by the presence of c. In other words, $\{b, c, m\} \sim \{b, m\}$. We shall formalize this below. Let us say that $\beta \in x$ is *tempted* if there exists $y \subset x$ that tempts β.

AXIOM 7: (Separability of Menus: SoM) If x is finite and $\beta \notin B(x \cup \{\beta\})$ is tempted,

$$x \cup \{\beta\} \sim x.$$

(Here $B(\{\beta\} \cup x) := \{\alpha \in \{\beta\} \cup x : \{\alpha\} \succcurlyeq \{\gamma\} \text{ for all } \gamma \in \{\beta\} \cup x\}$ is the set of best singletons in x.) *SoM* says that the only alternative that matters in a menu (other than the decision maker's best alternative on the menu) is the object that is maximally tempting. We want to express the idea that *if* the agent succumbs to temptation, she will fall all the way and choose the most tempting alternative (from the perspective of the alter ego).

B. *Theorem*

DUAL SELF THEOREM: *A binary relation* $\succcurlyeq$ *satisfies Axioms* 1, 2A–C*, and* 3–7 *if and only if it admits an essentially unique dual self representation, so that the induced U is of the form*

$$U(x) = (1 - \rho) \max_{\beta \in x} u(\beta) + \rho \max_{\beta \in B_v(x)} u(\beta).$$

PROOF:

See Appendix.

The dual self representation is essentially unique in the following sense, u and v are unique up to (not necessarily the same) positive affine transformation, and ρ is unique. That U is unique up to positive affine transformation follows immediately from the representation. We now turn to a brief sketch of the proof of dual self theorem.

The dual self theorem says that when faced with choices of menus, the decision maker who satisfies Axioms 1–7 behaves *as if* she has an alter ego who has a utility function over lotteries given by v. Moreover, in the event that the alter ego makes the choice, he chooses the lottery in his most preferred set (in x), which maximizes the decision maker's utility. Also, the decision maker behaves *as if* she will be tempted (i.e., the probability that the choice will be made by the alter ego) with a probability of ρ when faced with the menu x.

One way to imagine the representation is that the decision maker expects an internal battle for self-control with her alter ego, and ρ represents the probability that she loses this battle. The function v has an immediate behavioral interpretation. Consider a preference $\succcurlyeq$ that admits a dual self representation. Then, for lotteries $\alpha, \beta \in \Delta$, $\{\beta\} \succ \{\beta, \alpha\} \succcurlyeq \{\alpha\}$ if and only if (i) $\rho > 0$, (ii) $u(\beta) > u(\alpha)$, and (iii) $v(\alpha) > v(\beta)$. Thus, if the alter ego's most preferred choice is different from the decision maker's best choice, then there is a battle for self-control, the decision maker anticipates losing this battle with some probability and consequently values the menu with the tempting option less.

It is entirely conceivable that ρ is not menu independent. For instance, consider a decision maker whose choices for dessert include fruit (f), sorbet (s), and chocolate cake (c). Suppose the decision maker's preferences are $u(f) > u(s) > u(c)$, and suppose the alter ego is such that $v(c) > v(s) > v(f)$. Then, the decision maker can have $U(\{f, c\}) > U(\{f, s, c\})$, i.e., the presence of more temptations increases the probability that the decision maker will lose the battle for self-control. Such a model is studied in Chatterjee and Krishna (2005).

We end with a simple proposition that allows us to compare the strength of the alter ego across different individuals. Let us suppose that $\succcurlyeq_1, \succcurlyeq_2$ are preferences that satisfy Axioms 1–7 and so admit dual self representations. Moreover, suppose the representations are (u, v, ρ_i) for $i = 1, 2$. Thus, α tempts β for $\succcurlyeq_1$ if and only if α tempts β for $\succcurlyeq_2$.

Let us say that $\succcurlyeq_1$ has *greater self-control* than $\succcurlyeq_2$ if for all $\alpha, \beta, \gamma \in \Delta$ such that α tempts β (for preferences $\succcurlyeq_1$ and $\succcurlyeq_2$), $\{\gamma\} \succcurlyeq_1 \{\beta, \alpha\}$ implies $\{\gamma\} \succcurlyeq_2 \{\beta, \alpha\}$. Intuitively, while both preferences have the same set of temptations, the cost of temptation is lower for $\succcurlyeq_1$. The following proposition formalizes this intuition.

PROPOSITION: *Let* $\succcurlyeq_1$ *and* $\succcurlyeq_2$ *have dual self representations of the form* (u, v, ρ_i), $i = 1, 2$. *Then,* $\succcurlyeq_1$ *has greater self-control than* $\succcurlyeq_2$ *if and only if* $\rho_1 \leq \rho_2$.

PROOF:

Necessity.—Let α, β such that $\{\beta\} \succ_i \{\beta, \alpha\}$. Then, $(1 - \rho_i)\{\beta\} + \rho_i(\alpha) \sim_i \{\beta, \alpha\}$ for $i = 1, 2$. But this implies $(1 - \rho_1)\{\beta\} + \rho_1(\alpha) \succcurlyeq_2 \{\beta, \alpha\} \sim_2 (1 - \rho_2)\{\beta\} + \rho_2\{\alpha\}$, i.e., $\rho_1 \leq \rho_2$.

Sufficiency.—Suppose $\rho_1 \leq \rho_2$ and suppose $\{\gamma\} \succcurlyeq_1 \{\beta, \alpha\} \sim_1 \{(1 - \rho_1)\beta + \rho_1\alpha\}$. But since $\rho_1 \leq \rho_2$, $\{\gamma\} \succcurlyeq_2 \{(1 - \rho_1)\beta + \rho_1\alpha\} \succcurlyeq_2 \{(1 - \rho_2)\beta + \rho_2\alpha\} \sim_2 \{\beta, \alpha\}$ as desired.

C. *Proof Sketch of Dual Self Theorem*

The "if" part of the proof and the uniqueness of the representation in the "only if" part have been described above. Here, we sketch the "only if" part. The proof proceeds through a series of simple arguments that we describe next.

Representing $\succcurlyeq$.—An application of the mixture space theorem (Lemma V2) shows that Axioms 1–3 guarantee the existence of a continuous linear functional U, unique up to affine transformation, which represents $\succcurlyeq$. Also U restricted to singletons is continuous.

The Alter Ego's Preferences.—For lotteries α, β such that $\{\beta\} \succ \{\beta, \alpha\} \succcurlyeq \{\alpha\}$, we stipulate that this must be because the alter ego strictly prefers α to β. From *Regularity*, we see that for each $\beta \in \Delta$, the set $\beta_+ := \{\alpha : \{\beta\} \succ \{\beta, \alpha\} \succcurlyeq \{\alpha\}\}$ is convex. Repeated application of *AoM* tells us that $\beta_- :=$ cl $(\{\alpha : \{\beta\} \sim \{\beta, \alpha\} \succ \{\alpha\}\})$ is also convex (where cl A is the closure of the set A). Thus, β_+ and β_- are disjoint, convex sets. Suppose that Z is finite, so that Δ is finite dimensional. Then, there exists a linear functional v that separates β_+ and β_-. Now, for the infinite dimensional Δ, we show that the separation argument described can be carried out on certain finite dimensional convex subsets, and there exists a linear functional that performs the separation for all these finite dimensional subsets and hence for Δ.

Translation Invariance.—We say that U is *translation invariant* if $U(x) \geq U(y)$ if and only if $U(x + c) \geq U(y + c)$ for all signed measures c such that $c(Z) = 0$ and $x + c, y + c \in \mathcal{A}$. That U is translation invariant follows from Axioms 2A–C, which ensure that U is sufficiently continuous and *Independent*. We use this property to show that there is an essentially unique linear functional that performs the separation described in the previous step for each lottery β. Thus, there exists a continuous linear functional that represents the alter ego's preference.

Finite Menus.—For any finite menu x, let $\beta_x^* \in x$ be such that $u(\beta_x^*) = \max_{\beta \in x} u(\beta)$ and let $\hat{\beta}_x \in B_v(x)$ be such that $u(\hat{\beta}_x) = \max_{B_v(x)} u(\beta)$. *Temptation* and repeated application of *AoM* implies that $u(\beta_x^*) \geq U(x) \geq u(\hat{\beta}_x)$.

Excising More Items from the Menu.—We now apply *SoM* to show that for each x, $x \sim \{\beta_x^*, \hat{\beta}_x\}$. Thus, $U(x) = (1 - \rho_x)u(\beta_x^*) + \rho_x u(\hat{\beta}_x)$. Finally, we show that ρ_x is independent of x and hence constant, giving us the desired representation.

III. Related Literature

GP (2001) pioneered the axiomatic treatment of temptation and self-control. They view self-control as the agent making a choice from a set of alternatives, and the choice reflects the agent's compromise between choosing what he would ideally have chosen (according to his *normative preference*) and the psychological cost (of self-control) that the agent faces when he doesn't choose the most tempting alternative. To this end, they offer two representations, the first of which is a *self-control representation*, which consists of a pair of expected utility functionals (u, v) (where $u, v : \Delta \to \mathbb{R}$) such that the utility of a menu is given by

$$\begin{aligned} V_{SC}(x) &= \max_{\beta \in x} (u(\beta) + v(\beta)) - \max_{\beta \in x} v(\beta) \\ &= \max_{\beta \in x} (u(\beta) - c(\beta; x)), \end{aligned}$$

where $c(\beta; x) := (\max_{\alpha \in x} v(\alpha)) - v(\beta)$. The second representation suggests that the agent expects to choose the item β from the menu x that maximizes $u(\beta) - c(\beta; x)$, where $c(\beta; x)$ is the cost of resisting temptation. Chatterjee and Krishna (2005) show that such a self-control representation admits a *menu dependent* dual self representation (u, v, ρ_x), where the probability of temptation, ρ_x, depends on the menu. They also provide an axiomatization of such preferences. Unfortunately, their requirement that the representation that U be linear imposes some conditions on ρ_x that are hard to interpret behaviorally, i.e., the conditions are hard to understand independently of the linearity of the representation.[6] Dekel, Lipman, and Rustichini (2007) consider more general cost functions that allow for a richer set of temptations.

GP (2001) have another representation they term the *overwhelming temptation representation*, where the utility of a menu is given by

$$V_{OT}(x) = \max_{\beta \in B_v(x)} u(\beta),$$

where $B_v(x) = \arg\max_{\gamma \in x} v(\gamma)$. Temptation is overwhelming because the choice is always made according to the *temptation utility* v with ties being broken in favor of the normative utility function u. Clearly, an overwhelming temptation representation can be thought of as a dual self representation (u, v, ρ) with $\rho = 1$. Thus, the dual self representation presented in this paper can be viewed as the natural generalization of the overwhelming temptation representation of GP.

Subsequent to Chatterjee and Krishna (2005), Dekel and Lipman (2007) have addressed the possibility of one of multiple selves making the choice in the second period in a different way. They introduce another generalization of self-control and the overwhelming temptation representations of GP, which they call the *random Strotz representation*. It consists of a collection of expected utility functions $(u, (w(t))_{t \in T})$, where T is a (Borel measurable) index set, and a measure μ on T, so that the utility of a menu is given by

$$V_{RS}(x) = \int_T \max_{\beta \in \mathrm{B}_{w(t)}(x)} u(\beta)\mu(dt).$$

DL show that for any self control representation (u, v) and the corresponding utility function V_{SC}, there exists a random Strotz representation V_{RS}, such that $V_{SC} = V_{RS}$. They achieve this by letting $w(t) := v + tu$, $t \in [0, 1] =: T$, and taking μ to be the uniform measure over $[0, 1]$.

Such a representation is termed a random Strotz representation after Strotz's consistent planning approach (see Strotz 1955), where an agent makes a choice today, expecting to make a choice tomorrow according to a different utility function from the one he has today. Unfortunately, the random Strotz representation of V_{SC}, while analytically appealing, is hard to interpret behaviorally, since the "selves" in the random Strotz representation are not selves in the sense of Strotz (1955). Strotz (1955) envisions selves who are independent, while the selves in the random Strotz representation are not entirely independent.

[6] In particular, Chatterjee and Krishna (2005) show that if ρ_x is continuous in x and the representation is linear, then ρ is constant, and that if the preference is continuous, then ρ_x cannot be constant. Thus, continuous preferences that are linear and admit (menu dependent) dual self representations must impose some additional conditions on ρ_x.

To make this point analytically, let us say that the utility function $w(t)$ is *relevant*[7] if for any $\varepsilon > 0$ there exist menus x and y such that $\max_{\beta \in x} w_s(\beta) = \max_{\beta \in y} w_s(\beta)$ for all $s \notin (t-\varepsilon, t+\varepsilon)$ and $x \not\sim y$. (Notice that we do not require $B_{w(s)}(x) = B_{w(s)}(y)$.) Intuitively, a state is relevant if for some menu x changing the (maximum) utilities in only a neighborhood of state $w(t)$ results in a menu y that is valued differently than the menu x.

By this definition, there is no $t \in (0, 1)$ such that $w(t)$ is relevant, making it difficult to interpret the random Strotz representation.[8] Nevertheless, by this definition, both u and v are relevant in the dual self representation presented in this paper and the self-control representation of GP.

A. *Other Papers*

Several recent papers have focused on the problems raised by Schelling. The paper closest in spirit to ours is the innovative paper by B. Douglas Bernheim and Antonio Rangel (2004) in which they specifically deal with addiction and are clear that, in their view, the individual who takes drugs is making a mistake caused by overestimating the amount of pleasure consumption would involve relative to the long-term costs of such consumption. The selves are not treated symmetrically. Drug consumption is anomalous and abstaining from it rational. Their model also explicitly takes into account the effect of environmental cues in triggering the change of the controlling self from *cold* to *hot*. Here the cold self is supposed to be the preference that usually represents the agent, while the hot self is the one who makes the anomalous choices. In the next section, we suggest a way of presenting their insights in the framework of our model.

Roland Benabou and Marek Pycia (2002) note that GP's interpretation of their self-control representation (that of choosing from a menu in the second period using a utility function $u + v$ while being tempted to choose according to utility function v) can also be considered as the Nash equilibrium outcome of a conflict between divided selves, where there is some risk of succumbing to temptation in the second period. They do not provide an axiomatic account of this interpretation, adopt an explicitly dual self model for these dynamic choice problems, and focus on the game between the selves rather than on the axiomatization of a virtual dual self model, as we do here. Kfir Eliaz and Ran Spiegler (2006) study contracting issues with several preference representations, including dual selves.[9]

The spirit of our approach (i.e., considering the planning stage of a choice problem) is introduced in Kreps (1979) who characterizes preferences that value flexibility while looking at a finite set of prizes. DLR (2001) extend Kreps' model by looking at lotteries over a finite set of prizes and menus that are sets of lotteries. GP (2001) extend this environment to a compact metric space of prizes to provide a characterization of temptation. They also emphasize the importance of not breaking the link between choice and welfare. Larry Samuelson and Jeroen Swinkels (2006) explore

[7] The notion of relevance introduced is the natural analogue of a *relevant state* introduced by DLR (2001) in their study (and axiomatization) of subjective state models.

[8] The states $w(t)$ for $t \in (0, 1)$ are not relevant in the sense of DLR (2001) either.

[9] In this context, also see Susanna Esteban and Eiichi Miyagawa (2005, 2006), and Esteban, Miyagawa, and Matthew Shum (2007).

the evolutionary foundations of temptation. They develop a model where endowing humans with utilities of menus that depend on unchosen alternatives is an optimal choice for nature from an evolutionary perspective.

B. *Exogenous States of the World*

Our representation admits a straightforward extension to finite exogenous states. This would be similar in spirit to the model studied by Bernheim and Rangel (2004) when limited to two periods. Formally, let S be a finite set of states with the probability that state $s \in S$ occurs being given by π_s. The state is realized after the decision maker chooses the menu. We take this to be some set of exogenous circumstances that affect the agent only inasmuch as they affect the likelihood of her getting tempted. Note that the agent's utility function does not change across states nor does her alter ego's. The only thing that changes is the probability of getting tempted. In particular, we are looking for a utility function (over menus) that looks like the following:

$$U(x) = \sum_{s \in S} \pi_s \left((1 - \rho_s) \max_{\beta \in x} u(\beta) + \rho_s \max_{\beta \in B_v(x)} u(\beta) \right).$$

If we let $S := \{0, 1, \dots, n\}$ and let ρ_s be the probability of getting tempted in the state of the world s, then one specification could be the following:

$$\rho_s < \rho_{s+1},$$

as in the Bernheim-Rangel model.

IV. Conclusion

In this paper, we consider a decision maker who has to decide on the set of feasible choices from which an actual choice will be made at a later point in time. We rule out the case where the decision maker may prefer larger sets of feasible choices due to a preference for flexibility.

Our main contribution is to provide axioms on first period preferences that enable us to interpret this problem as a decision maker who behaves as if he has an alter ego (with preferences different from her own), who makes the actual choice from the menu with some probability. Doing so enables us to address problems where decision makers demonstrate apparent dynamic inconsistency (i.e., make ex post choices that are inferior from an ex ante perspective) and make unambiguous welfare statements in these situations. A clear shortcoming of the paper is that we can only talk about a single alter ego. Unfortunately, our methods prevent us from extending the results directly to the analysis of multiple alter egos. It seems that discontinuous preferences pose analytical challenges distinct from the kind already encountered (with continuous preferences), pointing toward interesting questions for future research.

Appendix: Proofs

A. *Proof of Dual Self Theorem*

Here, we shall demonstrate the "only if" part of the proof. We first begin with a crucial lemma which is an extension of Lemma 1 in DLR. Recall that for any set x, its (closed) convex hull is denoted by conv(x). (We shall only use the closed convex hull in what follows, and so shall refer to the closed convex hull as the convex hull.) Recall that $\mathcal{A}$ is the set of all essentially finite sets, i.e., $x \in \mathcal{A}$ if (and only if) x is finite, or x is the convex hull of finitely many elements.

LEMMA V1: *Let* $\succcurlyeq$ *satisfy* Independence. *Then for all finite* $x \in \mathcal{A}$, $x \sim \text{conv}(x)$.

PROOF:

From Lemma 1 in DLR, it follows that for any finite set x, $x \sim \text{conv}(x)$.

We shall now show that there exists a continuous linear functional that represents preferences. (This is essentially Proposition 2 in DLR.)

LEMMA V2: *If* $\succcurlyeq$ *satisfies Axioms* 1, 2A–C *and* 3, *there exists an upper semicontinuous linear functional* $U : \mathcal{A} \to \mathbb{R}$ *that represents* $\succcurlyeq$. *Furthermore, U is unique up to positive affine transformation.*

PROOF:

Let $X \subset \mathcal{A}$ be the space of all closed *convex hulls* of finitely many points in Δ. Notice that X endowed with the Minkowski sum is a mixture space. It only remains to verify the mixture space axioms (see Kreps 1988, 52). By assumption, *Independence* holds. Axioms 2A–C ensure that von Neumann-Morgenstern (vN-M) continuity is also satisfied. Thus, by the mixture space theorem, there exists a linear functional $V : X \to \mathbb{R}$ so that for all $x, y \in X$, $V(x) \geq V(y)$ if and only if $x \succcurlyeq y$.

We now extend V to all menus. Let us define $U : \mathcal{A} \to \mathbb{R}$ as follows: for all $x \in \mathcal{A}$, let $U(x) := V(\text{conv}(x))$. It is easily seen that U represents $\succcurlyeq$. All that remains to be shown is that U is linear.

From Lemma V1, it follows that $\lambda x + (1 - \lambda)y \sim \text{conv}(\lambda x + (1 - \lambda)y)$. Also $x \sim \text{conv}(x)$ and $y \sim \text{conv}(y)$. From *Independence*, it follows that $\lambda x + (1 - \lambda)y \sim \lambda\, \text{conv}(x) + (1 - \lambda)y$ and $\lambda\, \text{conv}(x) + (1 - \lambda)y \sim \lambda\, \text{conv}(x) + (1 - \lambda)\, \text{conv}(y)$, i.e., $\lambda x + (1 - \lambda)y \sim \lambda\, \text{conv}(x) + (1 - \lambda)\, \text{conv}(y)$. Therefore,

$$\begin{aligned} U(\lambda x + (1 - \lambda)y) &= U(\lambda\, \text{conv}(x) + (1 - \lambda)\, \text{conv}(y)) \\ &= V(\lambda\, \text{conv}(x) + (1 - \lambda)\, \text{conv}(y)) \\ &= \lambda V(\text{conv}(x)) + (1 - \lambda)\, V(\text{conv}(y)) \\ &= \lambda U(x) + (1 - \lambda)\, U(y), \end{aligned}$$

which is the desired result.

Let us define $u(\alpha) := U(\{\alpha\})$ and interpret it to be the decision maker's utility from a lottery (in the untempted state). It is clear that u is a continuous, linear function. Another important property of preferences that we shall make use of is translation invariance. This is made precise below.

DEFINITION 1: *A binary relation $\succcurlyeq$ is translation invariant if $x \succcurlyeq y$ implies $x + c \succcurlyeq y + c$ for all signed measures c such that $c(Z) = 0$ and $x + c$, $y + c \in \mathcal{A}$. Such a c will be called an admissible translate.*

LEMMA V4: *Let $\succcurlyeq$ satisfy Axioms* 1, 2A–C *and* 3. *Then, $\succcurlyeq$ is translation invariant.*

PROOF:

As before, let $X \subset \mathcal{A}$ be the space of all closed convex hulls of finitely many points in Δ. Notice, again, that X endowed with the Minkowski sum is a mixture space.

Take any $x, y \in X$ and admissible translate c. For an arbitrary $\beta \in \Delta$, $(½)(x + c) + (½)\{\beta\} = (½)x + (½)(\{\beta\} + c)$ so that $(½)U(x + c) + (½)U(\{\beta\}) = (½)U(x) + (½)U(\{\beta\} + c)$. A similar equality holds for $y + c$ and β, which gives us $U(x) \geq U(y)$ if and only if $U(x + c) \geq U(y + c)$. Since $\mathrm{conv}(x + c) = \mathrm{conv}(x) + c$, it follows that $\mathrm{conv}(x + c) \sim \mathrm{conv}(x) + c$.

Suppose $x, y \in \mathcal{A}$ and $x \succcurlyeq y$. Then, $\mathrm{conv}(x) \succcurlyeq \mathrm{conv}(y)$ and for an admissible translate c, $\mathrm{conv}(x) + c \succcurlyeq \mathrm{conv}(y) + c$. Then $x + c \sim \mathrm{conv}(x + c) \sim \mathrm{conv}(x) + c \succcurlyeq \mathrm{conv}(y) + c \sim \mathrm{conv}(y + c) \sim y + c$.

LEMMA V5: *Suppose Axioms* 1, 2A–C, 3–6 *hold. Then there exists a continuous, linear functional $v : \Delta \to \mathbb{R}$ such that $\{\beta\} \succ \{\alpha, \beta\} \succcurlyeq \{\alpha\}$ if and only if $v(\beta) < v(\alpha)$ and $u(\beta) > u(\alpha)$, and for all x, there exists $\hat{\beta}_x \in x$ such that $U(x) \geq u(\hat{\beta}_x)$ and $u(\hat{\beta}_x) = \max_{\beta \in B_v(x)} u(\beta)$.*

PROOF:

See Appendix Section B.

Now, there exists $\beta_x^* \in x$ so that $u(\beta_x^*) \geq U(x)$ from where we can determine a ρ_x using the Intermediate Value Theorem, such that

$$U(x) = (1 - \rho_x) \max_{\beta \in x} u(\beta) + \rho_x \max_{\beta \in B_v(x)} u(\beta).$$

Let $\beta_x^* \in B_u(x)$ and $\hat{\beta}_x \in B_u(B_v(x))$. Then, in order that the U be linear, the function ρ_x must satisfy

$$\rho_{\lambda x + (1-\lambda) y} = \frac{\lambda \rho_x \delta_x + (1 - \lambda)\rho_y \delta_y}{\lambda \delta_x + (1 - \lambda)\delta_y},$$

where $\delta_x := u(\beta_x^*) - u(\hat{\beta}_x)$ and δ_y is defined similarly.

We now prove a simple lemma which shows that we can restrict attention to essentially finite menus that lie entirely in the relative interior of Δ. (In what follows, we shall denote the ε-neighborhood (in Δ) of a point $\beta \in \Delta$ by $N_\varepsilon(\beta)$ and the diameter of a set x by diam(x). We shall also repeatedly use the fact that ρ must be consistent with the linearity of U, i.e., (♣) holds.) For any $y \in \mathcal{A}$, $F_y := \text{aff}(y) \cap \Delta$ is a compact convex subset of Δ (where aff(y) is the affine hull of y in the vector space $\mathcal{M}(Z)$).

The lemma shows that for any essentially finite set y, there exists another menu x that lies in the relative interior of F_y. Also, the diameter of x can be made arbitrarily small, and $\rho_y = \rho_x$. This result is useful because restricting attention to such an x enables us to consider arbitrary perturbations of x that lie in aff(x). In what follows, $ri(F)$ refers to the relative interior of the set F.

LEMMA V6: *For all $y \in \mathcal{A}_0$ and for all $\overline{\varepsilon} > 0$, there exists $x \in \mathcal{A}_0$ so that $x \subset ri\ F_y$, $\rho_y = \rho_x$ and diam$(x) < \overline{\varepsilon}$.*

PROOF:

Let $y \in \mathcal{A}_0$ and let $\hat{\beta}$ be an extreme point of conv(y). Also, let $\overline{\varepsilon} > 0$. Then, for all $\lambda \in (0, 1)$, $\rho_{\lambda\{\hat{\beta}\}+(1-\lambda)y} = \rho_y$. Moreover, for all $\varepsilon > 0$, there exists $\lambda_\varepsilon \in (0, 1)$ such that $\lambda_\varepsilon\{\hat{\beta}\} + (1 - \lambda)y \subset N_\varepsilon(\hat{\beta}) \cap F_y$. Let us now take $\varepsilon \in (0, \overline{\varepsilon}/2)$ so that for some $\beta^* \in ri\ F_y$, $N_\varepsilon(\beta^*) \subset ri\ F_y$. Let $c := \beta^* - \hat{\beta}$ be a signed measure so that $c(Z) = 0$. By the translation invariance property of U (and therefore of ρ), it follows that for $x := \lambda_\varepsilon\{\hat{\beta}\} + (1 - \lambda_\varepsilon)y + c$, $\rho_x = \rho_{\lambda_\varepsilon\{\hat{\beta}\}+(1-\lambda)y} = \rho_y$.

LEMMA V7: *Let $\succcurlyeq$ have a dual self representation and satisfy SoM. Then, for all finite x, for any $\beta \in B_u(x)$, and for any $\alpha \in B_u(B_v(x))$, $x \sim \{\beta, \alpha\}$.*

PROOF:

Let $\hat{x} := \{\beta_1, \ldots, \beta_m\} \cup x' \cup \{\alpha_1, \ldots, \alpha_n\} \cup y$, where $\beta_i \in B_u(\hat{x})$for $i = 1, \ldots, m$, $\alpha_j \in B_u(B_v(\hat{x}))$ for $j = 1, \ldots, n$, $u(\beta_1) > u(\gamma) \geq u(\alpha_1)$ for all $\gamma \in x'$, and $u(\alpha_1) > u(\gamma')$ for all $\gamma' \in y$. (Note that by definition, $v(\alpha_1) > v(\beta_i)$, and $v(\alpha_1) \geq v(\gamma')$for all $\gamma' \in y$.)

Recall that $\beta \in \hat{x}$ is *untempted in* $\hat{x}$ if there is no $\alpha \in \hat{x}$ that tempts β. Since α_1 is untempted in $\hat{x}$ (which means, among other things, that $\{\alpha_1\} \sim \{\alpha_1\} \cup \{\alpha_2, \ldots, \alpha_m\} \cup y$), by *AoM*, $\hat{x} \sim \{\beta_1, \ldots, \beta_n\} \cup x' \cup \{\alpha_1\}$. Also, by *AoM*, $\hat{x} \sim \{\beta_1, \ldots, \beta_m\} \cup x' \cup \{\alpha_1\}$. By *SoM*, $\hat{x} \sim \{\beta_1, \ldots, \beta_m\} \cup \{\alpha_1\}$. Let $x := \{\beta_1, \ldots, \beta_m\} \cup \{\alpha_1\}$. From Lemma V6, we can assume, without loss of generality, that $x \subset ri\ F_x$.

Let (β_i^k) be a sequence in $ri\ F_x$ such that $\{\beta_i^k\} \succ \{\beta_i^{k+1}\} \succ \{\beta_i\}$, and $\beta_i^k \in \text{conv}(\{\beta_i^1, \alpha_1\})$ for each k and $\lim_{k\to\infty} \beta_i^k = \beta_i$. Since $x \subset ri\ F_x$, it is clear that such a sequence always exists.

Let $x_i^k = \{\beta_i^k, \alpha_1\}$. By, *SoM*, it follows that $\{\beta_1, \ldots, \beta_i^k, \ldots, \beta_m\} \cup \{\alpha_1\} \sim x_i^k$. Furthermore, $x_i^k = \lambda_k x_i^1 + (1 - \lambda_k)\{\alpha_1\}$ for some $\lambda_k \in (0, 1)$ and $\lambda_k + \lambda_{k+1}$. Therefore, $\rho_{x_i^k} = \rho_{x_i^1}$. This implies that $x_i^k \succ x_i^{k+1}$. By *Upper-Semicontinuity*, it now follows that

$$x = \lim_k \{\beta_1, \ldots, \beta_i^k, \ldots, \beta_m\} \cup \{\alpha_1\} \sim \lim_k x_i^k = x_i,$$

which gives us the desired result.

LEMMA V8: *Let $\succcurlyeq$ have a dual self representation and satisfy SoM. Then for all β, α, α' such that $\{\beta\} \succ \{\alpha\} \sim \{\alpha'\}$ and α, α' tempt β, it is the case that $\{\beta, \alpha\} \sim \{\beta, \alpha'\}$. Hence, $\rho_{\{\beta,\, \alpha\}} = \rho_{\{\beta,\, \alpha'\}}$.*

PROOF:

By Lemma V6, we can assume that $\beta \in ri\ F_{\{\beta,\alpha,\alpha'\}}$, so there exists $\varepsilon > 0$ such that $N_\varepsilon(\beta) \subset ri\ F_{\{\beta,\alpha,\alpha'\}}$. We can also assume that $\alpha, \alpha' \in N_{\varepsilon/4}(\beta)$.

Let $v(\beta) < v(\alpha) < v(\alpha')$, and let $c := \alpha' - \alpha$, $c' := \alpha - \beta$. By hypothesis, α tempts β, so that $\alpha + c = \alpha'$ tempts $\beta + c = \beta'$. By *Translation Invariance*, $\{\beta, \alpha\} \sim \{\beta', \alpha'\}$. Also, $\{\beta\} \sim \{\beta'\}$. To see this, suppose the contrary, i.e., suppose $\{\beta\} \nsim \{\beta'\}$. By definition, $\beta' = \beta + c = \beta + (\alpha' - \alpha)$. By *Translation Invariance*, $\{\beta\} + c' \nsim \{\beta'\} + c'$, i.e., $\{\beta\} + (\alpha - \beta) \nsim \{\beta\} + (\alpha' - \alpha) + (\alpha - \beta)$, which is equivalent to $\{\alpha\} \nsim \{\alpha'\}$, contradicting the hypothesis.

Hence, β, β', α', and α form the vertices of a parallelogram. By Lemma V7, $\{\beta, \alpha'\} \sim \{\beta, \beta', \alpha'\} \sim \{\beta', \alpha'\}$. This proves that $\{\beta, \alpha\} \sim \{\beta, \alpha'\}$. Since $\{\alpha\} \sim \{\alpha'\}$, it follows from the representation that $\rho_{\{\beta,\, \alpha\}} = \rho_{\{\beta,\, \alpha'\}}$.

PROOF OF DUAL SELF THEOREM:

From Lemma V7, it follows that for any x, there exist elements $\beta, \alpha \in x$ such that $\{\beta, \alpha\} \sim x$. Therefore, we can restrict attention to two element subsets. Let $x = \{\beta, \alpha\}$ and $y = \{\beta', \alpha'\}$, where $x, y, \in ri\ F_{x \cup y}$, $\varepsilon = \text{diam}(x) \geq \text{diam}(y) > 0$ and $N_\varepsilon(\beta) \cap \Delta \subset ri\ F_{x \cup y}$.

Let $c := \beta - \beta'$. Then, $y + c \subset N_\varepsilon(\beta)$ and $\rho_y \sim \rho_{y+c}$. If $u(\alpha) > u(\alpha' + c)$, then there exists $\lambda \in (0, 1)$ so that $u(\lambda\beta + (1 - \lambda)(\alpha' + c)) = u(\alpha)$. Appealing to Lemma V8 now gives us the desired result. (The case where $u(\alpha) \leq u(\alpha' + c)$ is dealt with in a similar fashion.)

B. *The Alter Ego's Preferences*

In this section, we shall construct the alter ego's preferences via some revealed preference arguments thereby providing a proof of Lemma V5.

Let us define $\beta_+ := \{\alpha : \{\beta\} \succ \{\beta, \alpha\} \succcurlyeq \{\alpha\}\}$. From *Regularity*, it follows that β_+ is convex. Let us also define $\beta_- := \text{cl}\ (\{\alpha : \{\beta, \alpha\} \sim \{\beta\} \succ \{\alpha\}\})$. The lemma below shows that β_- is also convex.

LEMMA V9: *Suppose $\succcurlyeq$ satisfies Axioms* 1*,* 3 *and* 6*. Then, β_- is convex.*

PROOF:

Let $\alpha_1, \alpha_2 \in \{\alpha : \{\beta, \alpha\} \sim \{\beta\} \succ \{\alpha\}\}$. By *Independence* and $\{\beta\} \sim \{\beta, \alpha_2\}$,

$$\{\beta\} \sim \lambda\ \{\beta\} + (1 - \lambda)\ \{\beta, \alpha_2\}.$$

Independence and $\{\beta\} \sim \{\beta, \alpha_1\}$ also implies

$$\lambda\ \{\beta\} + (1 - \lambda)\ \{\beta, \alpha_2\} \sim \lambda\ \{\beta, \alpha_1\} + (1 - \lambda)\ \{\beta, \alpha_2\}.$$

Transitivity of $\succcurlyeq$ implies

$$\{\beta\} \sim \lambda\{\beta, \alpha_1\} + (1 - \lambda)\{\beta, \alpha_2\}.$$

But note that

$$\lambda\{\beta, \alpha_1\} + (1 - \lambda)\{\beta, \alpha_2\} = \{\beta, \lambda\alpha_1 + (1 - \lambda)\beta, \lambda\beta + (1 - \lambda)\alpha_2, \lambda\alpha_1 + (1 - \lambda)\alpha_2\}.$$

Applying *AoM* twice, we find $\{\beta\} \sim \{\beta, \lambda\alpha_1 + (1 - \lambda)\alpha_2\}$. Since $\{\beta\} \succ \{\alpha_1\}$, *Independence* gives us $\{\beta\} \succ (1 - \lambda)\{\beta\} + \lambda\{\alpha_1\}$. Also, $\{\beta\} \succ \{\alpha_2\}$ and *Independence* implies $(1 - \lambda)\{\beta\} + \lambda\{\alpha_1\} \succ \lambda\{\alpha_1\} + (1 - \lambda)\{\alpha_2\}$. By the transitivity of $\succ$, $\{\beta\} \succ \{\lambda\alpha_1 + (1 - \lambda)\alpha_2\}$. Thus, $\{\alpha : \{\beta, \alpha\} \sim \{\beta\} \succ \{\alpha\}\}$ is convex and so its closure β_- is also convex.

Let us recall some definitions of objects in linear spaces. An *affine subspace* (or linear variety) of a vector space is a translation of a subspace. A *hyperplane* is a *maximal* proper affine subspace. If H is a hyperplane in the vector space $\mathcal{M}(Z)$, then there is a linear functional f on $\mathcal{M}(Z)$ and a constant c such $H = \{x : f(x) = c\}$. Moreover, H is closed if and only if f is continuous by Lemma 5.42 (Charalambos D. Aliprantis, and Kim C. Border 1999). For notational ease, we shall write H as $[f = c]$. Similarly, (two of) the negative and positive half spaces are represented as $[f \leq c]$ and $[f > c]$, respectively. For any subset $S \subset \mathcal{M}(Z)$, let aff(S) denote the (closed) affine subspace generated by S, i.e., the smallest (closed) affine subspace that contains S. (In what follows, if $\beta_+^* = \varnothing$, let $v = u = U|_\Delta$. Henceforth, we shall assume β_+^* is not empty.) We shall also denote the zero vector by $\boldsymbol{\theta}$. We shall first demonstrate that the alter ego's preferences can be represented on any finite dimensional subset of Δ. For this, we need some definitions. Let $x \in \mathcal{A}$, and define $F_x := \text{aff}(x) > \Delta$. If x is essentially finite, i.e., if x is the convex hull of a finite subset of Δ, then F_x is a finite dimensional convex subset of Δ. (This is because the affine hull of a finite set in a topological vector space is finite dimensional.)

LEMMA V10: *For any $x \in \mathcal{A}_0$, let $\beta^* \in ri\ F_x$. Then there exists $v_x : F_x \to \mathbb{R}$ which is continuous and linear so that $\beta_-^* \cap F_x \subset [v_x \leq v_x(\beta^*)]$ and $\beta_+^* \cap F_x \subset [v_x > v_x(\beta^*)]$.*

PROOF:

Since $\beta_-^* \cap F_x$ and $\beta_+^* \cap F_x$ are disjoint convex subsets of a finite dimensional Hausdorff linear space (which we can take to be $\text{aff}(F_x) - \beta^* = \text{aff}(x) - \beta^*$), there exists a continuous linear functional that separates them. We denote this functional by v_x.

We have thus far established that, for some $\beta^* \in F_x$, there exists a continuous linear functional v_x that represents the alter ego's preferences at that point. We will now show that there is a single continuous, linear functional that represents the alter ego's preferences over all of F_x. (We shall use *Translation Invariance* towards this end.) Note that for $\beta \in ri\ F_x$, there exists $\varepsilon > 0$ such that $N_\varepsilon(\beta) \cap \Delta \subset ri\ F_x$. Also, recall a fact about the Hausdorff metric, d_h. For all $\lambda \in [0, 1]$, $d_h(\{\beta\}, \{\beta, \lambda\alpha + (1 - \lambda)\beta\}) = \lambda d_h(\{\alpha\}, \{\beta\})$.

LEMMA V11: *For all* $\beta \in F_x$, $[v_x = v_x(\beta)]$ *separates* $\beta_- \cap F_x$ *and* $\beta_+ \cap F_x$.

PROOF:

Suppose not. Then, there exists $\beta \in F_x$, such that either

(i) $\exists\, \alpha \in \beta_- \cap F_x$ such that $v_x(\alpha) \geq v_x(\beta)$, or

(ii) $\exists\, \alpha \in \beta_+ \cap F_x$ such that $v_x(\beta) \geq v_x(\alpha)$.

Let us consider the first possibility.

Let $c = \beta^* - \beta$. Since $\beta^* \in ri\ F_x$, there exists $\varepsilon > 0$ such that $N_\varepsilon(\beta^*) \cap \Delta \subset ri\ F_x$. From *Translation Invariance*, we can assume $\alpha \in N_\varepsilon(\beta) \cap F_x$. Thus, $\alpha + c \in N_\varepsilon(\beta^*) \cap F_x$. Since v_x is continuous and linear, $v_x\,(\alpha + c) > v_x\,(\beta + c) = v_x\,(\beta^*)$. This implies that $\{\beta + c\} \succ \{\beta + c,\, \alpha + c\}$ which, by *Translation Invariance*,[10] is equivalent to $\{\beta\} \succ \{\beta, \alpha\}$, which is a contradiction of the hypothesis that $\alpha \in \beta_-$.

The second possibility is taken care of with a similar argument, thus establishing the desired result.

We now define the binary relation $R \subset \Delta \times \Delta$ as follows. Suppose $\{\beta\} \succsim \{\alpha\}$. Define $\alpha\ R\ \beta$ if (i) $\exists\ (\alpha_n)$ such that $\alpha_n \to \alpha$ and α_n tempts β for each n, or $\exists\ (\beta_n)$ such that $\beta_n \to \beta$ and α tempts β_n for all n. Also define $\beta\ R\ \alpha$ if $\exists\ (\alpha_n)$ such that $\alpha_n \to \alpha$ and α_n does not tempt β for any n, or $\exists\ (\beta_n)$ such that $\beta_n \to \beta$ and α does not tempt β_n for any n. (Recall that α tempts β if $\{\beta\} \succ \{\beta, \alpha\}$ and α does not tempt β if $\{\beta\} \sim \{\beta, \alpha\}$.)

Notice that for any $x \in \mathcal{A}_0$, $R|_{F_x}$ is represented by v_x. We shall show that R has an expected utility representation. We first claim that R is complete and transitive.

CLAIM V12: *R is complete and transitive.*

PROOF:

For any $\alpha, \beta \in \Delta$, let $\mathcal{A}_0 \ni x \ni \alpha, \beta$ so that $\dim(F_x) \geq 2$. It is easy to see now that either $\alpha\ R\ \beta$ or $\beta\ R\ \alpha$ so that R is complete. To see that R is transitive, suppose not, so that there exists $\alpha, \beta, \gamma \in \Delta$ such that $\alpha\ R\ \beta$ and $\beta\ R\ \gamma$ but $\neg(\alpha\ R\ \gamma)$. Consider again $x \in \mathcal{A}_0$ such that $\alpha, \beta, \gamma \in x$ and $\dim(F_x) \geq 2$. Since $R|_{F_x}$ is represented by v_x, this is impossible.

We shall now show that R is continuous.

CLAIM V13: *R is continuous.*

PROOF:

For any $\beta \in \Delta$, define $\beta^- := \{\alpha \in \Delta : u(\alpha) > u(\beta)$ and $\{\alpha\} \succ \{\alpha, \beta\}\}$ and $\beta^+ := \{\alpha \in \Delta : u(\alpha) \geq u(\beta)$ and $\{\alpha\} \sim \{\alpha, \beta\}\}$. Thus, the R-lower contour set at β is $\beta_- \cup \mathrm{cl}(\beta^-)$ and similarly for the upper contour set. We shall show that R is

[10] Notice that we only require *Translation Invariance* to hold for two-element subsets.

continuous by demonstrating that β_- is closed and β_+ is open. A symmetric argument for the other regions will complete the proof.

We first recall that $U : \mathcal{A} \to \mathbb{R}$ is upper semicontinuous. To show that β_- is closed, consider a convergent sequence (α_n) in β_- and let $\alpha_n \to \alpha$. Then, by definition of β_- and from the upper semicontinuity of U, $U(\{\alpha, \beta\}) = \limsup_n U(\{\alpha_n, \beta\}) = U(\{\beta\})$ so that $\alpha \in \beta_-$ which proves that β_- is closed.

To show that β_+ is open, fix $\alpha \in \beta_+$ and let $\varepsilon = U(\{\beta\}) - U(\{\alpha, \beta\}) > 0$. Then, by the upper semicontinuity of U, there exists $\delta > 0$ such that for all $\alpha' \in N_\delta(\alpha)$, $U(\{\alpha', \beta\}) < U(\{\alpha, \beta\}) + \varepsilon = U(\{\beta\})$. Hence, β_+ is open in Δ.

Recall that R, by definition, satisfies *Translation Invariance*. Chatterjee and Krishna (2008) show that a preference relation R that is continuous and is translation invariant has an expected utility representation.[11] Let $v : \Delta \to \mathbb{R}$ represent R.

LEMMA V14: *For all finite x,*

$$\max_{\beta \in x} u(\beta) \geq U(x) \geq \max_{\beta \in B_v(x)} u(\beta).$$

PROOF:

Let $\beta^* \in B_u(x)$ and let $\hat{\beta} \in B_u B_v(x)$. Let $x' := \{\alpha \in x : u(\alpha) \geq u(\hat{\beta})\}$and $y := \{\alpha \in x : u(\alpha) < u(\hat{\beta})\}$. Then, $x = x' \cup y$ and by *Temptation*, $\beta^* \succcurlyeq x' \succcurlyeq \hat{\beta}$. Let $y := \{\alpha_1, \alpha_2, \dots, \alpha_n\}$. It then follows that for each $\alpha_i \in y$, $\{\hat{\beta}\} \sim \{\hat{\beta}, \alpha_i\}$. By *AoM*, it follows that $\{\hat{\beta}\} \sim \{\hat{\beta}, \alpha_1, \alpha_2\}$. Repeatedly applying *AoM* implies $\{\hat{\beta}\} \sim \{\hat{\beta}\} \cup y$. Once again applying *AoM* implies $\{\hat{\beta}\} \cup x' \cup y \sim \{\hat{\beta}\} \cup x' = x'$. Thus, $x \sim x'$ and $\{\beta\} \succcurlyeq x \succcurlyeq \{\hat{\beta}\}$.

REFERENCES

Aliprantis, Charalambos D., and Kim C. Border. 1999. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. 2nd ed. New York: Springer.

Amador, Manuel, Iván Werning, and George-Marios Angeletos. 2006. "Commitment versus Flexibility." *Econometrica*, 74(2): 365–96.

Anscombe, Frank J., and Robert J. Aumann. 1963. "A Definition of Subjective Probability." *Annals of Mathematical Statistics*, 34(1): 199–205.

Benabou, Roland, and Marek Pycia. 2002. "Dynamic Inconsistency and Self-Control: A Planner-Doer Interpretation." *Economics Letters*, 77(3): 419–24.

Bernheim, B. Douglas, and Antonio Rangel. 2004. "Addiction and Cue-Triggered Decision Processes." *American Economic Review*, 94(5): 1558–90.

Chatterjee, Kalyan, and R. Vijay Krishna. 2005. "Menu Choice, Environmental Cues and Temptation: A 'Dual Self' Approach to Self-control." NayEcon Working Paper.

Chatterjee, Kalyan, and R. Vijay Krishna. 2008. "A Geometric Approach to Continuous Expected Utility." *Economics Letters*, 98(1): 89–94.

Dekel, Eddie. 1986. "An Axiomatic Characterization of Preferences under Uncertainty: Weakening the Independence Axiom." *Journal of Economic Theory*, 40(2): 304–18.

Dekel, Eddie, and Barton L. Lipman. 2007. "Self-control and Random Strotz Representations." Unpublished.

[11] It follows, therefore, that for a continuous preference relation, *Independence* is equivalent to *Translation Invariance*. This is easy to see once one notices that *Translation Invariance* implies *Betweenness* in the sense of Dekel (1986), from which it is easy to show that *Independence* follows.

Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini. Forthcoming. "Temptation-Driven Preferences." *Review of Economic Studies.*

Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini. 2001. "Representing Preferences with a Unique Subjective State Space." *Econometrica*, 69(4): 891–934.

Eliaz, Kfir, and Ran Spiegler. 2006. "Contracting with Diversely Naïve Agents." *Review of Economic Studies*, 73(3): 689–714.

Esteban, Susanna, and Eiichi Miyagawa. 2005. "Optimal Menu of Menus with Self-Control Preferences." Unpublished.

Esteban, Susanna, and Eiichi Miyagawa. 2006. "Temptation, Self-Control, and Competitive Nonlinear Pricing." *Economics Letters*, 90(3): 348–55.

Esteban, Susanna, Eiichi Miyagawa, and Matthew Shum. 2007. "Nonlinear Pricing with Self-Control Preferences." *Journal of Economic Theory*, 135(1): 306–38.

Fudenberg, Drew, and David K. Levine. 2006. "A Dual-Self Model of Impulse Control." *American Economic Review*, 96(5): 1449–76.

Gul, Faruk, and Wolfgang Pesendorfer. 2001. "Temptation and Self-Control." *Econometrica*, 69(6): 1403–35.

Hammond, Peter J. 1976. "Changing Tastes and Coherent Dynamic Choice." *Review of Economic Studies*, 43(1): 159–73.

Kreps, David M. 1979. "A Representation Theorem For 'Preference for Flexibility.'" *Econometrica*, 47(3): 565–77.

Kreps, David M. 1988. *Notes on the Theory of Choice*. Boulder, CO: Westview Press.

Samuelson, Larry, and Jeroen Swinkels. 2006. "Information and the Evolution of the Utility Function." *Theoretical Economics*, 1(1): 119–42.

Schelling, Thomas C. 1978. "Egonomics, or the Art of Self-Management." *American Economic Review*, 68(2): 290–94.

Schelling, Thomas C. 1984. "Self-Command in Practice, in Policy, and in a Theory of Rational Choice." *American Economic Review*, 74(2): 1–11.

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

This article has been cited by:

1. S. N. Ali. 2011. Learning Self-Control. *The Quarterly Journal of Economics* **126**, 857-893. [CrossRef]
2. Wojciech Olszewski. 2011. A model of consumption-dependent temptation. *Theory and Decision* **70**, 83-93. [CrossRef]