Menu Choice, Environmental Cues and Temptation: A "Dual Self" Approach to Self-control*

Kalyan Chatterjee† R. Vijay Krishna‡

June, 2007

Abstract

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. This eventuality need not occur with certainty but might be triggered by the nature of the set chosen in the first period. 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 with some probability. We provide a discussion of the behavioural implications of our model as compared with existing theories.

§ 1 Introduction

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

*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 and especially Bart Lipman for helpful comments and Avantika Chowdhury for research assistance.

†Department of Economics, Pennsylvania State University. *email*: `kchatterjee@psu.edu`, *web page*: `http://www.econ.psu.edu/people/biographies/chatterjee_bio.htm`

‡Department of Economics, University of North Carolina, Chapel Hill. *email*: `rvk@email.unc.edu`, *web page*: `http://www.unc.edu/~rvk`

cannot make anyone strictly worse off. Strotz (1955), who discusses intertemporal choice, calls this constraining of one's future choices the "strategy of pre-commitment" and suggests that it results from the agent today being a "different person with a different discount function from the agent of the past" (p 168). Schelling (1978, 1984) gives several examples of this kind of behaviour and also speaks of "multiple selves," though he writes also that he is somewhat hesitant to use this terminology outside the circle of professional economists. 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 maximise 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 he might be in the mood for either seafood or vegetarian food for dinner (depending on his changed taste in the evening), even though he would go for seafood if he were forced to choose a menu item in the morning. It is clear that he would prefer making a reservation at a restaurant that serves both rather than one that serves just seafood. If, instead, the agent perceives that future choices might be made that maximise a utility function different from his, he might prefer to limit his options. For instance, if someone had to choose between going to a party with tea and coffee being served, versus another one in which tea, coffee and cocaine would be served (to modify Amartya Sen's example), the fact that this person might choose cocaine in the evening would possibly lead to him pre-committing in the morning to going 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 maximise a different utility function.

One way to conceptualise both types of problems is to consider, following 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 behaviour, which the individual considers a possibility. The menu chosen might trigger temptation in the next period. Temptation is thought of as the (utility maximising) choice of a *virtual* alternate self or *alter ego*. The probability of the alternate self taking over is menu-dependent.[1] 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

[1]We can think of this as a generalisation of Strotz, in that the alternate self appears here with some probability instead of with certainty, as in Strotz.

about the *choice* made (although we shall assume that the alter ego breaks ties in favour 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 (i) a long-run self's utility function $u(\cdot)$; (ii) a virtual self's utility function $v(\cdot)$ and (iii) a probability of getting tempted ρ, which could be menu dependent. 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. If it turns out, for example, that $u(\cdot)$ and $v(\cdot)$ give the same choices from a menu x, then ρ_x is irrelevant for that menu (there are no temptations in it) and we assume $\rho_x = 0$ in this case.

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-maximisers in x. The decision maker believes she will be tempted with probability ρ_x when faced with menu x, which results in the alter ego making a choice. This induces a preference, indeed a utility function, over menus, given by

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

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 is most preferred by the decision-maker. Since we are characterising 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 the equation above. Thus, a decision-maker who satisfies our axioms behaves *as if* there is a probability of her getting 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 emphasised that the alter ego (and his utility function v) is subjective, as is the probability, ρ_x, of getting tempted. The only observables are first period choices over menus and v and ρ_x 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) $v(a) > v(b)$, ie the alter ego prefers b to a and (ii) $\rho_{\{a,b\}} > 0$, ie 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 either of the requirements above does not hold, then temptation would not be an issue.

We will have more to say later on about the nature of the dependence on x of ρ_x. At this stage, we can mention that $U(x)$ could depend solely on the best and the maximally tempting elements in a menu (ie the u- and v-maximisers respectively) in a very simple way, namely with a constant ρ (this is analysed in Theorem 3.6); however, ρ_x also, more generally, could depend on elements in the menu x other than the u- and v-maximisers (for example, the presence of sorbet in a menu makes it easier to be tempted by full-fledged ice cream). This is our Dual Self Theorem, Theorem 3.3. Our axioms will also imply certain restrictions on ρ_x resulting from the linearity of the utility function, given the dual self representation. We consider ρ_x to be potentially dependent on the menu x. The case where ρ_x is constant (independent of x) is an important *special case*, which we consider in detail in §3.1. It is worthwhile to note that the well-known representation of Gul and Pesendorfer (2001) of temptation preferences implies our representation with ρ_x dependent on the menu. The case where $\rho > 0$ is constant and not equal to 1 is not capable of being captured by the Gul-Pesendorfer preferences (as discussed in §4.1).

We emphasise that this paper seeks to characterise behaviour that exhibits temporary loss of self-control by determining a set of axioms on 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" α is just the von Neumann-Morgenstern utility $u(\alpha)$.[2]

Our model consists of the two-period environment of GP and DLR and axioms designed to translate the motivating story into a utility representation. These axioms are therefore somewhat different from those present in the existing literature. However, since the different motivating stories relate to the same environment, there should be some relation between the different sets of axioms. We explore this in §4.

[2]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.

The remainder of the paper is structured as follows. In §2 we introduce our model, in §3 the axioms and our representation theorem and sketch the proof of the representation theorem in §3.2. In §4 we first compare our axioms with those of Gul and Pesendorfer (2001) (henceforth GP) in §4.1 and with those of Dekel, Lipman and Rustichini (2006) in §4.2 (henceforth DLR05) and then review other related literature. §5 concludes and proofs are in the Appendix.

§ 2 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 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 $\alpha, \beta, \ldots$) and is endowed with the topology of weak convergence. This topology is metrisable and we let d_p be a metric which generates this topology. The objects of analysis are subsets of Δ. Let $\mathscr{A}$ be the set of all closed subsets of Δ (with generic elements, called *menus*, denoted by $x, y, \ldots$) endowed with the Hausdorff metric

$$d_h(x,y) := \max\left\{\max_x \min_y d_p(\alpha,\beta), \max_y \min_x d_p(\alpha,\beta)\right\}.$$

Convex combinations of elements $x, y \in \mathscr{A}$ 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 $\mathscr{A} \times \mathscr{A}$. (In the sequel, read $A \to B$ as "A implies B" unless the arrow is a limit. In either case, it should be clear from the context what the intended arrow denotes and no confusion should arise.)

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 Anscombe and 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, pp. 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, a seafood fancier who goes to a restaurant not knowing the quality of the shrimp he is about to get, is doing the same. It is also possible that the menu of lotteries could arise from a non-degenerate 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.

§ 3 Axioms and Representations

We first define the linear functionals relevant to a dual self representation. As is standard, we shall say that $U : \mathscr{A} \to \mathbb{R}$ is *linear* if $U(\lambda x + (1-\lambda)y) = \lambda U(x) + (1-\lambda)U(y)$ for all $x, y \in \mathscr{A}$ and $\lambda \in (0,1)$ and that it *represents* $\succcurlyeq$ if it is the case that $U(x) \geqslant U(y)$ 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-maximisers in x (with a similar definition for B_u). Let $\beta^*_x \in B_u(x)$ and let $\widehat{\beta}_x \in B_u\left(B_v(x)\right)$.

For any menu x, we shall say in what follows that the decision-maker, when confronted with a choice from the menu x, gets "tempted" with probability ρ_x. If ρ is to be consistent with linearity of U, then it must be the case that for all $\lambda \in (0,1)$,

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

where $\delta_x := u(\beta^*_x) - u(\widehat{\beta}_x)$ and $\delta_y := u(\beta^*_y) - u(\widehat{\beta}_y)$. (The expression follows from the linearity of u and v and the observation that if β^*_x (resp. $\widehat{\beta}_x$) maximises u (resp. v) over x and if β^*_y (resp. $\widehat{\beta}_y$) maximises u (resp. v) over y, then $\lambda\beta^*_x + (1-\lambda)\beta^*_y$ (resp. $\lambda\widehat{\beta}_x + (1-\lambda)\widehat{\beta}_y$) maximises u (resp. v) over $\lambda x + (1-\lambda)y$. We return to preferences and impose the following axioms on them.

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

Axiom 2 (Continuity) The sets $\{y : y \succcurlyeq x\}$ and $\{y : x \succcurlyeq y\}$ 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. Axiom 2 is a continuity requirement in the Hausdorff topology. The motivation for *Independence* is the familiar one and some normative arguments in its favour 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 Fudenberg and Levine (2006), this may not be an innocuous assumption.) Clearly, axioms 1–3 deliver a continuous linear functional $U : \mathscr{A} \to \mathbb{R}$ that represents $\succcurlyeq$. Our next axiom captures the essence of temptation.

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

Temptation says that insofar as the presence of alternatives different from the best alternative in the menu affects the decision-maker, it does not make the decision-maker worse off than her worst choice in the menu. It also says that the cost of temptation (ie the cost of not being able to choose the best alternative) is bounded. In particular, it rules out situations like Sen's rational donkey, which starves because it is unable to make a choice between two equally acceptable alternatives. In other words, it is never the case that "analysis is paralysis."

At this point, it is useful to provide a preliminary definition of a dual self representation and relate it to axioms 1–4.

3.1 Definition. A ***weak dual self representation*** is a triple (u, v, ρ), where $u, v : \Delta \to \mathbb{R}$ are continuous and linear, $\rho : \mathscr{A} \to [0, 1]$ satisfies (♣) and $U : \mathscr{A} \to \mathbb{R}$ given by

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

represents $\succcurlyeq$.

In what follows, we shall refer to U as being induced by the dual self representation. It is easy to see that any $U : \mathscr{A} \to \mathbb{R}$ that satisfies axioms 1–4 has a weak dual self representation. Indeed, let $u(\alpha) := U(\{\alpha\})$ and $v = -u$. By *Temptation*, there exists ρ such that $U(x) = (1-\rho_x) \max_{\beta \in x} u(\beta) + \rho_x \max_{\beta \in B_v(x)} u(\beta) = (1-\rho_x) \max_{\beta \in x} u(\beta) + \rho_x \min_{\beta \in x} u(\beta)$. Moreover, from the linearity of U, it follows that ρ satisfies (♣). But this weak dual self representation is not particularly useful as it tells us nothing about which lotteries are tempting. Consider, for instance, a preference $\succcurlyeq$ satisfying Axioms

1–4 with lotteries α, β such that $\{\beta\} \sim \{\alpha, \beta\} \succ \{\alpha\}$. Here, $\succcurlyeq$ has a weak dual self representation and $u(\beta) > u(\alpha)$, $v(\alpha) > v(\beta)$ and $\rho_{\{\alpha,\beta\}} = 0$ and yet $\{\beta\} \sim \{\alpha, \beta\}$. This ascribes a meaning to the alter ego other than merely representing the presence of suboptimal choices (since the alter ego never makes the choice), which is definitely not our intention.

Moreover, there may be other triples (u', v', ρ') that represent U, leading to a very unsatisfactory representation. We shall now provide additional axioms on preferences and strengthen our definition of a weak dual self representation to rule out trivial representations as described above and deliver a unique representation.

Consider three prizes, broccoli (b), rich chocolate cake (c) and deep-fried Mars bars (m). 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 in 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, ie $\{\beta\} \succ \{\alpha, \beta\} \succcurlyeq \{\alpha\}$. More generally, say that y *tempts* β if $\{\beta\} \succ \{\beta\} \cup y$. Thus, at the very least, y contains some elements that tempt β.

The first part of our 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 $1/2 - 1/2$ 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.[3]

Axiom 5 (Regularity) $\{\alpha_1\}$ and $\{\alpha_2\}$ tempt $\{\beta\}$ implies (i) $\{\lambda\alpha_1 + (1-\lambda)\alpha_2\}$ 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 once again the three prizes, broccoli (b), rich chocolate cake (c) and deep-fried Mars bars (m). As before, 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, ie $\{b\} \succ$

[3] 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.

$\{b,c\},\{b,m\}$. 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, ie $\{c\} \sim \{c,m\}$. 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, (ie $\{b,c\} \sim \{b,c,m\}$). 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 (AoM: Additivity of Menus) 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 $\{\beta\} \cup x$.[4] This has the flavour 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. Requiring it to hold for arbitrary menus will neither change any of the theorems below nor simplify their proofs.

Now that we have stated our axioms, we present the definition of a dual self representation that is relevant to us. But let us first impose another requirement on ρ. For x, y finite, if $\beta \in B_u B_v(\{\beta\} \cup x \cup y)$, $u(\beta) > u(\alpha)$ for all $\alpha \in y$ and $\beta \notin B_u(\{\beta\} \cup x \cup y)$,

$$\rho_{\{\beta\} \cup x \cup y} = \rho_{\{\beta\} \cup x}. \qquad (\spadesuit)$$

To see the point of this requirement, let us use the dual self point of view. With our interpretation, temptation is the transfer of decision making rights to the alter ego. This transfer is stochastic and depends on the set of alternatives present. The requirement ($\spadesuit$) says that the probability that this event occurs does not depend on elements of the menu that can never be chosen by either the agent or his alter ego. Indeed, both the agent and the alter ego would prefer β to any $\alpha \in y$. This brings us to the following definition.

3.2 Definition. A ***dual self representation*** is a triple (u, v, ρ), where $u, v : \Delta \to \mathbb{R}$ are continuous and linear, $\rho : \mathscr{A} \to [0,1]$ satisfies ($\clubsuit$), ($\spadesuit$), has $B_u(\{\alpha,\beta\}) \cap B_v(\{\alpha,\beta\}) = \varnothing$ if and only if $\rho_{\{\alpha,\beta\}} > 0$ and $U : \mathscr{A} \to \mathbb{R}$ given by

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

represents $\succcurlyeq$. It is called *continuous* if the U defined above is continuous.

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

As mentioned above, (♣) is a structural requirement that ensures the linearity of U. The motivation for (♠) is the same as that for *AoM*. Finally, the condition $B_u(\{\alpha,\beta\}) \cap B_v(\{\alpha,\beta\}) = \varnothing$ if and only if $\rho_{\{\alpha,\beta\}} > 0$ says that the alter ego's utility function corresponds exactly to our notion that the alter ego represents temptation. We can now state our main representation theorem.

3.3 Theorem (Dual Selves)**.** *A binary relation $\succcurlyeq$ satisfies Axioms 1–6 if and only if it admits an essentially unique dual self representation (u, v, ρ). Moreover, the induced U is continuous.*

The dual self representation is essentially unique in the following sense: u and v are unique up to positive affine transformation and ρ_x is unique[5] for all x where temptation is meaningful.

Proof. We shall sketch the "if" part of the proof here and the uniqueness result. The "only if" part is in the appendix. First, for any (u, v, ρ) with $U(x) := (1 - \rho_x)\max_{\beta\in x} u(\beta) + \rho_x \max_{\beta\in B_v(x)} u(\beta)$, it is clear that U is continuous and linear and that $U(\{\alpha\}) = u(\alpha)$ for all $\alpha \in \Delta$. Notice that $U(\{\beta\}) > U(\{\beta,\alpha\}) > U(\{\alpha\})$ if and only if $u(\beta) > u(\alpha)$, $v(\alpha) > v(\beta)$ and $\rho_{\{\alpha,\beta\}} > 0$. Thus, α tempts β if and only if $v(\alpha) > v(\beta)$ and $\rho_{\{\alpha,\beta\}} > 0$. It is now easy to see that U satisfies *AoM* (from (♠)) and *Regularity*. It is also easily seen that $U(x) \geqslant U(\widehat{\beta}_x)$ where $\widehat{\beta}_x \in B_u B_v(x)$ so that U also satisfies *Temptation*. Now to prove uniqueness.

Suppose, by way of contradiction, $\succcurlyeq$ admits another dual self representation (u', v', ρ'). Clearly, it must be the case that u' is not a positive affine transformation of u (else we get an immediate contradiction), so assume that $u' = u$. Now suppose v' is not a positive affine transformation of v. Then, there exist α, β such that $v(\beta) > v(\alpha)$ and $v'(\alpha) > v'(\beta)$. Without loss of generality, we can take α, β such that $u(\beta) > u(\alpha)$. Define $U'(x) = (1-\rho'_x)\max_{\beta\in x} u(\beta) + \rho_x \max_{\beta\in B_{v'}(x)} u(\beta)$. We have claimed that $U' = U$. Notice that $U(\{\beta\}) = U(\{\alpha,\beta\})$, whereas $U'(\{\beta\}) > U'(\{\alpha,\beta\})$, thus yielding the desired contradiction. □

The dual self theorem below says that when faced with choices of menus, the decision-maker who satisfies Axioms 1–6 behaves *as if* he 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 maximises the decision-maker's utility. Also, the decision-maker behaves *as if* he will be

[5] This follows since $U|_\Delta = u$ so that u and U are unique up to the same positive affine transformation.

tempted (ie the probability that the choice will be made by the alter ego) with a probability of ρ_x 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_x represents the probability that she loses this battle. It is entirely conceivable that ρ_x depends on more than the u and v maximisers in x. For instance, consider an 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 $U(\{f, c\}) > U(\{f, s, c\})$, ie the presence of more temptations increases the probability that the decision maker will lose the battle for self control.

It should be emphasised that ρ_x and v are subjective and hence unobservable. They have to be inferred from first period choices. The only observables here are first period behaviour. Furthermore, the decision-maker behaves as if second period choice from a convex menu is from an extreme face of the menu. (That the decision-maker is indifferent between any menu and its convex hull is proved in Lemma 6.1 in the Appendix.)

We should point out that in the dual self representation above, ρ cannot be constant as this would violate the continuity of $\succcurlyeq$. To see this, suppose ρ is constant and $U(\cdot)$ is continuous. Consider an α and β such that $u(\beta) > u(\alpha)$ and $v(\alpha) = v(\beta)$ so that $U(\{\beta\}) = U(\{\beta, \alpha\})$. Let (α_k) be a sequence such that (i) $\lim_k \alpha_k = \alpha$, (ii) $u(\alpha_k) = u(\alpha)$ for all k and (iii) $v(\alpha_k) > v(\alpha_{k+1}) > v(\alpha)$ for all k. Then, $U(\{\beta\}) > U(\{\beta, \alpha_k\}) = U(\{\beta, \alpha_{k+1}\})$ for all k but, as noted above, $U(\{\beta\}) = U(\{\beta, \alpha\})$ which contradicts the continuity of $U(\cdot)$. Moreover, the following is also true.

3.4 Proposition. *Let ρ satisfy equation (♣) and be continuous. Then ρ is constant.*

Proof. Suppose ρ is continuous and suppose it is not constant. Then, there exist menus x, y such that $\rho_x \neq \rho_y$. Then, for any singleton $\{\beta\}$, $\rho_{\lambda\{\beta\}+(1-\lambda)x} = \rho_x$ and $\rho_{\lambda\{\beta\}+(1-\lambda)y} = \rho_y$. (This follows from equation (♣) above.) But for any $\varepsilon > 0$, there exists λ large enough such that

$$d_h\left(\lambda\{\beta\} + (1-\lambda)x, \lambda\{\beta\} + (1-\lambda)y\right) < \varepsilon$$

but $\rho_{\lambda\{\beta\}+(1-\lambda)x}$ and $\rho_{\lambda\{\beta\}+(1-\lambda)y}$ remain just as far apart, contradicting the continuity of ρ. □

Thus, no matter what additional axioms we choose, we cannot have continuous ρ except if ρ is constant. Furthermore, if ρ is not constant, then it must be the case that $\succcurlyeq$ is discontinuous. We examine the case of constant ρ in the next section.

§ 3.1 *Other Representations*

The dual self representation allows many possibilities. Indeed, it encompasses the GP representation (see §4.1 below) and allows ρ to depend on arbitrarily many items in a menu. We can, nevertheless, say that there can be *no* continuous selection ρ. This was demonstrated in Proposition 3.4. Moreover, the dual self theorem rules out situations where the decision maker has no self control, ie decision makers who have a dual self representation with $\rho_x = 1$ for all x. This is because such a decision maker's preferences are typically only upper semicontinuous (and not continuous). In this section, we introduce another excision axiom which will give us this possibility and also the case where $\rho_x \in [0, 1]$ is a constant for all x. Let us first weaken *Continuity* appropriately.

Axiom 2a: (Upper Semicontinuity) The sets $\{y \in \mathscr{A} : y \succcurlyeq x\}$ 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.

Axioms 2a–c are identical to Axioms 2a–c in §3 of GP. They weaken *Continuity* just enough to enable us to have a linear utility representation that admits a dual self representation. Notice also that Axioms 2a–c are strictly weaker than Axiom 2. It is worthwhile to record what we lose by not requiring the preference relation to satisfy *Continuity*. Let $\mathscr{A}_0 \subset \mathscr{A}$ be the collection of all sets x that are either finite or the convex hulls of finite sets. Thus, $x \in \mathscr{A}_0$ if and only if x is finite or x is the convex hull of a finite set. We shall refer to $\mathscr{A}_0$ as the collection of *essentially finite* subsets of Δ. This leads to the following representation.

3.5 Theorem. *A binary relation $\succcurlyeq$ satisfies Axioms 1, 2a–c, 3–6 if and only if it admits an essentially unique dual self representation. Moreover, the induced U is upper semicontinuous.*

Proof. The proof follows from the proof of Theorem 3.3. □

Observe that the representation above permits ρ to be constant. We now proceed towards the axiom which isolates this property. We have already seen one kind of excision in *AoM*. Another kind of excision is the notion that the only items in a menu that matter to a decision-maker are the alternative he would have chosen were he not tempted and the item in the menu that causes him 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 formalise this below. Let us say that $\beta \in x$ is *tempted* if there exists $y \subset x$ such that $\{\beta\} \succ \{\beta\} \cup y$.

Axiom 7 (SoM: Separability of Menus) 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 in 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). This is a strong assumption, but it has the advantage of providing a lot of structure to the dual self representation. As with *AoM*, we only require *SoM* to hold for finite menus. Here too, requiring it for arbitrary menus neither changes Theorem 3.6 below nor simplifies its proof.

3.6 Theorem. *A binary relation $\succcurlyeq$ satisfies Axioms 1, 2a–c, 3–7 if and only if it admits an essentially unique dual self representation where $\rho_x = \rho$ for all $x \in \mathscr{A}$ 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 6.2. □

We note again that *SoM* is a strong axiom and some readers may not find it particularly compelling. Nevertheless, from a consequentialist perspective, this axiom gives us the utility of a menu as characterised by (a fixed convex combination of) the best option and the most tempting option. For the readers who find behaviour of this type intuitively appealing as a description, *SoM* is necessary and sufficient to generate it. We now turn to a brief sketch of the proof of Theorem 3.3.

§ 3.2 *Proof-sketch of Theorem 3.3*

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 of arguments which we describe below.

1. *Representing* $\succcurlyeq$. An application of the mixture space theorem (lemma 6.2) shows that *Preferences, Continuity* and *Independence* guarantee the existence of a continuous linear functional U unique up to affine transformation which represents $\succcurlyeq$. Also U restricted to singletons is continuous.

2. *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 $\beta_- := \text{cl}\left(\{\alpha : \{\beta\} \sim \{\beta, \alpha\} \succ \{\alpha\}\}\right)$ is also convex. 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.

3. *Translation Invariance.* We say that U is *translation invariant* if $U(x) \geqslant U(y)$ if and only if $U(x+c) \geqslant U(y+c)$ for all signed measures c such that $c(Z) = 0$ and $x+c, y+c \in \mathscr{A}$. That U is translation invariant follows from *Continuity* and *Independence*. We use this property to show that there is an essentially unique linear functional which performs the separation described in the previous step for each lottery β. Thus, there exists a continuous linear functional which represents the alter ego.

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

5. *Arbitrary menus.* Using *Continuity*, we show that for all x, it is the case that $u(\beta_x^*) \geqslant U(x) \geqslant u(\widehat{\beta}_x)$ where β_x^* and $\widehat{\beta}_x$ are defined as above.

6. *Representation.* A simple application of the intermediate value theorem gives us the desired representation and ρ_x for each x.

§ 4 Related Literature

§ 4.1 *The Requirement of* Set Betweenness

Gul and Pesendorfer (2001) consider an agent who faces the same two-period problem as above. They introduce a condition on preferences called *Set Betweenness* (SB). This requirement says that for all menus $x, y \in \mathscr{A}$,

$$x \succcurlyeq y \longrightarrow x \succcurlyeq x \cup y \succcurlyeq y. \tag{SB}$$

Their representation theorem (as it pertains to us) says that if $\succcurlyeq$ satisfies axioms 1, 2a–c, 3 and *Set Betweenness*, then the utility of a menu is given by a function U_{GP} defined either as

$$U_{GP}(x) = \max_{\beta \in x} \{u(\beta) + v(\beta)\} - \max_{\beta \in x} v(\beta)$$

or

$$U_{GP}(x) = \max_{\beta \in B_v(x)} u(\beta)$$

where $U_{GP} : \mathscr{A} \to \mathbb{R}$ is linear and upper-semicontinuous, $u, v : \Delta \to \mathbb{R}$ are continuous, linear functionals unique up to the same affine transformation and $u = U_{GP}|_{\Delta}$. GP refer to u as the *commitment utility* and to v as the *temptation utility*. The first representation obtains if $\succcurlyeq$ is continuous and the second obtains if $\succcurlyeq$ is upper semi-continuous but not continuous. The second kind of representation described above is to account for the possibility of "overwhelming temptation" where the agent always succumbs to temptation. To better understand the representation, let us assume, for the moment, that $U_{GP}(x)$ is continuous. This immediately rules out the overwhelming temptation representation. Let us write

$$c(\beta, x) := \max_{\beta' \in x} v(\beta') - v(\beta)$$

and interpret $c(\beta, x)$ to be the cost imposed by the temptation whenever β is chosen from x. We can now rewrite

$$U_{GP}(x) = \max_{\beta \in x} \{u(\beta) - c(\beta, x)\}$$

which says that the utility to the decision maker of a menu is determined (additively) by the utility of the best choice in the menu and the cost it imposes through its selection. The interpretation is that in the second period, the agent makes the choice which maximises the utility function $u + v$, which represents a compromise between the agent's commitment utility and his temptation utility.[6]

In spite of the very different motivation underlying GP's model, it is surprising to note that an agent satisfying axioms 1, 2a–c, 3 and *Set Betweenness* also satisfies our axioms. (Whether he satisfies *SoM* or not depends on whether or not the preferences are continuous.) To see this, suppose $U_{GP}(\cdot)$ is continuous so that for any menu x,

$$\max_{\beta\in x} u(\beta) \;\geqslant\; U_{GP}(x) \;\geqslant\; \max_{\beta\in B_v(x)} u(\beta)$$

whereby there exists some ρ_x such that

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

Moreover, ρ also satisfies the other requirements of a dual self representation so that $U_{GP}(x)$ admits a dual self representation.[7] The case where $U_{GP}(\cdot)$ is only upper semi-continuous, ie where

$$U_{GP}(x) = \max_{\beta\in B_v(x)} u(\beta)$$

corresponds to the case where $\rho_x = 1$ for all x, ie the decision-maker gets tempted with probability 1. Here too, $U_{GP}(\cdot)$ admits a dual self representation.

Thus, GP's representation implies our representation. Put another way, a decision-maker who satisfies Axioms 1–3 and *Set Betweenness* also behaves as if he has an alter ego who makes a choice with probability ρ_x when the menu chosen is x. Thus any differences between the two models can only be detected by looking at second period choice. Nevertheless, there exist preferences which have dual self representations but do not satisfy GP's axioms. This is demonstrated in Example 4.1 below. Before proceeding to the example, we need a little more simplifying algebra.

[6] However, this is just one interpretation. We can also write GP's utility function as $U_{GP} = \max_{\beta\in x} w(\beta) - \max_{\beta\in x} v(\beta)$ where $w = u + v$ and interpret this (following DLR) as follows: The decision maker believes that in the second period, she can be in two states of mind, that she will have utility function w in one state and utility function v in the other, and that she will make her second period choice from the menu according to which state of mind she is in. Her first period payoff is the weighted sum of the utility she receives in each state with weight $+1$ being assigned to the first state and weight -1 assigned to the second state.

[7] This can also be seen directly. It is easy to see that U_{GP} satisfies *Temptation* and *Regularity*. To see that it satisfies *AoM*, consider the finite menu $\{\beta\}\cup x\cup y$ where $\beta \in B_u B_v(\{\beta\}\cup x\cup y)$ and $\{\beta\} \succcurlyeq \alpha$ for all $\alpha \in y$. Then, $U_{GP}(\{\beta\}\cup x\cup y) = U_{GP}(\{\beta\}\cup x)$ since there is no $\alpha \in y$ such that α is the unique maximiser of either $u + v$ or v.

Consider any utility function U_{GP}. For any x, let $\beta^* \in B_u(x)$, $\overline{\beta} \in B_{u+v}(x)$ and $\widehat{\beta} \in B_uB_v(x)$. A little algebra then shows that since U_{GP} admits a dual self representation and there exists a ρ such that $U_{GP}(x) = (1-\rho_x)\max_{\beta\in x} u(\beta) + \rho_x \max_{\beta\in B_v(x)} u(\beta)$. Such a ρ_x is therefore given by

$$\rho_x = \min\left\{ \frac{[u(\beta^*) - u(\overline{\beta})] + [v(\widehat{\beta}) - v(\overline{\beta})]}{u(\beta^*) - u(\widehat{\beta})}, 1 \right\}. \qquad (\bigstar)$$

For a two-element menu $x = \{\alpha, \beta\}$ where $\beta \in B_{u+v}(x)$ and $\alpha \in B_u(x)$, $\rho_{\{\alpha,\beta\}} = \min\left\{ \frac{v(\alpha)-v(\beta)}{u(\beta)-u(\alpha)}, 1 \right\}$.

It is also useful to consider the behaviour of ρ in three element menus. Consider $x := \{\alpha, \beta, \gamma\}$ where $u(\alpha) > u(\beta) > u(\gamma)$, $v(\gamma) > v(\beta) > v(\alpha)$ and $\{\beta\} = B_{u+v}(x)$. Also define $\alpha_\lambda := (1-\lambda)\alpha + \lambda\beta$ and $x_\lambda := \{\alpha_\lambda, \beta, \gamma\}$. Then,

$$\rho_{x_\lambda} = \frac{[u(\alpha_\lambda) - u(\beta)] + [v(\gamma) - v(\beta)]}{u(\alpha_\lambda) - u(\gamma)}.$$

Notice that as $\lambda \uparrow 1$, $x_\lambda \to \{\beta, \gamma\}$ and $\rho_{x_\lambda} \downarrow \rho_{\{\beta,\gamma\}}$. In other words, if $1 > \lambda > \lambda' > 0$, $\rho_{x_\lambda} < \rho_{x_{\lambda'}}$, which is what we would expect. We are now ready to present the aforementioned example (from Dekel, Lipman and Rustichini (2006)).

4.1 Example. Consider a weak-willed dieter who faces choices over broccoli (b) multiple temptations in the form of rich chocolate ice cream (c) and low-fat yogurt (y). A natural set of rankings over menu is:

$$\{b, y\} \succ \{y\} \quad \text{and} \quad \{b, c, y\} \succ \{b, c\}.$$

Dekel, Lipman and Rustichini (2006) show that there is no GP representation that is consistent with the above ordering. In other words, any extension of the ordering above to the space of all menus would violate one of GP's axioms.

Nevertheless, there is a dual self representation that is consistent with the above ordering. To see this most transparently, let u and v be as following:

	b	c	y
u	6	0	4
v	0	8	6

so that utility over menus is given by

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

where $\widetilde{\rho}_x = (1/2)\rho_x$ and ρ_x is given by equation ($\bigstar$). We can interpret the utility function U as representing an agent who is uncertain about the probability with which he will be tempted, wherein with probability $1/2$ he has $\rho_x = 0$ for all x and with equal probability has ρ_x as in ($\bigstar$). $\diamondsuit$

We reiterate that for an agent who satisfies *Set Betweenness*, the implied second-period behaviour is vastly different under GP's interpretation as compared to ours. Under GP's interpretation (in the no overwhelming temptation case), second period choice is made according to the utility function $u + v$ and the agent knows with certainty in the first period that this will happen. Under our interpretation, the second period choice is made with some (menu dependent) probability according to one utility function and with complementary probability according to another and the agent does not know in the first period which event will occur. (Needless to say, this does not apply in case the agent has the overwhelming temptation representation in which case she knows with certainty that the choice will be made by her alter-ego.)

§ 4.2 *Temptation and Multiple States*

In another recent paper, Dekel, Lipman and Rustichini (2006) consider generalisations of GP preferences. They want to look at the class of preferences where the decision-maker is certain about the preferences of his untempted self. The only uncertainty is about what form the temptation takes. A further generalisation that they consider is in the form of the temptation cost. This is described below.

To model the uncertainty in temptation, DLR05 look at different linear functions $v_j : Z \to \mathbb{R}$ which give rise to different cost functions

$$c_i(\beta, x) := \sum_{j \in J_i} \max_{\beta' \in x} v_j(\beta') - \sum_{j \in J_i} v_j(\beta).$$

The key to uncertainty about temptation then, is that the decision-maker is uncertain about which cost function he will be facing. Thus, the generalisation of GP that obtains, namely the *temptation representation*, can be written as

$$U(x) := \sum q_i \max_{\beta \in x} \{u(\beta) - c_i(\beta, x)\}$$

where $q_i > 0$ and $\sum q_i = 1$ which means that q_i can be interpreted as the probability that the decision-maker is faced with the ith temptation in the form of the cost function c_i. (The temptation representation is a special case of a *finite state additive*

EU representation. For details, the reader is referred to Dekel, Lipman and Rustichini (2006). It suffices here to point out that a consequence of the finite state additive EU representation is that there exists an $N > 0$ such that for all menus x, there exists a submenu $x' \subset x$ such that $x' \sim x$ and $|x'| \leqslant N$.) The function u is such that $u(\beta) := U(\{\beta\})$ and is called the *commitment utility*. (All the functions u and v_j are expected utility functions.)

To characterise such a preference, DLR05 introduce two more axioms. The first says that if the decision-maker could commit himself to a certain item in a menu, he would. This is made precise in the following axiom.

Axiom (DFC: Desire for Commitment) There exists $\alpha \in x$ such that $\{\alpha\} \succcurlyeq x$.

DFC is equivalent to the first half of *Temptation*. (There exists preferences that admit temptation representations but do not satisfy *Temptation*. This is demonstrated in example 4.3 below.) To state the second axiom, we need some more definitions. Call β an *approximate improvement for* x if $\beta \in \operatorname{cl}(\{\beta' : x \cup \{\beta'\} \succ x\})$ where cl denotes closure. As before $B(x) := \{\alpha \in x : \{\alpha\} \succcurlyeq \{\beta\}, \forall\, \beta \in x\}$ is the set of best singletons in x. This gives us:

Axiom (AIC: Approximate Improvements are Chosen) If β is an approximate improvement for x, $x' \subset x$, and $\alpha \in B(x')$ satisfies $\{\alpha\} \succ \{\beta\}$, then $\{\alpha\} \succ x' \cup \{\beta\}$.

The main representation theorem in DLR05 is then the following:

4.2 Theorem. *A preference relation $\succcurlyeq$ has a temptation representation if and only if it has a finite state additive EU representation and satisfies* DFC *and* Dominance*.*

It is straightforward to show that our representation implies *AIC*. Therefore, it would seem that any *continuous* preference (in the Hausdorff topology) which has a dual self representation strictly belongs to the class of preferences identified in DLR05. But this is not the case, primarily because we do not have a finiteness axiom. But let us first look at a a couple of examples where the decision-maker satisfies the DLR05 axioms but does not have a dual self representation.

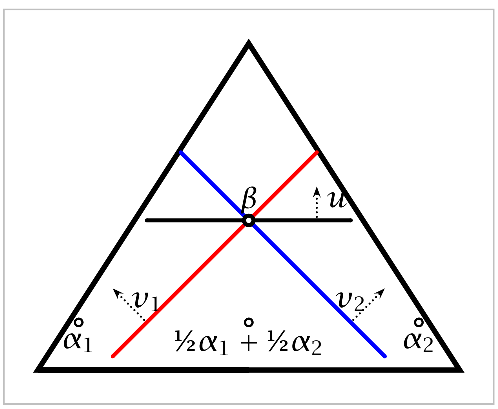
4.3 Example. Let Z be a finite set and let Δ be the space of lotteries over Z. Define utility function u, v_1 and v_2 so that

	u	v_1	v_2
α	0	2	2
β	0	1	6

Now let $c(\gamma, x) := \sum_j [\max_{\gamma' \in x} v_j(\gamma')] - \sum_j v_j(\gamma)$. Then $U(x) := \max_{\beta \in x} [u(\gamma) - c(\gamma, x)]$ is a temptation representation.

Let $x := \{\alpha, \beta\}$. Then $c(\alpha, x) = v_1(\alpha) + v_2(\beta) - v_1(\alpha) - v_2(\alpha) = 4$ and $c(\beta, x) = v_1(\alpha) - v_1(\beta) = 1$. This means that $U(x) = \max_{\gamma \in x} \{u(\gamma) - c(\gamma, x)\} = -1$. Thus, $\{\alpha\} \sim \{\beta\} \succ x$ which is not possible in a dual self representation (as *Temptation* is violated). ♢

Notice that if in the temptation representation, each cost function depends on only one temptation ie each J_i is a singleton, such an example could not arise. In a dual self representation, the decision maker does not care about the utility level of the alter ego. She only cares about the choice made by the alter ego insofar as it affects her own utility level. If we were to interpret the different v_j's in a temptation representation as belonging to different selves who are the cause of the temptation, then we could say that a temptation representation is an interdependent utilities model, where the decision maker (or the "commitment self" in the terminology of DLR05) cares about the utility levels of the different selves. We shall now see another example of a temptation representation which does not admit a dual self representation. This time our axiom *Regularity* will be violated.



4.4 Example. Suppose the decision-maker has utility function u and is faced with two temptations denoted by expected utility functions v_1 and v_2. Also, suppose that v_j is not a convex combination of v_i and u for $i \neq j$. Then, such a configuration might look as in the figure above. But such a configuration would violate our axiom *Regularity*. The violation is because the lottery α_1 and α_2 which both tempt β, but there also exist lotteries over α_1 and α_2 which do not tempt β. ♢

It should be remarked that the construction above holds in any temptation representation where there is uncertainty about the form the temptation will take. We shall now show that there exist preferences that have dual self representations but do not have temptation representations. Recall that GP preferences of the following form

$$U_{GP}(x) = \max_{\beta \in x}\{u(\beta) + kv(\beta)\} - k \max_{\beta \in x} v(\beta)$$

for $k \in [0, \infty)$, have a dual self representation (u, v, ρ) with

$$\rho_x(k) = \min\left\{ \frac{[u(\beta^*) - u(\overline{\beta})] + k[v(\widehat{\beta}) - v(\overline{\beta})]}{u(\beta^*) - u(\widehat{\beta})}, 1 \right\}.$$

where $\beta^* \in B_u(x)$, $\overline{\beta} \in B_{u+kv}(x)$ and $\widehat{\beta} \in B_v(x)$. Now, let (k_i) be a sequence such that $k_i \in [0, \infty)$ and let $y := \bigcup_i \{\beta_i\} \cup \{\alpha, \gamma\}$ where $\{\beta_i\} = B_{u+k_i v}(y)$, $\{\alpha\} = B_u(y)$ and $\{\gamma\} = B_v(y)$.

Let (λ_i) be another sequence such that $\lambda_i \in (0, 1)$ for each i and $\sum_i \lambda_i = 1$. Now, it easy to see that

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

where $\rho_x = \sum_i \lambda_i \rho_x(k_i)$, is part of a dual self representation with ρ_x satisfying all the necessary conditions for this to be so. By choosing the k_i's and λ_i's appropriately, it is therefore possibly to construct utility functions that admit dual self representations but not temptation representations. In other words, these utility functions are such that the utility of a menu can depend on an arbitrarily large (countable) sub-menu. This is because each $\rho(k_i)$ depends on $\{\alpha, \beta_i, \gamma\}$ and we can make the set of β_i's arbitrarily large. Indeed, this is the spirit of Example 4.1. We can further generalise this construction (as suggested by Bart Lipman). Let μ be a Borel probability measure on $(0, 1)$ and let $\rho_x = \int \rho_x(k)\, \mathrm{d}\mu(k)$ so that we have a dual self representation but not a temptation representation.

§ 4.3 *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 Bernheim and Rangel (2004), who 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. Fudenberg and Levine (2006) adopt an explicitly dual self model for these dynamic choice problems and focus on the game between the selves rather than on the axiomatisation of a virtual dual self model, as we do here. Eliaz and Spiegler (2006) study contracting issues with several preference representations, including dual selves.[8]

The spirit of our approach (ie considering the planning stage of a choice problem) is introduced in Kreps (1979) who characterises preferences that value flexibility while looking at a finite set of prizes. Dekel, Lipman and Rustichini (2001) extend Kreps' model by looking at lotteries over a finite set of prizes and menus that are sets of lotteries. Gul and Pesendorfer (2001) extend this environment to a compact metric space of prizes to provide a characterisation of temptation. They also emphasise the importance of not breaking the link between choice and welfare. Samuelson and Swinkels (2006) explore 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.

§ 4.4 *Exogenous States of the World*

Our representation admits a straightforward extension to finite exogenous states. This would be the formal equivalent of the model studied by Bernheim and Rangel (2004) 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 realised 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:

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

In particular, note that the probability of getting tempted can depend on *both the menu and the state*.

[8] In this context, also see Esteban and Miyagawa (2005a,b) and Esteban, Miyagawa and Shum (2003).

4.5 Example. Let $S := \{0, 1, \ldots, n\}$ and let ρ_i^x be the probability of getting tempted in the state of the world i. Then one specification could be the following:

$$\rho_x^i < \rho_x^{i+1}$$

for all x and $\rho_x^0 = 0$ for all x. If in addition we assumed $\rho_x^i = \rho_y^i$ for all $x, y \in \mathscr{A}$ for all $i \in S$, we would get the Bernheim-Rangel model. ♢

§ 5 Conclusion

In this paper, we consider a decision-maker faced 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 (ie make ex-post choices that are inferior from an ex-ante perspective) and make unambiguous welfare statements in these situations. We also relate our model to the papers of Gul and Pesendorfer (2001) and Dekel, Lipman and Rustichini (2006).

One possible extension of our model would be to allow the decision-maker to have multiple alter egos. With two alter egos, such a representation may be written as $U(x) = (1 - \rho_x - \widetilde{\rho}_x) \max_{\beta \in x} u(\beta) + \rho_x \max_{\beta \in B_v(x)} u(\beta) + \widetilde{\rho}_x \max_{\beta \in B_{\widetilde{v}}(x)} u(\beta)$, where the alter egos have utility functions v and $\widetilde{v}$ respectively. It can be shown that a modified version of *AoM* will hold in such a model. However, *Regularity* will not hold. In particular, consider utilities where there exist $\alpha, \widetilde{\alpha}, \beta$ such that $u(\beta) > u(\alpha) = u(\widetilde{\alpha})$, $v(\alpha) > v(\beta) > v(\widetilde{\alpha})$ and $\widetilde{v}(\widetilde{\alpha}) > \widetilde{v}(\beta) > \widetilde{v}(\alpha)$. Also, $v(\beta) > v(\frac{1}{2}\alpha + \frac{1}{2}\widetilde{\alpha})$ and $\widetilde{v}(\beta) > \widetilde{v}(\frac{1}{2}\alpha + \frac{1}{2}\widetilde{\alpha})$. Then, α and $\widetilde{\alpha}$ tempt β but $\frac{1}{2}\alpha + \frac{1}{2}\widetilde{\alpha}$ does not tempt β. Such choices can always be made if for some β, β_- is convex and has a non-empty algebraic interior. Thus *Regularity* can be seen to rule out an interesting class of preferences. Recall that we use the Intermediate Value theorem to determine ρ_x for each menu. The major difficulty in extending our results to a multiple self representation lies in our inability to infer the ρ's for each alter ego. We leave such investigations for future research.

§ 6 Appendix: Proofs

§ 6.1 *Proof of Theorem 3.3*

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 $\mathrm{conv}\,(x)$. (We shall only use the closed convex hull in what follows, and so shall refer to the closed convex hull as the convex hull.)

6.1 Lemma. *Let* $\succcurlyeq$ *satisfy* Independence*. Then for all finite* x*,* $x \sim \mathrm{conv}\,(x)$*. Furthermore, if* $\succcurlyeq$ *satisfies* Continuity*, then* $x \sim \mathrm{conv}\,(x)$ *for all* $x \in \mathscr{A}$*.*

Proof. >From Lemma 1 in DLR, it follows that for any finite set x, $x \sim \mathrm{conv}\,(x)$. Let x be an arbitrary closed set. Then (because x is compact), there exists a sequence (x_k) of finite sets such that $x_k \subset x$ and $x_k \to x$ (in the Hausdorff metric). Thus, $\mathrm{conv}\,(x_k) \to \mathrm{conv}\,(x)$. From the continuity of $\succcurlyeq$, we see that $x \sim \mathrm{conv}\,(x)$. □

We shall now show that there exists a continuous linear functional that represents preferences. (This is Proposition 2 in DLR.) Recall that $\mathscr{A}$ is the space of all closed subsets of Δ.

6.2 Lemma. *If* $\succcurlyeq$ *satisfies Axioms 1, 2 and 3, then there exists a continuous linear functional* $U : \mathscr{A} \to \mathbb{R}$ *that represents* $\succcurlyeq$*. Furthermore,* U *is unique up to affine transformations.*

Proof. Let $X \subset \mathscr{A}$ be the space of all closed *convex* subsets of 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, page 52). By assumption, *Independence* holds. *Continuity* ensures that 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 $x, y \in X$, $V(x) \geqslant V(y)$ if and only if $x \succcurlyeq y$.

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

>From lemma 6.1, it follows that $\lambda x + (1-\lambda)y \sim \mathrm{conv}\,(\lambda x + (1-\lambda)y)$. Also $x \sim \mathrm{conv}\,(x)$ and $y \sim \mathrm{conv}\,(y)$. From *Independence* it follows that $\lambda x + (1-\lambda)y \sim \lambda\,\mathrm{conv}\,(x) + (1-\lambda)y$ and $\lambda\,\mathrm{conv}\,(x) + (1-\lambda)y \sim \lambda\,\mathrm{conv}\,(x) + (1-\lambda)\,\mathrm{conv}\,(y)$, ie

$\lambda x + (1-\lambda)y \sim \lambda \operatorname{conv}(x) + (1-\lambda)\operatorname{conv}(y)$. Therefore,

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

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 us of is translation invariance. This is made precise below.

6.3 Definition. 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 $x+c, y+c \in \mathscr{A}$. Such a c will be called an **admissible translate**.

6.4 Lemma. *Let $\succcurlyeq$ satisfy Axioms 1 – 3. Then, $\succcurlyeq$ is translation invariant.*

Proof. As before, let $X \subset \mathscr{A}$ be the space of closed convex subsets of Δ. Notice again that X endowed with the Minkowski sum is a mixture space.

Take any $x, y \in X$ and admissible translate c.[9] For an arbitrary $\beta \in \Delta$, $(1/2)(x+c) + (1/2)\{\beta\} = (1/2)x + (1/2)(\{\beta\}+c)$ so that $(1/2)U(x+c) + (1/2)U(\{\beta\}) = (1/2)U(x) + (1/2)U(\{\beta\}+c)$. A similar equality holds for $y+c$ and β which gives us $U(x) \geqslant U(y)$ if and only if $U(x+c) \geqslant U(y+c)$.

Since the space of all measures on (Z,d), denoted by $\mathscr{M}(Z)$, is a locally convex topological vector space (endowed with the weak* topology), it is the case that for all $x \in \mathscr{A}$, and for all admissible translates c, $\operatorname{conv}(x+c) = \operatorname{conv}(x) + c$. (This follows from the fact that in a locally convex topological vector space, the convex hull of a compact set is closed. See Aliprantis and Border, 1999, lemma 5.57, pp. 193.) From the reflexivity of $\sim$, $\operatorname{conv}(x+c) \sim \operatorname{conv}(x) + c$.

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

[9] We thank an anonymous referee for providing us with this transparent proof, which simplifies an earlier argument.

6.5 Lemma. *Suppose Axioms 1–6 hold. Then there exists a continuous, linear functional* $v : \Delta \to \mathbb{R}$ *such that (i)* $\{\beta\} \succ \{\alpha, \beta\} \succcurlyeq \{\alpha\}$ *if and only if* $v(\beta) < v(\alpha)$ *and* $u(\beta) > u(\alpha)$*, and (ii) for all* x*, there exists* $\widehat{\beta}_x \in x$ *such that* $U(x) \geqslant u(\widehat{\beta}_x)$ *and* $u(\widehat{\beta}_x) = \max_{\beta \in B_v(x)} u(\beta)$*.*

Proof. See §6.1.1 below. □

Now, there exists $\beta^*_x \in x$ so that $u(\beta^*_x) \geqslant U(x)$ from where we can determine ρ_x using the Intermediate Value Theorem, which completes the proof. The other properties of ρ are also easily obtained.

§ 6.1.1 *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 6.5.

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

6.6 Lemma. *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 $\mathscr{M}(Z)$, then there is a linear functional f on $\mathscr{M}(Z)$ and a constant c such $H = \{x : f(x) = c\}$. Moreover, H is closed if and only if f is continuous (Lemma 5.42, Aliprantis and Border, 1999). For notational ease, we shall write H as $[f = c]$. Similarly, (two of) the negative and positive half spaces are represented as $[f \leqslant c]$ and $[f > c]$ respectively. For any subset $S \subset \mathscr{M}(Z)$, let aff (S) denote the (closed) affine subspace generated by S, ie 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 θ. 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 $x \in \mathscr{A}$ and define $F_x := \text{aff}\,(x) \cap \Delta$. If x is essentially finite, ie 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.)

6.7 Lemma. *For any $x \in \mathscr{A}_0$, let $\beta^* \in \text{ri } F_x$. Then there exists $v_x : F_x \to \mathbb{R}$ which is continuous and linear so that $\beta^*_- \cap F_x \subset [v_x \leqslant v_x(\beta^*)]$ 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 aff $(F_x) - \beta^* =$ 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 which represents the alter ego's preferences over all of F_x. (We shall use *Translation Invariance* towards this end.) Note that for $\beta \in \text{ri } F_x$, there exists $\varepsilon > 0$ such that $N_\varepsilon(\beta) \cap \Delta \subset \text{ri } F_x$. Also recall a fact about the Hausdorff metric, d_h. For all $\lambda \in [0, 1]$, $d_h\left(\{\beta\}, \{\beta, \lambda\alpha + (1 - \lambda)\beta\}\right) = \lambda d_h\left(\{\alpha\}, \{\beta\}\right)$.

6.8 Lemma. *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 $v_x(\alpha) \geqslant v_x(\beta)$, or
(ii) $\exists\ \alpha \in \beta_+ \cap F_x$ such that $v_x(\beta) \geqslant v_x(\alpha)$.
Let us consider the first possibility.

Let $c = \beta^* - \beta$. Since $\beta^* \in \mathrm{ri}\ F_x$, there exists $\varepsilon > 0$ such that $N_\varepsilon(\beta^*) \cap \Delta \subset \mathrm{ri}\ F_x$. >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 $\{\beta\} \not\succcurlyeq \{\alpha\}$. 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 $x \in \mathscr{A}_0$, $R|_{F_x}$ is represented by v_x. We shall show that R has an expected utility representation. We first claim that R is complete and transitive.

6.9 Claim. R is complete and transitive.

Proof. For any $\alpha, \beta \in \Delta$, let $\mathscr{A}_0 \ni x \ni \alpha, \beta$ so that $\dim(F_x) \geqslant 2$. 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 $x \in \mathscr{A}_0$ such that $\alpha, \beta, \gamma \in x$ and $\dim(F_x) \geqslant 2$. Since $R|_{F_x}$ is represented by v_x, this is impossible. □

We shall now show that R is continuous.

6.10 Claim. R is continuous.

Proof. For any $\beta \in \Delta$, define $\beta^- := \big\{\alpha \in \Delta\ :\ u(\alpha) > u(\beta) \text{ and } \{\alpha\} \succ \{\alpha, \beta\}\big\}$ and $\beta^+ := \big\{\alpha \in \Delta\ :\ u(\alpha) \geqslant u(\beta) \text{ and } \{\alpha\} \sim \{\alpha, \beta\}\big\}$. 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 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 $U : \mathscr{A} \to \mathbb{R}$ 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.

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

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 (2007) 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.

6.11 Lemma. *For all finite x*

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

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

6.12 Lemma. *For any x,*

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

Proof. Let $\beta^* \in B_u(x)$ and let $\widehat{\beta} \in B_u B_v(x)$. Let (x_k) be a sequence of finite sets where $|x_k| = k$, $x_k \subset x$, $x_k \to x$ and $\beta^*, \widehat{\beta} \in x_k$ for each k.

By *Temptation*, $u(\beta^*) \geqslant U(x_k)$ for each k. Hence, by *Continuity*, $u(\beta^*) \geqslant U(x)$. Also, for each k, $U(x_k) \geqslant u(\widehat{\beta})$. Once again, *Continuity* implies that $U(x) \geqslant u(\widehat{\beta})$. This gives us the desired result. □

§ 6.2 *Proof of Theorem 3.6*

We shall prove the theorem for finite menus. Towards this end, we show that there exists a linear functional that represents preferences over essentially finite menus (which are defined below). Also, for finite menus x, $x \sim \operatorname{conv}(x)$ and *Translation Invariance*

[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.

holds. *SoM* is then used to derive the representation for finite menus. A straightforward continuity argument then extends the result to arbitrary menus.

Recall that a set $x \subset \Delta$ is *essentially finite* if x is finite or if x is the convex hull of a finite set and $\mathscr{A}_0$ is the space of all essentially finite menus, ie the space of all essentially finite subsets of Δ. Also define $X_0 \subset X$ as the space of all closed, essentially finite, convex subsets of Δ. Before we begin, recall that lemma 6.1 shows that for each finite x, $x \sim \text{conv}\,(x)$.

6.13 Lemma. *Let $\succcurlyeq$ satisfy Axioms 1, 2a–c and 3. Then there exists an upper-semicontinuous linear functional $U : \mathscr{A}_0 \to \mathbb{R}$ such that for all $x, y \in \mathscr{A}_0$, $x \succcurlyeq y$ if and only if $U(x) \geqslant U(y)$. Also, U is unique up to affine transformation.*

Proof. The proof is similar to that of lemma 6.2. We shall only provide a sketch here. Notice that X_0 is a mixture space. Since *Independence* and von Neumann-Morgenstern continuity also hold, there exists a $V : X_0 \to \mathbb{R}$, unique up to affine transformation, so that for all $x, y \in X_0$, $x \succcurlyeq y$ if and only if $V(x) \geqslant V(y)$. Now define $U(x) := V(\text{conv}\,(x))$ for all $x \in \mathscr{A}_0$. Linearity of U is demonstrated as in lemma 6.2. Upper-semicontinuity follows from Axiom 2a. □

Another property that we shall establish for essentially finite menus is translation invariance.

6.14 Lemma. *Let $\succcurlyeq$ satisfy Axioms 1, 2a–c and 3. Then for all $x, y \in \mathscr{A}_0$, $x \succcurlyeq y$ implies $x + c \succcurlyeq y + c$ where c is an admissible translate.*

Proof. The proof is similar to the proof of lemma 6.4. Notice that the proof of lemma 6.4 only relied on the existence of linear functional that represented preferences. From lemma 6.13, such a functional exists which gives us the desired result. (Indeed, all that is to be changed in the proof of lemma 6.4 is to read X_0 for X and $\mathscr{A}_0$ for $\mathscr{A}$.) □

We shall now construct the alter ego's preferences. The construction in §6.1.1 goes through without change. Recall that in the construction of v, we only used the *Upper Semicontinuity* of preferences and *Translation Invariance* for two-element subsets. Repeated application of *AoM* as in §6.1.1 gives us the following lemma.

6.15 Lemma. *Suppose Axioms 1, 2a–c, 3–6 hold. Then there exists a continuous, linear functional $v : \Delta \to \mathbb{R}$ such that (i) $\{\beta\} \succ \{\alpha, \beta\} \succcurlyeq \{\alpha\}$ if and only if $u(\beta) > u(\alpha)$ and $v(\beta) < v(\alpha)$, and (ii) for all $x \in \mathscr{A}_0$, there exists $\widehat{\beta}_x \in x$ such that $U(x) \geqslant u(\widehat{\beta}_x)$ and $u(\widehat{\beta}_x) = \max\limits_{\beta \in B_v(x)} u(\beta)$.*

It follows from *Temptation* that for all $x \in \mathscr{A}_0$, $\max_{\beta \in x} u(\beta) \geqslant U(x) \geqslant u(\widehat{\beta}_x)$. The dual self representation follows immediately giving us ρ_x for each x. The properties of ρ which are required for it to be part of a dual self representation are easily verified.

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-neighbourhood (in Δ) of a point $\beta \in \Delta$ by $N_\varepsilon(\beta)$ and the diameter of a set x by $\operatorname{diam}(x)$. We shall also repeatedly use the fact that ρ must be consistent with the linearity of U, ie (♣) holds.) For any $y \in \mathscr{A}$, $F_y := \operatorname{aff}(y) \cap \Delta$ is a compact convex subset of Δ (where $\operatorname{aff}(y)$ is the affine hull of y in the vector space $\mathscr{M}(Z)$).

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 $\operatorname{aff}(x)$.

6.16 Lemma. *For all $y \in \mathscr{A}_0$ and for all $\bar{\varepsilon} > 0$, there exists $x \in \mathscr{A}_0$ so that $x \subset \operatorname{ri} F_y$, $\rho_y = \rho_x$ and $\operatorname{diam}(x) < \bar{\varepsilon}$.*

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

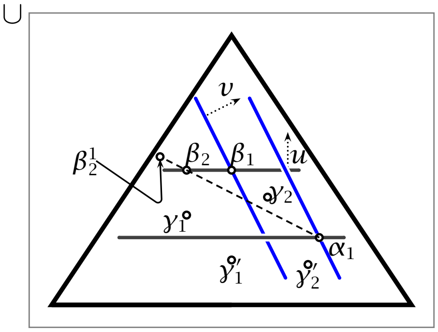
6.17 Lemma. *Let $\succcurlyeq$ have a dual self representation and satisfy* SoM. *Then, for all finite x, for any $\beta \in B_u(x)$ and for any $\alpha \in B_u(B_v(x))$, $x \sim \{\beta, \alpha\}$.*

Proof. Let $\widehat{x} := \{\beta_1, \ldots, \beta_m\} \cup x' \cup \{\alpha_1, \ldots, \alpha_n\} \cup y$ where $\beta_i \in B_u(\widehat{x})$ for $i = 1, \ldots, m$, $\alpha_j \in B_u(B_v(\widehat{x}))$ for $j = 1, \ldots, n$, $u(\beta_1) > u(\gamma) \geqslant u(\alpha_1)$ 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 $v(\alpha_1) \geqslant v(\gamma')$ for all $\gamma' \in y$.)

Recall that $\beta \in \widehat{x}$ is *untempted in* $\widehat{x}$ if there is no $\alpha \in \widehat{x}$ that tempts β. Since α_1 is untempted in $\widehat{x}$ (which means, among other things, that $\{\alpha_1\} \sim \{\alpha_1\} \cup \{\alpha_2, \ldots, \alpha_m\} \cup y$), by *AoM*, $\widehat{x} \sim \{\beta_1, \ldots, \beta_n\} \cup x' \cup \{\alpha_1\}$. Also, by *AoM*, $\widehat{x} \sim \{\beta_1, \ldots, \beta_m\} \cup x' \cup \{\alpha_1\}$. By *SoM*, $\widehat{x} \sim \{\beta_1, \ldots, \beta_m\} \cup \{\alpha_1\}$. Let $x := \{\beta_1, \ldots, \beta_m\} \cup \{\alpha_1\}$. From lemma 6.16, we can assume, without loss of generality, that $x \subset \operatorname{ri} F_x$.

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 $\beta_i^k \in \operatorname{conv}(\{\beta_i^1, \alpha_1\})$ 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 $\{\beta_1, \dots, \beta_i^k, \dots, \beta_m\} \cup \{\alpha_1\} \sim x_i^k$. 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



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

which gives us the desired result. □

6.18 Lemma. *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 6.16, 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, ie suppose $\{\beta\} \nsim \{\beta'\}$. By definition, $\beta' = \beta + c = \beta + (\alpha' - \alpha)$. By *Translation Invariance*, $\{\beta\} + c' \nsim \{\beta'\} + c'$, ie $\{\beta\} + (\alpha - \beta) \nsim \{\beta\} + (\alpha' - \alpha) + (\alpha - \beta)$ which is equivalent to $\{\alpha\} \nsim \{\alpha'\}$ which contradicts the hypothesis.

Hence β, β', α' and α form the vertices of a parallelogram. By Lemma 6.17, $\{\beta, \alpha'\} \sim \{\beta, \beta', \alpha'\} \sim \{\beta', \alpha'\}$. 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 Theorem 3.6 for finite x. >From lemma 6.17, 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 $F_{x\cup y}$, $\varepsilon = \operatorname{diam}(x) \geqslant \operatorname{diam}(y) > 0$ 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 6.18 now gives us the desired result. (The case where $u(\alpha) \leqslant u(\alpha' + c)$ is dealt with in a similar fashion.) □

We can now prove Theorem 3.6 for arbitrary menus.

Proof of Theorem 3.6. Let $x \in \mathscr{A}$, $\beta \in B_u(x)$ and $\alpha \in B_u(B_v(x))$. If $\beta \in B_u(B_v(x))$, then by *Temptation*, we are done. Let us assume this isn't the case.

Consider a sequence (x_k) such that for each k, $x_k \in \mathscr{A}_0$, $x_k \subset x$, $|x_k| < |x_{k+1}|$ and $\lim_k x_k = x$. Define $\alpha_k := \lambda_k\beta + (1 - \lambda_k)\alpha$ for $\lambda_k \in (0, 1)$. We will also require that for each k, $\beta \in x_k$ and $\alpha_k \in B_u(B_v(x_k))$. Then, $x_k \sim \{\beta, \alpha_k\}$ and $U(x_k) = \rho U(\{\beta\}) + (1 - \rho)U(\{\alpha_k\})$. Now, $\lim_k\{\beta, \alpha_k\} = \{\beta, \alpha\}$. Also, for each k, $\{\beta, \alpha_k\} \succ \{\beta, \alpha_{k+1}\}$, ie $x_k \succ x_{k+1}$. From *Upper Semicontinuity*, it follows that $U(x) = \lim_k U(x_k) = \lim_k U(\{\beta, \alpha_k\}) = U(\{\beta, \alpha\})$. □

References

AHN, DAVID (2005). "Ambiguity without a State Space," mimeo, University of California, Berkeley.

ALIPRANTIS, CHARALAMBOS D. AND KIM C. BORDER (1999). *Infinite Dimensional Analysis: A Hitchhiker's Guide*, 2nd ed., Springer, New York.

ANSCOMBE, FRANK J. AND ROBERT J. AUMANN (1963). "A Definition of Subjective Probability," *Annals of Mathematical Statistics*, **34**, 199–205.

BERNHEIM, B. DOUGLAS AND ANTONIO RANGEL (2004). "Addiction and Cue-Triggered Decision Processes," *American Economic Review*, **94(5)**, 1558–1590.

CHATTERJEE, KALYAN AND R. VIJAY KRISHNA (2007). "A Geometric Approach to Continuous Expected Utility," forthcoming, *Economics Letters*.

DEKEL, EDDIE (1986). "An Axiomatic Characterization of Preferences under Uncertainty: Weakening the Independence Axiom," *Journal of Economic Theory*, **40**, 304–318.

DEKEL, EDDIE, BARTON L. LIPMAN AND ALDO RUSTICHINI (2001). "Representing Preferences with a Unique Subjective State Space," *Econometrica*, **69(4)**, 891–934.

DEKEL, EDDIE, BARTON L. LIPMAN AND ALDO RUSTICHINI (2006). "Temptation-Driven Preferences," mimeo, Boston Unversity.

DEKEL, EDDIE, BARTON L. LIPMAN, ALDO RUSTICHINI AND TODD SARVER (2005). "Representing Preferences with a Unique Subjective State Space: Corrigendum," working paper.

ESTEBAN, SUSANNA AND EIICHI MIYAGAWA (2005), "Optimal Menu of Menus with Self-Control Preferences," mimeo, Pennsylvania State University.

ESTEBAN, SUSANNA AND EIICHI MIYAGAWA (2005), "Temptation, Self-Control, and Competitive Nonlinear Pricing," forthcoming, *Economics Letters*.

ESTEBAN, SUSANNA, EIICHI MIYAGAWA AND MATTHEW SHUM (2003), "Nonlinear Pricing with Self-Control Preferences," forthcoming, *Journal of Economic Theory*.

ELIAZ, KFIR AND RANI SPIEGLER (2006). "Contracting with Diversely Naïve Agents," *Review of Economic Studies*, **73**, 689–714.

FUDENBERG, DREW AND DAVID K. LEVINE (2006). "A Dual Self Model of Impulse Control," *American Economic Review*, **96**, 1449–1476.

GAUTHIER, DAVID (1987). *Morals by Agreement*, Oxford University Press.

GUL, FARUK AND WOLFGANG PESENDORFER (2001). "Temptation and Self-Control," *Econometrica*, **69**, 1403–1435.

HAMMOND, PETER J. (1976). "Changing Tastes and Coherent Dynamic Choice," *Review of Economic Studies*, **43**, 159–173.

KREPS, DAVID M. (1979). "A Representation Theorem for 'Preference for Flexibility'," *Econometrica*, **47**, 565–576.

KREPS, DAVID M. (1988). *Notes on the Theory of Choice,* Westview Press, Co., U.S.A.

SAMUELSON, LARRY AND JEROEN SWINKELS (2006). "Information and the Evolution of the Utility Function," *Theoretical Economics*, **1**, 119–142.

SCHELLING, THOMAS C. (1978). "Egonomics, or The Art of Self-management," *American Economic Review*, **68(2)**, 290–294.

SCHELLING, THOMAS C. (1984). "Self-command in Practice, in Policy and in a Theory of Rational Choice," (Richard T. Ely lecture), *American Economic Review*, **74(2)**, 1–11.

STROTZ, ROBERT H. (1955). "Myopia and Inconsistency in Dynamic Utility Maximisation," *Review of Economic Studies*, **23(3)**, 165–180.